181380 Mg

STIC-Biotech/ChemLib

From: CRF
Sent: Liu, Samuel
To: Monday, March 06, 2006 8:49 PM
Subject: STIC-Biotech/ChemLib
10757745

RECEIVED
MAR - 7 2006
(STIC)

Hi, Christina,

Would you please authorize the rush search for the following seqeunces: SEQ ID NO:2 [(1) resdiues 115-121, (2) resdiues 145-153, (3) resdiues 347-121, and (4) resdiues 54-140], SEQ ID NO: 6 (resdiues 2-45), and full-length sequences of SEQ ID NO: 3 (DNA) and SEQ ID NO:2 (polypeptide), against commercial **nucleic acid** and **interference** databases with at least 75 search results?

Very best,

Samuel Liu
AU 1653
Mail Box 3C70
3A64
571-272-0949

seq 2 - 362 AA
3 - 1312 NA
6 - 442 AA

Many thanks Chris

181380

Very best,

Samuel Liu
AU 1653
Mail Box 3C70

571-272-0949

3/13
EIC

****
Vendors and cost where applicabl
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM:_____
WWW/Internet:_____
Other (Specify):_____

1

Vendors
_____ IG
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
✓ Other CGN

Type of Search
_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Total time: _____
Number of Searches: _____
Number of Databases: _____

PTO-1590 (9-90)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2006, 16:04:45 ; Search time 512 Seconds
(without alignments)
5912.448 Million cell updates/sec

Title: US-10-757-745-3
Perfect score: 1312
Sequence: 1 agctattaatgattcgaatt.....tctcaacctttcaggacatc 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7673375 seqs, 1153648444 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	108.4	8.3	606	6	US-09-925-065A-783537	Sequence 783537,
	2	108.4	8.3	634	6	US-09-925-065A-783536	Sequence 783536,
c	3	97.6	7.4	952	8	US-10-750-185-25698	Sequence 25698, A
c	4	97.6	7.4	952	8	US-10-750-623-25698	Sequence 25698, A
	5	55	4.2	703	12	US-11-000-463-28	Sequence 28, Appl
	6	54.2	4.1	814	12	US-11-000-463-123	Sequence 123, App
c	7	53.2	4.1	629	12	US-11-136-527-1281	Sequence 1281, Ap
	8	53	4.0	647	12	US-11-000-463-230	Sequence 230, App
	9	52	4.0	7659	12	US-11-091-883-498	Sequence 498, App
	10	51	3.9	580	12	US-11-000-463-126	Sequence 126, App
	11	51	3.9	2954	12	US-11-241-347-2	Sequence 2, Appli
	12	49.4	3.8	2304	12	US-11-052-554A-538	Sequence 538, App
	13	44.8	3.4	1086	12	US-11-052-554A-553	Sequence 553, App
	14	44	3.4	3038	7	US-10-501-035-194	Sequence 194, App
	15	43.8	3.3	1056	12	US-11-136-527-3921	Sequence 3921, Ap
	16	43	3.3	746	12	US-11-099-691-26	Sequence 26, Appl
	17	42.8	3.3	1598	12	US-11-136-527-2759	Sequence 2759, Ap
	18	42.6	3.2	1713	9	US-11-096-568A-26076	Sequence 26076, A
	19	42.4	3.2	5706	12	US-11-052-554A-519	Sequence 519, App
	20	42.2	3.2	1821	12	US-11-052-554A-547	Sequence 547, App
c	21	42	3.2	172111	12	US-11-121-086-28	Sequence 28, Appl
c	22	41.6	3.2	44617	7	US-10-330-773-928	Sequence 928, App
c	23	41.4	3.2	1658	9	US-11-096-568A-18102	Sequence 18102, A
	24	41.2	3.1	201	8	US-10-995-561-35900	Sequence 35900, A
c	25	41.2	3.1	201	8	US-10-995-561-73926	Sequence 73926, A
	26	41.2	3.1	1015	9	US-11-096-568A-2100	Sequence 2100, Ap
	27	41	3.1	1210	9	US-11-096-568A-326	Sequence 326, App
	28	41	3.1	1854	12	US-11-143-401-103	Sequence 103, App
	29	41	3.1	1857	12	US-11-052-554A-534	Sequence 534, App
c	30	41	3.1	2187	9	US-11-245-147-141	Sequence 141, App
	31	40.8	3.1	1491	9	US-11-096-568A-22316	Sequence 22316, A
	32	40.8	3.1	4983	12	US-11-052-554A-521	Sequence 521, App
	33	40.8	3.1	5679	12	US-11-075-185-36	Sequence 36, Appl
	34	40.8	3.1	78869	12	US-11-075-185-1	Sequence 1, Appli
	35	40.6	3.1	201	8	US-10-995-561-35803	Sequence 35803, A
c	36	40.6	3.1	201	8	US-10-995-561-73911	Sequence 73911, A
	37	40.6	3.1	2486	8	US-10-967-648A-1	Sequence 1, Appli
	38	40.4	3.1	702	6	US-09-925-065A-8726	Sequence 8726, Ap
c	39	40.4	3.1	12309	8	US-10-995-561-13450	Sequence 13450, A
	40	40.4	3.1	95832	8	US-10-995-561-13273	Sequence 13273, A
	41	40	3.0	1699	12	US-11-136-527-1009	Sequence 1009, Ap
	42	39.8	3.0	588	12	US-11-136-527-1488	Sequence 1488, Ap
	43	39.8	3.0	588	12	US-11-136-527-5584	Sequence 5584, Ap
c	44	39.8	3.0	1196	8	US-10-750-185-34334	Sequence 34334, A
c	45	39.8	3.0	1196	8	US-10-750-623-34334	Sequence 34334, A
c	46	39.8	3.0	1931	8	US-10-453-372-1183	Sequence 1183, Ap
	47	39.8	3.0	3990	12	US-11-052-554A-520	Sequence 520, App
c	48	39.8	3.0	4697	8	US-10-453-372-1185	Sequence 1185, Ap
c	49	39.8	3.0	4760	8	US-10-453-372-1179	Sequence 1179, Ap
c	50	39.8	3.0	5688	8	US-10-453-372-1189	Sequence 1189, Ap
c	51	39.8	3.0	6224	8	US-10-453-372-1173	Sequence 1173, Ap
c	52	39.8	3.0	6494	8	US-10-453-372-1187	Sequence 1187, Ap
	53	39.6	3.0	15234	12	US-11-136-527-306	Sequence 306, App
c	54	39.4	3.0	571	6	US-09-925-065A-414314	Sequence 414314,
	55	39.4	3.0	774	8	US-10-632-150-39	Sequence 39, Appl
	56	39.4	3.0	774	12	US-11-073-457-39	Sequence 39, Appl
	57	39.4	3.0	774	12	US-11-073-460-39	Sequence 39, Appl
	58	39.4	3.0	997	9	US-11-096-568A-16503	Sequence 16503, A
	59	39.4	3.0	1278	12	US-11-000-463-95	Sequence 95, Appl
	60	39.4	3.0	1920	12	US-11-052-554A-549	Sequence 549, App
c	61	39.4	3.0	2518	8	US-10-750-185-35012	Sequence 35012, A
c	62	39.4	3.0	2518	8	US-10-750-623-35012	Sequence 35012, A
c	63	39.2	3.0	610	6	US-09-925-065A-869021	Sequence 869021,
c	64	39.2	3.0	1400	12	US-11-136-527-7671	Sequence 7671, Ap
c	65	39.2	3.0	3050	12	US-11-136-527-3575	Sequence 3575, Ap
c	66	39	3.0	619	6	US-09-925-065A-765916	Sequence 765916,
c	67	39	3.0	619	6	US-09-925-065A-832620	Sequence 832620,
c	68	39	3.0	619	6	US-09-925-065A-832621	Sequence 832621,
	69	39	3.0	980	12	US-11-000-463-231	Sequence 231, App
c	70	39	3.0	1653	9	US-11-096-568A-11538	Sequence 11538, A
	71	39	3.0	2196	12	US-11-052-554A-539	Sequence 539, App
	72	39	3.0	2833	7	US-10-501-035-137	Sequence 137, App
	73	39	3.0	2875	12	US-11-000-688-780	Sequence 780, App
	74	39	3.0	4146	12	US-11-052-554A-522	Sequence 522, App
	75	39	3.0	268685	8	US-10-933-025-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24

```
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783537

  Query Match             8.3%;  Score 108.4;  DB 6;  Length 606;
  Best Local Similarity   78.6%;  Pred. No. 2.5e-18;
  Matches  143;  Conservative    0;  Mismatches   36;  Indels    3;  Gaps    1;

Qy        401 GCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCA 460
              |||  |||||| |||||||| || ||  ||| |||| | |||| ||  ||  ||| |||
Db        408 GCCAGTGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCA 467

Qy        461 TCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGAT 517
              ||||  |   ||||  | |||| || ||||||    | ||| ||||||||||||||||||
Db        468 TCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGAT 527

Qy        518 GGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTG 577
              |||||||||  |  ||||||| | ||||||||||||||||||||||| | ||||||||
Db        528 GGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTG 587

Qy        578 TA 579
              ||
Db        588 TA 589


RESULT 2
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783536

  Query Match             8.3%;  Score 108.4;  DB 6;  Length 634;
  Best Local Similarity   78.6%;  Pred. No. 2.6e-18;
  Matches  143;  Conservative    0;  Mismatches   36;  Indels    3;  Gaps    1;

Qy        401 GCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCA 460
              |||  |||||| |||||||| || ||  ||| |||| | |||| ||  ||  ||| |||
Db        259 GCCAGTGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCA 318

Qy        461 TCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGAT 517
              ||||  |   ||||  | |||| || ||||||    | ||| ||||||||||||||||||
Db        319 TCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGAT 378

Qy        518 GGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTG 577
              |||||||||  |  ||||||| | ||||||||||||||||||||||| | ||||||||
Db        379 GGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTG 438

Qy        578 TA 579
              ||
Db        439 TA 440


RESULT 3
US-10-750-185-25698/c
; Sequence 25698, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25698
;   LENGTH: 952
;   TYPE: DNA
;   ORGANISM: Bovine  19866881510436
US-10-750-185-25698

  Query Match             7.4%;  Score 97.6;  DB 8;  Length 952;
  Best Local Similarity   82.4%;  Pred. No. 2.1e-15;
  Matches  112;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy        658 AATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGT 717
              |||  |  ||||  |||||||||||||||||||||||| | |||||||||  ||||||
Db        798 AATACCTGAAGGTCGTGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGT 739

Qy        718 GAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCT 777
              ||| |||||||| || || |||||||||||||||||||  ||||||||||||||  |
Db        738 GAAGTTTAAAAGCCAAGAAATTATTCCTTTTCCAAATACGCAAATGATGAGAAACCTTTT 679

Qy        778 ATGCGTAAATGTGAGT 793
              || ||  |||| |||
Db        678 GTGTGTTCATGTAAGT 663


RESULT 4
US-10-750-623-25698/c
; Sequence 25698, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
```

```
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25698
;    LENGTH: 952
;    TYPE: DNA
;    ORGANISM: Bovine  19866881510436
US-10-750-623-25698

  Query Match             7.4%;  Score 97.6;  DB 8;  Length 952;
  Best Local Similarity   82.4%;  Pred. No. 2.1e-15;
  Matches  112;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy        658 AATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGT 717
              |||   |  ||||  |||||||||||||||||||||||| | |||||||||  ||||||
Db        798 AATACCTGAAGGTCGTGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGT 739

Qy        718 GAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCT 777
              ||| ||||||||  || || ||||||||||||||||||||  ||||||||||||||  |
Db        738 GAAGTTTAAAAGCCAAGAAATTATTCCTTTTCCAAATACGCAAATGATGAGAAACCTTTT 679

Qy        778 ATGCGTAAATGTGAGT 793
              ||  ||  |||| |||
Db        678 GTGTGTTCATGTAAGT 663


RESULT 5
US-11-000-463-28
; Sequence 28, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Cao, Yi-Cheng
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 785CIP4CN
;  CURRENT APPLICATION NUMBER: US/11/000,463
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: 10/291,265
;  PRIOR FILING DATE: 2002-11-08
;  PRIOR APPLICATION NUMBER: PCT/US01/02623
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 09/617,746
;  PRIOR FILING DATE: 2000-07-17
;  PRIOR APPLICATION NUMBER: 09/631,451
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 09/633,870
;  PRIOR FILING DATE: 2000-09-15
;  NUMBER OF SEQ ID NOS: 944
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 28
;    LENGTH: 703
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (285)..(509)
US-11-000-463-28

  Query Match             4.2%;  Score 55;  DB 12;  Length 703;
  Best Local Similarity   100.0%;  Pred. No. 0.00027;
  Matches   55;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         22 ATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 55


RESULT 6
US-11-000-463-123
; Sequence 123, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Cao, Yi-Cheng
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 785CIP4CN
;  CURRENT APPLICATION NUMBER: US/11/000,463
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: 10/291,265
;  PRIOR FILING DATE: 2002-11-08
;  PRIOR APPLICATION NUMBER: PCT/US01/02623
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 09/617,746
;  PRIOR FILING DATE: 2000-07-17
;  PRIOR APPLICATION NUMBER: 09/631,451
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 09/633,870
;  PRIOR FILING DATE: 2000-09-15
;  NUMBER OF SEQ ID NOS: 944
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 123
;    LENGTH: 814
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (124)..(324)
US-11-000-463-123

  Query Match             4.1%;  Score 54.2;  DB 12;  Length 814;
  Best Local Similarity   94.9%;  Pred. No. 0.00047;
  Matches   56;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         15 CGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAG 73
              | | || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 CAATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAG 69


RESULT 7
US-11-136-527-1281/c
; Sequence 1281, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
```

```
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1281
;   LENGTH: 629
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-1281

  Query Match             4.1%;  Score 53.2;  DB 12;  Length 629;
  Best Local Similarity   94.8%;  Pred. No. 0.00076;
  Matches   55;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              || || ||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        629 TTAATCCGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGTC 572


RESULT 8
US-11-000-463-230
; Sequence 230, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Cao, Yi-Cheng
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 785CIP4CN
;  CURRENT APPLICATION NUMBER: US/11/000,463
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: 10/291,265
;  PRIOR FILING DATE: 2002-11-08
;  PRIOR APPLICATION NUMBER: PCT/US01/02623
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 09/617,746
;  PRIOR FILING DATE: 2000-07-17
;  PRIOR APPLICATION NUMBER: 09/631,451
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 09/633,870
;  PRIOR FILING DATE: 2000-09-15
;  NUMBER OF SEQ ID NOS: 944
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 230
;   LENGTH: 647
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (192)..(506)
US-11-000-463-230

  Query Match             4.0%;  Score 53;  DB 12;  Length 647;
  Best Local Similarity   100.0%;  Pred. No. 0.00087;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         22 ATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 53


RESULT 9
US-11-091-883-498
; Sequence 498, Application US/11091883
; Publication No. US20060024693A1
; GENERAL INFORMATION:
;  APPLICANT: CIBELLI, JOSE
;  APPLICANT:  FERNANDEZ, EMILIO O.
;  APPLICANT:  JORDAO DE MEGALHAES, GUILHERME
;  APPLICANT:  KOCABAS, ARIF
;  APPLICANT:  CROSBY, JAVIER A.
;  TITLE OF INVENTION: IDENTIFICATION OF GENES OR POLYPEPTIDES THE EXPRESSION OF WHICH
;  TITLE OF INVENTION:  CORRELATES TO FERTILITY, OVARIAN FUNCTION AND/OR FETAL/NEWBORN
;  TITLE OF INVENTION:  VIABILITY
;  FILE REFERENCE: 53942US
;  CURRENT APPLICATION NUMBER: US/11/091,883
;  CURRENT FILING DATE:  2005-03-29
;  PRIOR APPLICATION NUMBER: 60/556,875
;  PRIOR FILING DATE: 2004-03-29
;  NUMBER OF SEQ ID NOS: 513
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 498
;   LENGTH: 7659
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (518)..(518)
;   OTHER INFORMATION: a, c, g, or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (656)..(661)
;   OTHER INFORMATION: a, c, g, or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1026)..(1026)
;   OTHER INFORMATION: a, c, g, or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1102)..(1102)
;   OTHER INFORMATION: a, c, g, or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1179)..(1179)
;   OTHER INFORMATION: a, c, g, or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1282)..(1282)
;   OTHER INFORMATION: a, c, g, or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1342)..(1343)
;   OTHER INFORMATION: a, c, g, or t
US-11-091-883-498

  Query Match             4.0%;  Score 52;  DB 12;  Length 7659;
  Best Local Similarity   85.3%;  Pred. No. 0.0046;
  Matches   58;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || |||||||||| |||||||| ||||||||| ||| ||||||||||||||||||
Db         68 ATTTAATACGACTCATTATAGGGATTTTGGCCCTGGAGCCCAAGAATTCGGCACGAGGGT 127

Qy         77 GGGAAGCA 84
              |||   ||
Db        128 GGGTGTCA 135


RESULT 10
US-11-000-463-126
; Sequence 126, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y Tom
```

```
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Cao, Yi-Cheng
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 785CIP4CN
;  CURRENT APPLICATION NUMBER: US/11/000,463
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: 10/291,265
;  PRIOR FILING DATE: 2002-11-08
;  PRIOR APPLICATION NUMBER: PCT/US01/02623
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 09/617,746
;  PRIOR FILING DATE: 2000-07-17
;  PRIOR APPLICATION NUMBER: 09/631,451
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 09/633,870
;  PRIOR FILING DATE: 2000-09-15
;  NUMBER OF SEQ ID NOS: 944
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 126
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (163)..(318)
US-11-000-463-126

  Query Match             3.9%;  Score 51;  DB 12;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 0.0028;
  Matches   51;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         23 TACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAG 73
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAG 51


RESULT 11
US-11-241-347-2
; Sequence 2, Application US/11241347
; Publication No. US20060024665A1
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: PRMTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
;  FILE REFERENCE: EX02-096
;  CURRENT APPLICATION NUMBER: US/11/241,347
;  CURRENT FILING DATE:  2005-09-30
;  PRIOR APPLICATION NUMBER: US 60/296,076
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/328,605
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/338,733
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/357,253
;  PRIOR FILING DATE: 2002-02-15
;  PRIOR APPLICATION NUMBER: US 60/357,600
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 2954
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-241-347-2

  Query Match             3.9%;  Score 51;  DB 12;  Length 2954;
  Best Local Similarity   69.7%;  Pred. No. 0.0056;
  Matches   69;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy        147 CGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGG 206
              |||||| | | || | | |||| | |  ||||||| ||||| | |  |||||   |||||
Db          1 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCAGCGGCGGCGG 60

Qy        207 AGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
                  || |||||  | || |||||||||  | | ||||||
Db         61 CCTGGGCCCGGGCGCAGCGGCGGCGGCGGCGGGGCCTGG 99


RESULT 12
US-11-052-554A-538
; Sequence 538, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 538
;   LENGTH: 2304
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-538

  Query Match             3.8%;  Score 49.4;  DB 12;  Length 2304;
  Best Local Similarity   53.3%;  Pred. No. 0.013;
  Matches  104;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy         71 GAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCT 130
              | || |  |  |  ||  |  | |||||||  |     | |   |  ||| ||  ||| ||
Db        683 GGGGCCACGGGGGTGCCGGCGGGGCGGGTGGTGCCGGCGTGGCTGGCGCCGGCGGCTTCG 742

Qy        131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
              |  |  | ||  |   ||||||  ||| || | | ||||  | ||  |||||  |||||||
Db        743 AGGGGACGATCGGTGCCGGCGGTGCCGGCGGTGTCGGCGGTGCCGGCGGTGTCGGCGGTG 802

Qy        191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
                 ||  |  | |||||  |       || || |  |  || | |  |    |||  |
Db        803 CCGGCGGTGCCGGCGGGTGGCTGTACGGCGACGCCGGTGCCGGTGGGGATGGTGGTGTCG 862

Qy        251 GCGTTGGTGGGGGGA 265
              |||  ||  || || |
Db        863 GCGGTGCCGGCGGCA 877


RESULT 13
US-11-052-554A-553
; Sequence 553, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
```

```
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 553
;    LENGTH: 1086
;    TYPE: DNA
;    ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-553

  Query Match             3.4%;  Score 44.8;  DB 12;  Length 1086;
  Best Local Similarity   53.1%;  Pred. No. 0.15;
  Matches  119;  Conservative    0;  Mismatches  102;  Indels    3;  Gaps    1;

Qy          43 TTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTG-- 100
               ||| | ||| | ||   || |||| | ||| | |  | ||  | | || || | |
Db         361 TTGATCGGCGACGGCACCGATGCGGCGCCGGGGAGCGGGGGGGCCGGCGGGGCCGGCGGC 420

Qy         101 -TTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGC 159
                | |||| ||   | ||||    |||| |      | ||     |   |||   | ||
Db         421 ATCTTGATCGGCAACGGCGGGACCGGCGGCGCCAGCGACTTAGCCGGGACCGGCCGCGGC 480

Qy         160 AGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGT 219
               ||| |||| ||||||||||||| |||||  |    | |||  ||||| |  ||   |||
Db         481 GGGGTCGGCGGGGCGGGCGGCGCCGGCGGGCTCTTCGGCATCGGCGGCGCCGGCGGGGGC 540

Qy         220 GAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGG 263
                 | | ||||   |  |  ||    |||  | |||  ||
Db         541 TGCGGGTCCGCGGTGGCGATCGGGGGTGACGGCGGGGCTGGTGG 584


RESULT 14
US-10-501-035-194
; Sequence 194, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
;  TITLE OF INVENTION:  ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
;  TITLE OF INVENTION:  AND/OR PROTEIN TYROSINE KINASE PATHWAYS
;  FILE REFERENCE: D0185 PCT
;  CURRENT APPLICATION NUMBER: US/10/501,035
;  CURRENT FILING DATE:  2004-07-09
;  PRIOR APPLICATION NUMBER: US 60/350,061
;  PRIOR FILING DATE: 2002-01-18
;  NUMBER OF SEQ ID NOS: 795
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 194
;    LENGTH: 3038
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-501-035-194

  Query Match             3.4%;  Score 44;  DB 7;  Length 3038;
  Best Local Similarity   46.7%;  Pred. No. 0.38;
  Matches  140;  Conservative    0;  Mismatches  160;  Indels    0;  Gaps    0;

Qy          79 GAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAG 138
               | ||  |  || |||||| | |  | | ||||  | || |  ||  |  |  ||  |
Db        1779 GGAGGCGGCGGAGGAGCGAGAGAGTGCAAGGGAGGAGGAGGAGGCGGAGCAGGAGAGACG 1838

Qy         139 TTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGC 198
                  ||  | || |  |  |    ||  | |  ||||||    |||| | || ||||
Db        1839 CGGCGGGGAGGAGAGGGTGGGGGAAGAGGATGAGGAGGCGGCCGAGGCGGCGGAGGCAGA 1898

Qy         199 ACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGT 258
                 ||  |||| |||  ||| |  |  |     ||| | || |  ||      | |   |
Db        1899 GGCGGACGAGGCGGAGAGGGCGCGGCAGAACGCGCTCCTGTTCGCGGAGAACGACTGGCAGACGCA 1958

Qy         259 GGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCA 318
               ||  |   ||| |       | |  | ||   |   |  |  |  | || ||| || | |
Db        1959 GGAAGCCGGCGCCGAGGACAAGCGCTCCCAGGAGGAGACGCCGGGCCACCGGCGGAAGGA 2018

Qy         319 GAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCC 378
               |   |   ||  |   |  || |  |  ||| ||||| | | |  |   |  ||| |
Db        2019 GGCCGAGGGGACAGAGGAGGGCGGGGAGGAGGAGGACGACGAGGAGATGGATCCGCAGAC 2078


RESULT 15
US-11-136-527-3921
; Sequence 3921, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3921
;    LENGTH: 1056
;    TYPE: DNA
;    ORGANISM: Rattus norvegicus
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(1)
;    OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-3921

  Query Match             3.3%;  Score 43.8;  DB 12;  Length 1056;
  Best Local Similarity   87.3%;  Pred. No. 0.28;
  Matches   48;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          33 TATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCG 87
               ||| |||||||||||||||||||||||||||| | |||||||  ||| || |||
Db          37 TATTGGGAATTTGGCCCTCGAGGCCAAGAATTCAGGACGAGGGGCGGAGGCGGCG 91


RESULT 16
US-11-099-691-26
; Sequence 26, Application US/11099691
; Publication No. US20050260644A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE PHARMACEUTICALS, INC.
;  APPLICANT:  BANDMAN, Olga
;  APPLICANT:  HILLMAN, Jennifer L.
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  YUE, Henry
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  YANG, Junming
;  TITLE OF INVENTION: CELL SIGNALING PROTEINS
;  FILE REFERENCE: PF-0521 PCT
;  CURRENT APPLICATION NUMBER: US/11/099,691
;  CURRENT FILING DATE:  2005-04-06
;  PRIOR APPLICATION NUMBER: US/09/700,444
;  PRIOR FILING DATE: 2002-08-26
;  PRIOR APPLICATION NUMBER: 60/085,343
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: 60/098,010
;  PRIOR FILING DATE: 1998-08-26
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PERL Program
; SEQ ID NO 26
;    LENGTH: 746
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc-feature
;    OTHER INFORMATION: Incyte Clone 3315936
US-11-099-691-26
```

Query Match 3.3%; Score 43; DB 12; Length 746;
Best Local Similarity 83.1%; Pred. No. 0.39;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||| |||||||||||||||||||||||||||||||    ||| |  ||||||||||
Db          1 ATTTAATATGACTCACTATAGGGAATTTGGCCCTCGAGCTAGAGATTCGGGCACGAGGG 59
```

RESULT 17
US-11-136-527-2759
; Sequence 2759, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2759
; LENGTH: 1598
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2759

Query Match 3.3%; Score 42.8; DB 12; Length 1598;
Best Local Similarity 57.5%; Pred. No. 0.6;
Matches 77; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

```
Qy          71 GAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCT 130
               | ||  ||    |  ||| |  | |  || |    | | |||  | ||||||  |||| |
Db         464 GCGGCGGGAGCGGTGCGGGCGGCGGCGGCGGCGCGGGAGGAGGCGGCGGCGGCGGCGCG 523

Qy         131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
               || | |       || ||||||| ||  || | || ||| | |  ||||||| ||||| |
Db         524 GGAACCGTGCGGGAGGCGGCGGTGCTGGCGGAGCAGGCGGCGGCGGCGGCGGCGGCGCG 583

Qy         191 GAAGCAGCACAGGC 204
               |  ||  ||  | |
Db         584 GCGGCTCCAGGGCC 597
```

RESULT 18
US-11-096-568A-26076
; Sequence 26076, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26076
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1713)
; OTHER INFORMATION: Ceres Seq. ID no. 13498435
US-11-096-568A-26076

Query Match 3.2%; Score 42.6; DB 9; Length 1713;
Best Local Similarity 54.0%; Pred. No. 0.7;
Matches 87; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```
Qy         168 CAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGC 227
               | |||  ||||||| || ||    |  |      | |||||  |||     |    |   |
Db         312 CCGGGACGGCGGCGCCGTCGTGGGCGGGGTTCCTGGCGGACCAGGCGTCGCGTGACGCGT 371

Qy         228 GGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCA 287
               ||||||  | ||  ||||| ||     |  ||||  | ||  ||| || | ||   | ||||
Db         372 CGCGGCTGCTGTACCTGGACTCGCTGGCGGTGCGTGGCCGCGCCCGCGCGTACGCGCCCA 431

Qy         288 GCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTG 328
               ||    | ||  | |     |||||||| ||    | ||
Db         432 TCGCGTCGGGGCGGCAGCTGCTGCAGACGCTGACGTACGTG 472
```

RESULT 19
US-11-052-554A-519
; Sequence 519, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 519
; LENGTH: 5706
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-519

Query Match 3.2%; Score 42.4; DB 12; Length 5706;
Best Local Similarity 60.3%; Pred. No. 1.3;
Matches 70; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```
Qy          92 GAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCG 151
               | || || |       | ||   | ||||||| |||| | |           | | | |||||
Db        4091 GGGCCGGCGGGGCCGGCGGCAACGGCGGCGACGGCGACGGCTTTATCGGCGGGTCCGGCG 4150

Qy         152 GAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGA 207
               |  || || || ||||| | |  | |||||||||||| ||    || ||| ||||||
Db        4151 GCACCGGCGGGACCGGCGGCGACGCCGGCGTCGGCGGCCTGGCCAACACCGGCGGA 4206
```

RESULT 20
US-11-052-554A-547
; Sequence 547, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 547
; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-547

Query Match 3.2%; Score 42.2; DB 12; Length 1821;
Best Local Similarity 48.9%; Pred. No. 0.92;
Matches 113; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

```
Qy        122 ATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCG 181
              ||||| | |      |   | | | |||||| | | ||   |||||| | || ||||||
Db        416 ATGGCAACGGCGGCAATGGCGGGTCCGGCGGGGTCGGCCAAGCCGGCGGAGCCGGCGGCA 475

Qy        182 TCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGC 241
               || ||     | | |||  |||||     ||   ||  | || ||| |  ||     |
Db        476 GCGCCGGGTTGATCGGCATCGGCGGGACCGGGGGGGCCGGAGGGGCCGGCGCGGTCGGCG 535

Qy        242 CTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
                |      ||   || ||| || |      | | |  ||  ||    | |  | |||
Db        536 GCGTCGGCGGCAACGGGGGGTGGTTATACGGCAATGGCGGGGCCGGCGGGCTCGGGGGGA 595

Qy        302 AACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAG 352
                | |||| | ||   |  ||  |  | ||  |    ||| |||| | |
Db        596 CCGGGGTGGCCGGCGTCAATGGCGGTATGGGCGCAGCAGGAGGTGCCGGCG 646
```

RESULT 21
US-11-121-086-28/c
; Sequence 28, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28
; LENGTH: 172111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-28

Query Match 3.2%; Score 42; DB 12; Length 172111;
Best Local Similarity 50.8%; Pred. No. 7.1;
Matches 127; Conservative 0; Mismatches 120; Indels 3; Gaps 1;

```
Qy         85 GCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGA 144
              ||  |  | ||||  |    |  || |  | |||| |||  | | |  ||   ||
Db      34878 GCCGGCCGGGCGGCCGGGAGGCACGGTCTGGTGGACGATCGCGGGTCGGACGAGCGGCGC 34819

Qy        145 TGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGC 204
               ||| | | ||||||| | |||||| || |||||| | | |    |  ||       ||
Db      34818 AAGGGCTGGGGCCGCAGGCCGGGCAGGACGCGCGGCGCCCGGG---GCTGCGCGGGGCGC 34762

Qy        205 GGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGG 264
              || |  || ||||| |    |  |||||   | | ||   ||  || |    |   | |
Db      34761 GGGGCCGGGCCGGGCGGGCGGCGGGCGGGGGCGGGGCGGCGGAATTCCCCGGCGCCGCCG 34702

Qy        265 ATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGC 324
               | | | |||||  || | |  | | |    ||| ||||  |||  ||| | |
Db      34701 GCCTGCGCGCCCATTGGTCTGCACCCCGCCGTGACGTCACGCCTGGGGAACGGGAAAACC 34642

Qy        325 CCTGAGCGCC 334
              ||   ||| |
Db      34641 CCCACGCGTC 34632
```

RESULT 22
US-10-330-773-928/c
; Sequence 928, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 928
; LENGTH: 44617
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(44617)
; OTHER INFORMATION: n = A,T,C or G
US-10-330-773-928

Query Match 3.2%; Score 41.6; DB 7; Length 44617;
Best Local Similarity 55.6%; Pred. No. 5.1;
Matches 80; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

```
Qy        674 GAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAG 733
              |||||| |  ||  |     | ||||| |  ||| |  | | || || | |||  ||
Db      29413 GAAGAAAAAACTTAGAACAAGTGAATATTGTATGAAGCATTATTAAACTTCACAAGATAG 29354

Qy        734 GAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT 793
                | ||||  | | || | || |  | ||    |   ||||  | |||| | |||| |
Db      29353 CAAATTAAACATCTTTCTAAAAAGACAAATTTTCTAAACAGGTTATGTGGAAATATTCTA 29294

Qy        794 TTGGGTGGAAATGAATTTTGCCTT 817
              |   | ||  ||||| |||| | |
Db      29293 TCTAGGGGGCATGAAATTTGACCT 29270
```

RESULT 23
US-11-096-568A-18102/c
; Sequence 18102, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 18102
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1658)
; OTHER INFORMATION: Ceres Seq. ID no. 12362642
US-11-096-568A-18102

Query Match 3.2%; Score 41.4; DB 9; Length 1658;
Best Local Similarity 54.2%; Pred. No. 1.4;
Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
Qy        149 GCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAG 208
              | |||| | ||    |||||   || |||||||  || |    | |  ||   || || |
Db        379 GAGGAGTCGGCGACTCCGGCGCCGCCGGCGGCGGTGGGGGAACAGGAGGCGGCGGAGGCG 320

Qy        209 GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGC 268
              |  |  ||| || ||||| ||||     | |  ||||     || |||||  |||  ||
Db        319 GCCCAGTGGGCGATGAGGCTGCGGAGAGTGAGGTTGGGGACGAGGTCGGTGGAGGGGAGC 260

Qy        269 GACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAA 303
              |    || | ||| |     |||  ||  |  | |
Db        259 GGGAGCATGGTGGCCGGGCACGTGGTGTTGCCGGA 225
```

```
RESULT 24
US-10-995-561-35900
; Sequence 35900, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35900
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-35900

  Query Match             3.1%;  Score 41.2;  DB 8;  Length 201;
  Best Local Similarity   52.3%;  Pred. No. 0.66;
  Matches   91;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy         38 GGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGG 97
              ||    | || |   ||| | | ||   |    | |||||| ||  || ||| | || |  |
Db          2 GGGTGTAGGGCGGGGAGCCAAGGATCTGACCGCGAGGGGCGGGCGCGGCGGGGAGGGGCG 61

Qy         98 GTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCC 157
              | |    | |||   |  | |||          || || |  | |  | | |||||||
Db         62 GGGCGGGGCGGGCGGCGCGGGGCCTATTAAAGGCGCGGCSGGGCAGCGGGGCCGGAGCGG 121

Qy        158 GCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAG 211
              | |||||   | || | ||| || | ||||  |  || | |   |||| || |
Db        122 CCGGGGCCACGATGGAGCGCGACGGCTGCGCGGGGGGCGGGAGCCGCGGCGGCG 175


RESULT 25
US-10-995-561-73926/c
; Sequence 73926, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73926
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-73926

  Query Match             3.1%;  Score 41.2;  DB 8;  Length 201;
  Best Local Similarity   52.3%;  Pred. No. 0.66;
  Matches   91;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy         38 GGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGG 97
              ||    | || |   ||| | | ||   |    | |||||| ||  || ||| | || |  |
Db        200 GGGTGTAGGGCGGGGAGCCAAGGATCTGACCGCGAGGGGCGGGCGCGGCGGGGAGGGGCG 141

Qy         98 GTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCC 157
              | |    | |||   |  | |||          || || |  | |  | | |||||||
Db        140 GGGCGGGGCGGGCGGCGCGGGGCCTATTAAAGGCGCGGCSGGGCAGCGGGGCCGGAGCGG 81

Qy        158 GCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAG 211
              | |||||   | || | ||| || | ||||  |  || | |   |||| || |
Db         80 CCGGGGCCACGATGGAGCGCGACGGCTGCGCGGGGGGCGGGAGCCGCGGCGGCG 27


RESULT 26
US-11-096-568A-2100
; Sequence 2100, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 2100
;   LENGTH: 1015
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1015)
;   OTHER INFORMATION: Ceres Seq. ID no. 15180832
US-11-096-568A-2100

  Query Match             3.1%;  Score 41.2;  DB 9;  Length 1015;
  Best Local Similarity   52.3%;  Pred. No. 1.3;
  Matches   91;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy        146 GCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCG 205
              |||| | || | || |     |  |  ||||||  | ||  ||   | || |  | |||
Db        279 GCGGAGAAGGCGGCGGCCGACGACGCCCGGGCGCAGGCGCGGCGGCAGGCGGAGCTGGCC 338

Qy        206 GAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGA 265
              ||| |||| | || |      |||||  | | |   | ||     |||  | ||||  |
Db        339 GAGCAGGAGCTGGCGGCCGCCCGGCGCATGCGGCAGCAGGCGCAGGCGGAGCTGGGCCGC 398

Qy        266 TGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAG 319
                || |      || || |  ||| | || || | ||| |  || |  ||||
Db        399 GCGCACGCGCTCCGGGACCACGCCGTGCGGCAGGTGGACGCCACGCTGCTGCAG 452


RESULT 27
US-11-096-568A-326
; Sequence 326, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 326
;   LENGTH: 1210
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1210)
;   OTHER INFORMATION: Ceres Seq. ID no. 15179510
US-11-096-568A-326

  Query Match             3.1%;  Score 41;  DB 9;  Length 1210;
  Best Local Similarity   50.8%;  Pred. No. 1.6;
  Matches   98;  Conservative    0;  Mismatches   95;  Indels    0;  Gaps    0;

Qy        109 GGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGC 168
              | | ||  || ||  |||  |  || || |  ||| |  ||  || |  ||  | ||||
Db        400 GCAGCCGCCGCCGCCGGCTACGTGCTGCGGCGCCGCTTTGGACGACCGGGCGCCGGCGGC 459

Qy        169 AGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCG 228
              |   ||||| ||||| || ||||  | |  |||| || || ||  |  |         |
Db        460 GGCATGGGCGTCGTCGCCGTCCGAGCCTGACCAGGTGGCGGCGGCGCCTGCTGCTCAGGA 519
```

```
Qy        229 GCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAG 288
              | | |  ||  | | |  |||  | |  ||    |||  |  |   ||    | |
Db        520 GGCGGAGAAGAAGCAGAACAGCGCGACGACGACGGAGGCCGAAGCAGCAGCGGGAAGCGG 579

Qy        289 CGTCCTGCGGGAG 301
              || |  |||||||
Db        580 CGCCGGGCGGGAG 592
```

RESULT 28
US-11-143-401-103
; Sequence 103, Application US/11143401
; Publication No. US20060002953A1
; GENERAL INFORMATION:
; APPLICANT: Behr, Marcel
; APPLICANT: Small, Peter
; APPLICANT: Schoolnik, Gary
; APPLICANT: Wilson, Michael A.
; TITLE OF INVENTION: Molecular Differences Between Species of
; TITLE OF INVENTION: the M. Tuberculosis Complex
; FILE REFERENCE: STAN102CON
; CURRENT APPLICATION NUMBER: US/11/143,401
; CURRENT FILING DATE: 2005-06-01
; PRIOR APPLICATION NUMBER: US/10/647,089
; PRIOR FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: US/09/894,844
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/318,191
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/097,936
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 1854
; TYPE: DNA
; ORGANISM: Mycobacteria tuberculosis
US-11-143-401-103

Query Match 3.1%; Score 41; DB 12; Length 1854;
Best Local Similarity 50.2%; Pred. No. 1.9;
Matches 101; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
Qy         92 GAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCG 151
              |||| || |       | |   || |||| |  |||  | | |          | | |||||
Db        962 GAGCCGGCGGGGCAGGCGCTGCCGGCGGAGCCGGCACCGGCGACACCGAGAGCTTCGGCG 1021

Qy        152 GAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAG 211
              |   | || |||||||| | |  |||||||  |  | |   | | | |  ||||| |  |
Db       1022 GCCACGGCGGGGCCGGCGGTGATGGCGGCGCTGTTGGCTTGATCGGTAACGGCGGGGCCG 1081

Qy        212 GACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGAC 271
              |      | |   | |  |||   |||  |||     |||  |  || ||  || |
Db       1082 GCGGCACCGGATCTCCCGGCGCTGTGGTGGGTGGTAACGGCGGCGTCGGTGGTCTGGGTG 1141

Qy        272 CCCACGATGGTCCCCAGCGTC 292
              |  ||   |||||  | |||
Db       1142 GCGCCGGCAGTCCCGGGGGTC 1162
```

RESULT 29
US-11-052-554A-534
; Sequence 534, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 534
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-534

Query Match 3.1%; Score 41; DB 12; Length 1857;
Best Local Similarity 50.2%; Pred. No. 1.9;
Matches 101; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
Qy         92 GAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCG 151
              |||| || |       | |   || |||| |  |||  | | |          | | |||||
Db        962 GAGCCGGCGGGGCAGGCGCTGCCGGCGGAGCCGGCACCGGCGACACCGAGAGCTTCGGCG 1021

Qy        152 GAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAG 211
              |   | || |||||||| | |  |||||||  |  | |   | | | |  ||||| |  |
Db       1022 GCCACGGCGGGGCCGGCGGTGATGGCGGCGCTGTTGGCTTGATCGGTAACGGCGGGGCCG 1081

Qy        212 GACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGAC 271
              |      | |   | |  |||   |||  |||     |||  |  || ||  || |
Db       1082 GCGGCACCGGATCTCCCGGCGCTGTGGTGGGTGGTAACGGCGGCGTCGGTGGTCTGGGTG 1141

Qy        272 CCCACGATGGTCCCCAGCGTC 292
              |  ||   |||||  | |||
Db       1142 GCGCCGGCAGTCCCGGGGGTC 1162
```

RESULT 30
US-11-245-147-141/c
; Sequence 141, Application US/11245147
; Publication No. US20060030541A1
; GENERAL INFORMATION:
; APPLICANT: GARCIA, TERESA
; APPLICANT: ROMAN ROMAN, SERGIO
; APPLICANT: BARON, ROLAND
; APPLICANT: CALL, KATHERINE
; APPLICANT: THEILHABER, JOACHIM
; APPLICANT: CONNOLLY, TIMOTHY
; APPLICANT: JACKSON, AMANDA
; APPLICANT: BUSHNELL, STEVEN
; APPLICANT: RAWADI, GEORGES
; TITLE OF INVENTION: GENES INVOLVED IN OSTEOGENESIS, AND METHODS OF USE
; FILE REFERENCE: 37991-0023
; CURRENT APPLICATION NUMBER: US/11/245,147
; CURRENT FILING DATE: 2005-10-07
; PRIOR APPLICATION NUMBER: PCT/IB02/02211
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/281,400
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 141
; LENGTH: 2187
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-245-147-141

Query Match 3.1%; Score 41; DB 9; Length 2187;
Best Local Similarity 47.5%; Pred. No. 2;
Matches 122; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

```
Qy         46 GCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTG 105
              |||  ||  |||     |   ||| |  | || |  |  ||  |  ||   || |
Db        354 GCCGCCGCCGCCCGCGCTCTTGCAGGCAGCCGAGGGGGCGCTGGCCGAATTGGAGGAGGA 295

Qy        106 AGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCC 165
              | ||||  |||  | ||||  | |  |      || | ||| || |  | | ||| |
Db        294 CGAGGACGAGGCGCAGGAGGCGGAGGACGACGACGAGGAGGAGGAGGTGGTCTCCGGGGC 235
```

```
Qy        166 GGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAG 225
              ||| | |  ||||||| ||||| | |  |  |  |       || || | ||      |
Db        234 GGCGGCGGCGGCGGCGGCGGCGGGCGGCGGGCTCATCAGGGCGGCCTGGAGCGCGCC 175

Qy        226 GCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCC 285
              | |||      | |  |  | |          ||  || ||  ||   ||| || |||||
Db        174 GGGGACCGGGCTGCACGCGCGGCGGAGCGGGGCCGGCGGCGGCCCGCCCTCGGTGGTCAT 115

Qy        286 CAGCGTCCTGCGGGAGA 302
              | | | || | | | ||
Db        114 CTGGGACCCGGGAGCGA 98


RESULT 31
US-11-096-568A-22316
; Sequence 22316, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 22316
;   LENGTH: 1491
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1491)
;   OTHER INFORMATION: Ceres Seq. ID no. 12408460
US-11-096-568A-22316

  Query Match             3.1%;  Score 40.8;  DB 9;  Length 1491;
  Best Local Similarity   53.0%;  Pred. No. 2;
  Matches   87;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy        154 GCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGA 213
              | | ||  |  |   |||   ||| ||| ||||| | || | ||   |   || | ||||
Db        134 GACTGCCTGTACATGAGGAGCGGCTGCGGCGGCGGCGGAGGGAGACAACCTGGGGAAGGA 193

Qy        214 CCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCC 273
              || | ||  || | |||| | | ||   | |  || | ||||||| ||     |   |
Db        194 GGCGGGGGAGGCGCTGGGGCTGCTGGGCTGCGTCAATGAGGTGGTGGCGGCGGTGGCGGC 253

Qy        274 CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGC 317
              ||||  ||  ||||    |   | | |  || |  |  | |||
Db        254 CACGTGGGGGCCCATGACCGACGGCGCGCGCGTCGAGACGGCGC 297


RESULT 32
US-11-052-554A-521
; Sequence 521, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 521
;   LENGTH: 4983
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-521

  Query Match             3.1%;  Score 40.8;  DB 12;  Length 4983;
  Best Local Similarity   50.0%;  Pred. No. 3.3;
  Matches  102;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy         66 GGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGG 125
              ||   | |||   |  |  ||   |    ||||      |  |  | | | ||  || |||
Db       3288 GGACGGTGGGGTCGGCGGGGCCGCACCTTCGGGGACGGTCGGCAGTCACGGCACCGGTGG 3347

Qy        126 CGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGG 185
              |||  |      |   | ||  |||||| ||| || | |  ||| |||| ||||||  |||
Db       3348 GGTCGGTGGCGACGGCGGACTCGGCGGCGCCGGCGGCGTTGGCGGGGCCGGCGGCAACGG 3407

Qy        186 CGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
              || |    | | ||| | ||| || |      || |  |  |      |   | || ||
Db       3408 CGGCATTGGTATCACCGTCGGCGGCGCCGGCGGAGCGGGCGGTAACGGCGGTGACCCCGG 3467

Qy        246 GCTTTGCGTTGGTGGGGGGATGCG 269
                 ||| | | || ||  || |
Db       3468 CGCAGGCGGTCGCGGCGGTCTGGG 3491


RESULT 33
US-11-075-185-36
; Sequence 36, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
;  APPLICANT: REEVES, CHRISTOPHER  D
;  APPLICANT:  JULIEN, BRYAN
;  APPLICANT:  REID, RALPH
;  TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
;  FILE REFERENCE: 010099.03
;  CURRENT APPLICATION NUMBER: US/11/075,185
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/551,103
;  PRIOR FILING DATE: 2004-03-08
;  PRIOR APPLICATION NUMBER: US 60/568,290
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
;   LENGTH: 5679
;   TYPE: DNA
;   ORGANISM: Sorangium cellulosum
US-11-075-185-36

  Query Match             3.1%;  Score 40.8;  DB 12;  Length 5679;
  Best Local Similarity   52.3%;  Pred. No. 3.5;
  Matches   90;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy         65 CGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATG 124
              | || |  || | ||  || ||| |  | |      ||  |  || | |   |  | |
Db       1297 CCGCGCACGGACCGGCCGCGGCGCGCGGCGGTCTCGTCCTTCGGCGTCAGCGGCACCAAC 1356

Qy        125 GCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCG 184
              |||   | || ||     || ||| |||     ||   | |||| | || | |||||
Db       1357 GCGCACGTCATCATCGAGGAGGCGCCGGCCGAAGCGCCGACGGCCGCGCAGACGGCGGCA 1416

Qy        185 GCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTC 236
              ||| | |   |||  | ||||| || ||    |  |  || || || |||| |
Db       1417 GCGGCGGCGACAGAGCCGGCGGCGGCGGTGGTGCCGTCGACGCTGCCGCTGC 1468


RESULT 34
US-11-075-185-1
; Sequence 1, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
;  APPLICANT: REEVES, CHRISTOPHER  D
;  APPLICANT:  JULIEN, BRYAN
;  APPLICANT:  REID, RALPH
```

```
;  TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
;  FILE REFERENCE: 010099.03
;  CURRENT APPLICATION NUMBER: US/11/075,185
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/551,103
;  PRIOR FILING DATE: 2004-03-08
;  PRIOR APPLICATION NUMBER: US 60/568,290
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
;   LENGTH: 78869
;   TYPE: DNA
;   ORGANISM: Sorangium cellulosum
US-11-075-185-1

  Query Match             3.1%;  Score 40.8;  DB 12;  Length 78869;
  Best Local Similarity   52.3%;  Pred. No. 11;
  Matches   90;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy         65 CGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATG 124
             | || |  || | ||  || ||| | | ||    || |  || | |   |  | | |
Db      27528 CCGCGCACGGACCGGCCGCGGCGCGCGGCGGTCTCGTCCTTCGGCGTCAGCGGCACCAAC 27587

Qy        125 GCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCG 184
              |||   | || ||     || ||| |||     ||   | |||| | ||| | |||||| |
Db      27588 GCGCACGTCATCATCGAGGAGGCGCCGGCCGAAGCGCCGACGGCCGCGCAGACGGCGGCA 27647

Qy        185 GCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTC 236
              ||| | |   |||  | |||||| || ||    |  |  || || || ||| |
Db      27648 GCGGCGGCGACAGAGCCGGCGGCGGCGGTGGTGCCGTCGACGCTGCCGCTGC 27699


RESULT 35
US-10-995-561-35803
; Sequence 35803, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35803
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-35803

  Query Match             3.1%;  Score 40.6;  DB 8;  Length 201;
  Best Local Similarity   52.7%;  Pred. No. 0.95;
  Matches   88;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy         45 GGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTT 104
              || |   ||| | | ||  |   | ||||||  ||  || ||| | || |  || |
Db          1 GGGCGGGGAGCCAAGGATCTGACCGCGAGGGGCGGGCGCGGCGGGGAGGGGCGGGGCGGG 60

Qy        105 GAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGC 164
              | |||   |  | |||             || || |  | |  | | ||||||   | ||||
Db         61 GCGGGCGGCGCGGGGCCTATTAAAGGCGCGGCCGGGCAGCSGGGCCGGAGCGGCCGGGGC 120

Qy        165 CGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAG 211
              |   | || | ||| || | ||||  |  || | |   |||| || |
Db        121 CACGATGGAGCGCGACGGCTGCGCGGGGGGCGGGAGCCGCGGCGGCG 167


RESULT 36
US-10-995-561-73911/c
; Sequence 73911, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73911
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-73911

  Query Match             3.1%;  Score 40.6;  DB 8;  Length 201;
  Best Local Similarity   52.7%;  Pred. No. 0.95;
  Matches   88;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy         45 GGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTT 104
              || |   ||| | | ||  |   | ||||||  ||  || ||| | || |  || |
Db        201 GGGCGGGGAGCCAAGGATCTGACCGCGAGGGGCGGGCGCGGCGGGGAGGGGCGGGGCGGG 142

Qy        105 GAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGC 164
              | |||   |  | |||             || || |  | |  | | ||||||   | ||||
Db        141 GCGGGCGGCGCGGGGCCTATTAAAGGCGCGGCCGGGCAGCSGGGCCGGAGCGGCCGGGGC 82

Qy        165 CGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAG 211
              |   | || | ||| || | ||||  |  || | |   |||| || |
Db         81 CACGATGGAGCGCGACGGCTGCGCGGGGGGCGGGAGCCGCGGCGGCG 35


RESULT 37
US-10-967-648A-1
; Sequence 1, Application US/10967648A
; Publication No. US20050245473A1
; GENERAL INFORMATION:
;  APPLICANT: Saunders, Nicholas A
;  TITLE OF INVENTION: Differentiation- and/or proliferation-modulating agents and uses
;  TITLE OF INVENTION:  therefor
;  FILE REFERENCE: 12493972
;  CURRENT APPLICATION NUMBER: US/10/967,648A
;  CURRENT FILING DATE:  2004-10-15
;  PRIOR APPLICATION NUMBER: USSN 60/512010
;  PRIOR FILING DATE: 2003-10-16
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
;   LENGTH: 2486
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (124)..(1437)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (322)..(447)
;   OTHER INFORMATION: cyclin A/CDK2 binding domain
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (451)..(705)
;   OTHER INFORMATION: DNA binding domain
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (490)..(1104)
;   OTHER INFORMATION: Transcription factor E2F/dimerisation partner (TDP)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (580)..(645)
;   OTHER INFORMATION: leucine zipper domain
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (706)..(905)
;   OTHER INFORMATION: dimerisation domain
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1225)..(1434)
;   OTHER INFORMATION: transactivation domain
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1348)..(1401)
;   OTHER INFORMATION: retinoblastoma protein association domain
US-10-967-648A-1

  Query Match             3.1%;  Score 40.6;  DB 8;  Length 2486;
  Best Local Similarity   48.5%;  Pred. No. 2.8;
  Matches  112;  Conservative    0;  Mismatches  119;  Indels    0;  Gaps    0;

Qy        107 GGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCG 166
              ||| | |  || ||  | || ||| || | |  |||    ||| | |  ||    | |
Db         70 GGGCCCGCGCCGCCGCCGCCGCCTGTCACCCGGGCCGCGCGGGCCGTGAGCGTCATGGCC 129

Qy        167 GCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGG 226
                |  ||||  |  |||    |  |   || | ||||  ||||| || | |   |
Db        130 TTGGCCGGGGCCCCTGCGGGCGGCCCATGCGCGCCGGCGCTGGAGGCCCTGCTCGGGGCC 189

Qy        227 CGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCC 286
              |  ||| | |   || | ||   ||  | | |       |  | ||| ||  || | ||
Db        190 GGCGCGCTGCGGCTGCTCGACTCCTCGCAGATCGTCATCATCTCCGCCGCGCAGGACGCC 249

Qy        287 AGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTAC 337
              |||| || || ||    | | | |   || |   |||    || ||| |
Db        250 AGCGCCCCGCCGGCTCCCACCGGCCCCGCGGCGCCCGCCGCCGGCCCCTGC 300


RESULT 38
US-09-925-065A-8726
; Sequence 8726, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8726
;   LENGTH: 702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-8726

  Query Match             3.1%;  Score 40.4;  DB 6;  Length 702;
  Best Local Similarity   55.0%;  Pred. No. 1.8;
  Matches   77;  Conservative    1;  Mismatches   62;  Indels    0;  Gaps    0;

Qy        151 GGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGA 210
              |||| |||    ||| |  |||  || | ||   ||||||| |||||  |  ||| | |
Db        255 GGAGGCCGTGTTGCCCGTGGGGGCGGAGCCGAGCGCGCCCGGAGCAGGAGGGGCCGGGAG 314

Qy        211 GGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGA 270
              ||   |||||  | || | ||     :  ||  |  |   |||  |   ||| ||  |
Db        315 GGCGTGGGTGGGGCGGGGCCGATACGWCAGCAAGCTCGGGGCGGGGCCCGGGTGACCCCG 374

Qy        271 CCCCACGATGGTCCCCAGCG 290
                |||  |  |||||  ||| |
Db        375 ACCCCGGGCGGTCCGCAGTG 394


RESULT 39
US-10-995-561-13450/c
; Sequence 13450, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13450
;   LENGTH: 12309
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-13450

  Query Match             3.1%;  Score 40.4;  DB 8;  Length 12309;
  Best Local Similarity   51.1%;  Pred. No. 6.1;
  Matches   89;  Conservative    2;  Mismatches   83;  Indels    0;  Gaps    0;

Qy         38 GGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGG 97
              ||   | || |   ||| | | ||  |   | |:||||  ||  || :|| | || |  |
Db       6284 GGGTGTAGGGCGGGGAGCCAAGGATCTGACCGCRAGGGGCGGGCGCGKCGGGGAGGGGCG 6225

Qy         98 GTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCC 157
              | |   | |||   |  | |||            || || |  | |  | | ||||||
Db       6224 GRGCGGGGCGGGCGGCGCGGGGCCTATTAAAGGCGCGGCSGGGCAGCSGGGCCGGAGCGG 6165

Qy        158 GCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAG 211
              | |||||   | || | ||| || | ||||  |  || | |   |||| || |
Db       6164 CCGGGGCCACGATGGAGCGCGACGGCTGCGCGGGGGGCGGGAGCCGCGGCGGCG 6111


RESULT 40
US-10-995-561-13273
; Sequence 13273, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13273
;   LENGTH: 95832
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-13273

  Query Match             3.1%;  Score 40.4;  DB 8;  Length 95832;
  Best Local Similarity   51.1%;  Pred. No. 15;
  Matches   89;  Conservative    2;  Mismatches   83;  Indels    0;  Gaps    0;

Qy         38 GGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGG 97
              ||   | || |   ||| | | ||  |   | |:||||  ||  || :|| | || |  |
Db       5887 GGGTGTAGGGCGGGGAGCCAAGGATCTGACCGCRAGGGGCGGGCGCGKCGGGGAGGGGCG 5946

Qy         98 GTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCC 157
```

```
                | | |   | |||   |  | |||           || || |  | |  | | ||||||
Db       5947 GRGCGGGGCGGGCGGCGCGGGGCCTATTAAAGGCGCGGCSGGGCAGCSGGGCCGGAGCGG 6006

Qy        158 GCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAG 211
                | |||||   | || | ||| || | ||||  |  || | |   |||| || |
Db       6007 CCGGGGCCACGATGGAGCGCGACGGCTGCGCGGGGGGCGGGAGCCGCGGCGGCG 6060


RESULT 41
US-11-136-527-1009
; Sequence 1009, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1009
;   LENGTH: 1699
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1280)..(1280)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-1009

  Query Match             3.0%;  Score 40;  DB 12;  Length 1699;
  Best Local Similarity   40.7%;  Pred. No. 3.4;
  Matches   22;  Conservative   31;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         22 ATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              :|:|::|:|::|:||:::||::::|||::|:|| :|::::|:||::::|:|:::|
Db       1077 WTWCKRCKCMYTWTAKSRAAWKWRGCCYYCSAGAMCMWKAMTTMKRYAYGWRRG 1130


RESULT 42
US-11-136-527-1488
; Sequence 1488, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1488
;   LENGTH: 588
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-1488

  Query Match             3.0%;  Score 39.8;  DB 12;  Length 588;
  Best Local Similarity   57.4%;  Pred. No. 2.4;
  Matches   35;  Conservative   18;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGG 78
              || ||||||||||||||||||||||||::|::::||||::|: :  : : :  |::|::
Db          2 TTAATACGACTCACTATAGGGAATTTGGSSCWYKMGGCCWMGWKBCGSYMYBMGGSRCSR 61

Qy         79 G 79
              :
Db         62 D 62


RESULT 43
US-11-136-527-5584
; Sequence 5584, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5584
;   LENGTH: 588
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-5584

  Query Match             3.0%;  Score 39.8;  DB 12;  Length 588;
  Best Local Similarity   57.4%;  Pred. No. 2.4;
  Matches   35;  Conservative   18;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGG 78
              || ||||||||||||||||||||||||::|::::||||::|: :  : : :  |::|::
Db          2 TTAATACGACTCACTATAGGGAATTTGGSSCWYKMGGCCWMGWKBCGSYMYBMGGSRCSR 61

Qy         79 G 79
              :
Db         62 D 62


RESULT 44
US-10-750-185-34334/c
; Sequence 34334, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 34334
;   LENGTH: 1196
;   TYPE: DNA
;   ORGANISM: Bovine  19866881287916
US-10-750-185-34334

  Query Match             3.0%;  Score 39.8;  DB 8;  Length 1196;
  Best Local Similarity   46.5%;  Pred. No. 3.3;
  Matches  128;  Conservative    0;  Mismatches  147;  Indels    0;  Gaps    0;

Qy        691 AGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCC 750
              | || |||| | ||| | | |          |||| |    | | ||   |   |   |
Db        375 ACCTCTACTTATCCAGATATGGAAAACAGCTTTTAGACACGATGTGACATTAATTACCCA 316

Qy        751 AAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATT 810
              |||||| || ||  | | | |  ||||| | | |||||||||  ||  ||| |  | ||
```

```
Db        315 AAATACAAATATTTTTAAATAAATGCTAAGAGAAAATGTGACATTTTGTGAATTTTAAAA 256

Qy        811 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACA 870

Db        255 AATCTTTCTGAAGCTTTGCTACACGTTAATTCTCCCACACATGAGCCTTTGAAAACATTA 196

Qy        871 ATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGC 930

Db        195 AAAACGAAAATGCCAGCTGAATATTAAATACTTTCATTAGGCAACAACAGTATATACTGT 136

Qy        931 AGGAGATACAAATTTAAGAGATCAAGAAGTTATCA 965

Db        135 TAATATTCCAAATGTAAAAGATTAATAATTCACCA 101


RESULT 45
US-10-750-623-34334/c
; Sequence 34334, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 34334
;   LENGTH: 1196
;   TYPE: DNA
;   ORGANISM: Bovine  19866881287916
US-10-750-623-34334

  Query Match             3.0%;  Score 39.8;  DB 8;  Length 1196;
  Best Local Similarity   46.5%;  Pred. No. 3.3;
  Matches  128;  Conservative    0;  Mismatches  147;  Indels    0;  Gaps    0;

Qy        691 AGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCC 750

Db        375 ACCTCTACTTATCCAGATATGGAAAACAGCTTTTAGACACGATGTGACATTAATTACCCA 316

Qy        751 AAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATT 810

Db        315 AAATACAAATATTTTTAAATAAATGCTAAGAGAAAATGTGACATTTTGTGAATTTTAAAA 256

Qy        811 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACA 870

Db        255 AATCTTTCTGAAGCTTTGCTACACGTTAATTCTCCCACACATGAGCCTTTGAAAACATTA 196

Qy        871 ATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGC 930

Db        195 AAAACGAAAATGCCAGCTGAATATTAAATACTTTCATTAGGCAACAACAGTATATACTGT 136

Qy        931 AGGAGATACAAATTTAAGAGATCAAGAAGTTATCA 965

Db        135 TAATATTCCAAATGTAAAAGATTAATAATTCACCA 101


RESULT 46
US-10-453-372-1183/c
; Sequence 1183, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
;  APPLICANT: Alsobrook, et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;  FILE REFERENCE: 21402-589 A
;  CURRENT APPLICATION NUMBER: US/10/453,372
;  CURRENT FILING DATE:  2003-06-03
;  PRIOR APPLICATION NUMBER: 09/789390
;  PRIOR FILING DATE: 2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/823187
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195792
;  PRIOR FILING DATE: 2000-03-10
;  PRIOR APPLICATION NUMBER: 09/839446
;  PRIOR FILING DATE: 2001-03-19
;  PRIOR APPLICATION NUMBER: 60/199476
;  PRIOR FILING DATE: 2000-03-25
;  PRIOR APPLICATION NUMBER: 09/863776
;  PRIOR FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: 60/208263
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 09/939398
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: 60/227800
;  PRIOR FILING DATE: 2000-08-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1609
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1183
;   LENGTH: 1931
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (37)..(1902)
US-10-453-372-1183

  Query Match             3.0%;  Score 39.8;  DB 8;  Length 1931;
  Best Local Similarity   50.6%;  Pred. No. 4;
  Matches  127;  Conservative    0;  Mismatches  117;  Indels    7;  Gaps    1;

Qy        139 TTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGC 198

Db        817 TGCCGAAGCTGCAGGTCACGTTGGAGCAGGGGTCCCGCGAGCCGCACGGCCCGCGGCTGA 758

Qy        199 ACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGT 258

Db        757 GCAGGCGGATGCGGCGCTGCTGGCTGCACTGCGCCCGCTGCAGCTCGCATTCGTTGCTGT 698

Qy        259 GGGGGGATGCGACCC-------CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGC 311

Db        697 AGGTGGAGGCGTCCGACCCACACACAGGCGCCACCACGCTGGGGCACGGGCTCTTCTTGC 638

Qy        312 AGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGC 371

Db        637 AGACGCAGGACGCCCGGCCCGGCCCCTCCGCGTTGGGCTCGCACACGGCGCCGAAGCCGC 578

Qy        372 GCCAGCCTCCC 382

Db        577 ACAGCATTCCC 567


RESULT 47
US-11-052-554A-520
; Sequence 520, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
```

```
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 520
;   LENGTH: 3990
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-520

  Query Match             3.0%;  Score 39.8;  DB 12;  Length 3990;
  Best Local Similarity   53.5%;  Pred. No. 5.5;
  Matches   83;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy        109 GGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGC 168
              ||  ||  |||||  | |||  |   ||      | |  |||||| ||| || ||  ||||
Db       2758 GGTGCCAACGGCGTGGCCGTGGGTGCGCCCGGACAACCCGGCGGTGCCGGCGGGCACGGC 2817

Qy        169 AGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCG 228
              | || ||||||  |||||     || |||  ||||   |||    |     || |
Db       2818 GGCGCCGGCGGCAACGGCGGGGCCGGCGGCAACGGCGGTCAAGGCGTAGTCAGCGACGGC 2877

Qy        229 GCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGG 263
              ||||     |   ||||    |||  |  || ||
Db       2878 GCGGGCGGTGCCGGTGGGGCCGGCGGCGACGGCGG 2912


RESULT 48
US-10-453-372-1185/c
; Sequence 1185, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
;  APPLICANT: Alsobrook, et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;  FILE REFERENCE: 21402-589 A
;  CURRENT APPLICATION NUMBER: US/10/453,372
;  CURRENT FILING DATE:  2003-06-03
;  PRIOR APPLICATION NUMBER: 09/789390
;  PRIOR FILING DATE: 2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/823187
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195792
;  PRIOR FILING DATE: 2000-03-10
;  PRIOR APPLICATION NUMBER: 09/839446
;  PRIOR FILING DATE: 2001-03-19
;  PRIOR APPLICATION NUMBER: 60/199476
;  PRIOR FILING DATE: 2000-03-25
;  PRIOR APPLICATION NUMBER: 09/863776
;  PRIOR FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: 60/208263
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 09/939398
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: 60/227800
;  PRIOR FILING DATE: 2000-08-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1609
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1185
;   LENGTH: 4697
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (37)..(4668)
US-10-453-372-1185

  Query Match             3.0%;  Score 39.8;  DB 8;  Length 4697;
  Best Local Similarity   50.6%;  Pred. No. 5.8;
  Matches  127;  Conservative    0;  Mismatches  117;  Indels    7;  Gaps    1;

Qy        139 TTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGC 198
              | |||| || || |  | ||  ||  | |  |  | |    ||  |   |||||  || |
Db        817 TGCCGAAGCTGCAGGTCACGTTGGAGCAGGGGTCCCGCGAGCCGCACGGCCCGCGGCTGA 758

Qy        199 ACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGT 258
              ||||||||  | ||  | | ||       | |||  |  ||   || |  ||  |||
Db        757 GCAGGCGGATGCGGCGCTGCTGGCTGCACTGCGCCCGCTGCAGCTCGCATTCGTTGCTGT 698

Qy        259 GGGGGGATGCGACCC-------CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGC 311
              || |||  ||| ||           |||   |  | |||   |   ||  |  |  | || ||
Db        697 AGGTGGAGGCGTCCGACCCACACACAGGCGCCACCACGCTGGGGCACGGGCTCTTCTTGC 638

Qy        312 AGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGC 371
              ||||||||  | ||||  |  || |  || || | ||      | ||         |  |||||
Db        637 AGACGCAGGACGCCCGGCCCGGCCCCTCCGCGTTGGGCTCGCACACGGCGCCGAAGCCGC 578

Qy        372 GCCAGCCTCCC 382
              |    ||||
Db        577 ACAGCATTCCC 567


RESULT 49
US-10-453-372-1179/c
; Sequence 1179, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
;  APPLICANT: Alsobrook, et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD.
;  FILE REFERENCE: 21402-589 A
;  CURRENT APPLICATION NUMBER: US/10/453,372
;  CURRENT FILING DATE:  2003-06-03
;  PRIOR APPLICATION NUMBER: 09/789390
;  PRIOR FILING DATE: 2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/823187
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195792
;  PRIOR FILING DATE: 2000-03-10
;  PRIOR APPLICATION NUMBER: 09/839446
;  PRIOR FILING DATE: 2001-03-19
;  PRIOR APPLICATION NUMBER: 60/199476
;  PRIOR FILING DATE: 2000-03-25
;  PRIOR APPLICATION NUMBER: 09/863776
;  PRIOR FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: 60/208263
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 09/939398
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: 60/227800
;  PRIOR FILING DATE: 2000-08-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1609
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1179
;   LENGTH: 4760
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (37)..(4731)
US-10-453-372-1179

  Query Match             3.0%;  Score 39.8;  DB 8;  Length 4760;
  Best Local Similarity   50.6%;  Pred. No. 5.9;
  Matches  127;  Conservative    0;  Mismatches  117;  Indels    7;  Gaps    1;

Qy        139 TTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGC 198
              | |||| || || |  | ||  ||  | |  |  | |    ||  |   |||||  || |
Db        817 TGCCGAAGCTGCAGGTCACGTTGGAGCAGGGGTCCCGCGAGCCGCACGGCCCGCGGCTGA 758

Qy        199 ACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGT 258
              ||||||||  | ||  | | ||       | |||  |  ||   || |  ||  |||
```

```
Db        757 GCAGGCGGATGCGGCGCTGCTGGCTGCACTGCGCCCGCTGCAGCTCGCATTCGTTGCTGT 698

Qy        259 GGGGGGATGCGACCC-------CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGC 311
              || ||| ||| ||          |||    | | |||    |    ||  | |  |  || ||
Db        697 AGGTGGAGGCGTCCGACCCACACACAGGCGCCACCACGCTGGGGCACGGGCTCTTCTTGC 638

Qy        312 AGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGC 371
              |||||||| | |||| |  || |  || || | ||      | ||          | |||||
Db        637 AGACGCAGGACGCCCGGCCCGGCCCCTCCGCGTTGGGCTCGCACACGGCGCCGAAGCCGC 578

Qy        372 GCCAGCCTCCC 382
              |     ||||
Db        577 ACAGCATTCCC 567


RESULT 50
US-10-453-372-1189/c
; Sequence 1189, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
;  APPLICANT: Alsobrook, et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;  FILE REFERENCE: 21402-589 A
;  CURRENT APPLICATION NUMBER: US/10/453,372
;  CURRENT FILING DATE:  2003-06-03
;  PRIOR APPLICATION NUMBER: 09/789390
;  PRIOR FILING DATE: 2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/823187
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195792
;  PRIOR FILING DATE: 2000-03-10
;  PRIOR APPLICATION NUMBER: 09/839446
;  PRIOR FILING DATE: 2001-03-19
;  PRIOR APPLICATION NUMBER: 60/199476
;  PRIOR FILING DATE: 2000-03-25
;  PRIOR APPLICATION NUMBER: 09/863776
;  PRIOR FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: 60/208263
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 09/939398
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: 60/227800
;  PRIOR FILING DATE: 2000-08-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1609
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1189
;   LENGTH: 5688
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (37)..(4734)
US-10-453-372-1189

  Query Match             3.0%;  Score 39.8;  DB 8;  Length 5688;
  Best Local Similarity   50.6%;  Pred. No. 6.3;
  Matches  127;  Conservative    0;  Mismatches  117;  Indels    7;  Gaps    1;

Qy        139 TTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGC 198
              | |||| || || |  | ||  ||  | |  |  | |   || |   |||||  || |
Db        817 TGCCGAAGCTGCAGGTCACGTTGGAGCAGGGGTCCCGCGAGCCGCACGGCCCGCGGCTGA 758

Qy        199 ACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGT 258
              |||||||| | ||  | | ||        |||| |  | |   || |  ||      ||
Db        757 GCAGGCGGATGCGGCGCTGCTGGCTGCACTGCGCCCGCTGCAGCTCGCATTCGTTGCTGT 698

Qy        259 GGGGGGATGCGACCC-------CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGC 311
              || ||| ||| ||          |||    | | |||    |    ||  | |  |  || ||
Db        697 AGGTGGAGGCGTCCGACCCACACACAGGCGCCACCACGCTGGGGCACGGGCTCTTCTTGC 638

Qy        312 AGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGC 371
              |||||||| | |||| |  || |  || || | ||      | ||          | |||||
Db        637 AGACGCAGGACGCCCGGCCCGGCCCCTCCGCGTTGGGCTCGCACACGGCGCCGAAGCCGC 578

Qy        372 GCCAGCCTCCC 382
              |     ||||
Db        577 ACAGCATTCCC 567


RESULT 51
US-10-453-372-1173/c
; Sequence 1173, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
;  APPLICANT: Alsobrook, et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHO
;  FILE REFERENCE: 21402-589 A
;  CURRENT APPLICATION NUMBER: US/10/453,372
;  CURRENT FILING DATE:  2003-06-03
;  PRIOR APPLICATION NUMBER: 09/789390
;  PRIOR FILING DATE: 2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/823187
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195792
;  PRIOR FILING DATE: 2000-03-10
;  PRIOR APPLICATION NUMBER: 09/839446
;  PRIOR FILING DATE: 2001-03-19
;  PRIOR APPLICATION NUMBER: 60/199476
;  PRIOR FILING DATE: 2000-03-25
;  PRIOR APPLICATION NUMBER: 09/863776
;  PRIOR FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: 60/208263
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 09/939398
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: 60/227800
;  PRIOR FILING DATE: 2000-08-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1609
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1173
;   LENGTH: 6224
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (37)..(6195)
US-10-453-372-1173

  Query Match             3.0%;  Score 39.8;  DB 8;  Length 6224;
  Best Local Similarity   50.6%;  Pred. No. 6.6;
  Matches  127;  Conservative    0;  Mismatches  117;  Indels    7;  Gaps    1;

Qy        139 TTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGC 198
              | |||| || || |  | ||  ||  | |  |  | |   || |   |||||  || |
Db        817 TGCCGAAGCTGCAGGTCACGTTGGAGCAGGGGTCCCGCGAGCCGCACGGCCCGCGGCTGA 758

Qy        199 ACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGT 258
              |||||||| | ||  | | ||        |||| |  | |   || |  ||      ||
Db        757 GCAGGCGGATGCGGCGCTGCTGGCTGCACTGCGCCCGCTGCAGCTCGCATTCGTTGCTGT 698

Qy        259 GGGGGGATGCGACCC-------CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGC 311
              || ||| ||| ||          |||    | | |||    |    ||  | |  |  || ||
Db        697 AGGTGGAGGCGTCCGACCCACACACAGGCGCCACCACGCTGGGGCACGGGCTCTTCTTGC 638

Qy        312 AGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGC 371
              |||||||| | |||| |  || |  || || | ||      | ||          | |||||
Db        637 AGACGCAGGACGCCCGGCCCGGCCCCTCCGCGTTGGGCTCGCACACGGCGCCGAAGCCGC 578

Qy        372 GCCAGCCTCCC 382
              |     ||||
```

```
Db        577 ACAGCATTCCC 567


RESULT 52
US-10-453-372-1187/c
; Sequence 1187, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
;  APPLICANT: Alsobrook, et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;  FILE REFERENCE: 21402-589 A
;  CURRENT APPLICATION NUMBER: US/10/453,372
;  CURRENT FILING DATE:  2003-06-03
;  PRIOR APPLICATION NUMBER: 09/789390
;  PRIOR FILING DATE: 2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/823187
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195792
;  PRIOR FILING DATE: 2000-03-10
;  PRIOR APPLICATION NUMBER: 09/839446
;  PRIOR FILING DATE: 2001-03-19
;  PRIOR APPLICATION NUMBER: 60/199476
;  PRIOR FILING DATE: 2000-03-25
;  PRIOR APPLICATION NUMBER: 09/863776
;  PRIOR FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: 60/208263
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 09/939398
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: 60/227800
;  PRIOR FILING DATE: 2000-08-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1609
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1187
;   LENGTH: 6494
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (37)..(6465)
US-10-453-372-1187

  Query Match             3.0%;  Score 39.8;  DB 8;  Length 6494;
  Best Local Similarity   50.6%;  Pred. No. 6.7;
  Matches  127;  Conservative    0;  Mismatches  117;  Indels    7;  Gaps    1;

Qy        139 TTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGC 198
              | |||| || || |  | ||  ||  | |  | | |   ||  |  |||||  || |
Db        817 TGCCGAAGCTGCAGGTCACGTTGGAGCAGGGGTCCCGCGAGCCGCACGGCCCGCGGCTGA 758

Qy        199 ACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGT 258
               |||||||| | ||  | | ||      | ||| | | |   || |  || |      ||
Db        757 GCAGGCGGATGCGGCGCTGCTGGCTGCACTGCGCCCGCTGCAGCTCGCATTCGTTGCTGT 698

Qy        259 GGGGGGATGCGACCC-------CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGC 311
               || ||| ||| ||        |||    | | |||   |   ||  | |  |  || ||
Db        697 AGGTGGAGGCGTCCGACCCACACACAGGCGCCACCACGCTGGGGCACGGGCTCTTCTTGC 638

Qy        312 AGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGC 371
              |||||||| | |||| |  || |  || || | ||     | ||         |  |||||
Db        637 AGACGCAGGACGCCCGGCCCGGCCCCTCCGCGTTGGGCTCGCACACGGCGCCGAAGCCGC 578

Qy        372 GCCAGCCTCCC 382
              |  ||||
Db        577 ACAGCATTCCC 567


RESULT 53
US-11-136-527-306
; Sequence 306, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 306
;   LENGTH: 15234
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-306

  Query Match             3.0%;  Score 39.6;  DB 12;  Length 15234;
  Best Local Similarity   59.3%;  Pred. No. 11;
  Matches   89;  Conservative    0;  Mismatches   54;  Indels    7;  Gaps    1;

Qy        103 TTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGG 162
              | |||| | ||| | ||||   |||| |||||  |      || ||||  |   |||
Db       8075 TGGAGGAGGCCCGGAGGCGGCAGCGTGAGGCAGAGGAGGGTGTGAGGCGCAAGCAAGAGG 8134

Qy        163 GCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAG-------CACAGGCGGAGGAGGACC 215
                | |||| |   |  ||  | ||| | | |||||         || ||| |||||| |||
Db       8135 AACTGCAGCGTCTGGAGCAGCAGCGGCAGCAGCAGGAGAAACTACTGGCAGAGGAGAACC 8194

Qy        216 GGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
               |  |  | ||  |||||| ||| ||||||
Db       8195 AGAGGCTGCGGGAGCGGCTGCAGCGCCTGG 8224


RESULT 54
US-09-925-065A-414314/c
; Sequence 414314, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 414314
;   LENGTH: 571
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-414314

  Query Match             3.0%;  Score 39.4;  DB 6;  Length 571;
  Best Local Similarity   56.3%;  Pred. No. 3;
  Matches   94;  Conservative    0;  Mismatches   71;  Indels    2;  Gaps    1;

Qy         82 GCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTC 141
              |||||| |  | |  |          ||       | ||||||   ||      | ||| |  |
Db        530 GCAGCGGGGCGGGGAGCCTGGGGCAGCTACGGCGGCGGCGGACGCTGGCTGGAGCCGGGC 471

Qy        142 CGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACA 201
```

```
Db        470 AGAGTGAACGCACCTGGCGCCGCCGGGACCGCGGGCGGACTCAGCCCAGGAAGCGGCAGC 411

Qy        202 GGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCT 248
              ||| || ||| |||| | || |||||||||  | ||| | | | |
Db        410 GGC--AGCTGGAACGGGAGGAGGGGCGGCGGCGGCGGTGGCGGTGGT 366


RESULT 55
US-10-632-150-39
; Sequence 39, Application US/10632150
; Publication No. US20050251871A1
; GENERAL INFORMATION:
;  APPLICANT: Chiaur, D.
;  APPLICANT:  Pagano, M.
;  APPLICANT:  Latres, E.
;  TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
;  FILE REFERENCE: 5914-081
;  CURRENT APPLICATION NUMBER: US/10/632,150
;  CURRENT FILING DATE:  2003-07-30
;  PRIOR APPLICATION NUMBER: US/09/385,219
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: 60/098,355
;  PRIOR FILING DATE: 1998-08-28
;  PRIOR APPLICATION NUMBER: 60/118,568
;  PRIOR FILING DATE: 1999-02-03
;  PRIOR APPLICATION NUMBER: 60/124,449
;  PRIOR FILING DATE: 1999-03-15
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 774
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-632-150-39

  Query Match             3.0%;  Score 39.4;  DB 8;  Length 774;
  Best Local Similarity   52.8%;  Pred. No. 3.5;
  Matches   85;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy        141 CCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCAC 200
              |||  || ||| | |  |  | || | | | || |  |  |  |          || ||||
Db          8 CCGCCGCCGCGTACCTGGACGAGCTGCCCGAGCCGCTGCTGCTGCGCGTGCTGGCCGCAC 67

Qy        201 AGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGG 260
              | |||  |  || | |||| ||     | || ||   ||||||| |||    |  | |||
Db         68 TGCCGGCCGCCGAGCTGGTGCAGGCCTGCCGCCTGGTGTGCCTGCGCTGGAAGGAGCTGG 127

Qy        261 GGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
              ||   ||| |||   |  | | | ||    || || ||||
Db        128 TGGACGGCGCCCCGCTGTGGCTGCTCAAGTGCCAGCAGGAG 168


RESULT 56
US-11-073-457-39
; Sequence 39, Application US/11073457
; Publication No. US20050260556A1
; GENERAL INFORMATION:
;  APPLICANT: Pagano, M.
;  TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROLIFE
;  TITLE OF INVENTION:  AND DIFFERENTIATIVE DISORDERS
;  FILE REFERENCE: 5914-090-999
;  CURRENT APPLICATION NUMBER: US/11/073,457
;  CURRENT FILING DATE:  2005-03-04
;  PRIOR APPLICATION NUMBER: 10/042,417
;  PRIOR FILING DATE: 2002-01-07
;  PRIOR APPLICATION NUMBER: 60/260,179
;  PRIOR FILING DATE: 2001-01-05
;  NUMBER OF SEQ ID NOS: 92
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 774
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-073-457-39

  Query Match             3.0%;  Score 39.4;  DB 12;  Length 774;
  Best Local Similarity   52.8%;  Pred. No. 3.5;
  Matches   85;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy        141 CCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCAC 200
              |||  || ||| | |  |  | || | | | || |  |  |  |          || ||||
Db          8 CCGCCGCCGCGTACCTGGACGAGCTGCCCGAGCCGCTGCTGCTGCGCGTGCTGGCCGCAC 67

Qy        201 AGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGG 260
              | |||  |  || | |||| ||     | || ||   ||||||| |||    |  | |||
Db         68 TGCCGGCCGCCGAGCTGGTGCAGGCCTGCCGCCTGGTGTGCCTGCGCTGGAAGGAGCTGG 127

Qy        261 GGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
              ||   ||| |||   |  | | | ||    || || ||||
Db        128 TGGACGGCGCCCCGCTGTGGCTGCTCAAGTGCCAGCAGGAG 168


RESULT 57
US-11-073-460-39
; Sequence 39, Application US/11073460
; Publication No. US20050272066A1
; GENERAL INFORMATION:
;  APPLICANT: Pagano, M.
;  TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROLIF
;  TITLE OF INVENTION:  AND DIFFERENTIATIVE DISORDERS
;  FILE REFERENCE: 5914-090-999
;  CURRENT APPLICATION NUMBER: US/11/073,460
;  CURRENT FILING DATE:  2005-03-04
;  PRIOR APPLICATION NUMBER: 10/042,417
;  PRIOR FILING DATE: 2002-01-07
;  PRIOR APPLICATION NUMBER: 60/260,179
;  PRIOR FILING DATE: 2001-01-05
;  NUMBER OF SEQ ID NOS: 92
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 774
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-073-460-39

  Query Match             3.0%;  Score 39.4;  DB 12;  Length 774;
  Best Local Similarity   52.8%;  Pred. No. 3.5;
  Matches   85;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy        141 CCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCAC 200
              |||  || ||| | |  |  | || | | | || |  |  |  |          || ||||
Db          8 CCGCCGCCGCGTACCTGGACGAGCTGCCCGAGCCGCTGCTGCTGCGCGTGCTGGCCGCAC 67

Qy        201 AGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGG 260
              | |||  |  || | |||| ||     | || ||   ||||||| |||    |  | |||
Db         68 TGCCGGCCGCCGAGCTGGTGCAGGCCTGCCGCCTGGTGTGCCTGCGCTGGAAGGAGCTGG 127

Qy        261 GGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
              ||   ||| |||   |  | | | ||    || || ||||
Db        128 TGGACGGCGCCCCGCTGTGGCTGCTCAAGTGCCAGCAGGAG 168


RESULT 58
US-11-096-568A-16503
; Sequence 16503, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 16503
```

```
;   LENGTH: 997
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(997)
;   OTHER INFORMATION: Ceres Seq. ID no. 12352560
US-11-096-568A-16503

  Query Match             3.0%;  Score 39.4;  DB 9;  Length 997;
  Best Local Similarity   48.7%;  Pred. No. 3.9;
  Matches  135;  Conservative    0;  Mismatches  141;  Indels    1;  Gaps    1;

Qy        124 GGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTC 183
              ||   || |||||||   |||  || | || | || |  |||||   |||| | ||  ||
Db        176 GGGTTCCGGCAGCATCCCCGGAGCCGTGGCGGCCCCGTGGCCGTGCGGGC-GTCGCTGTT 234

Qy        184 GGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCT 243
               |||||  |  | ||   |||  |||| |  |||  ||  | |  || |   | || | |
Db        235 CTCGCCCAAGCCCGCGGCGGCCAAGGACGCGCGGCCGACCAAGGTGCAGGAGCTGTACGT 294

Qy        244 GGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAA 303
              |  |      |   | | |||   ||||  | || |  ||  || ||      ||| |
Db        295 GTACGAGATGAACGAGCGGGACCGCGAGAGCCCGGCGTACCTGCGCCTCAGCGCCAAGCA 354

Qy        304 CGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGG 363
                | | ||| |   | |||| | ||| |    |   |    | |||| |
Db        355 GACGGAGAACGCGCTGGGCGACCTGGTCCCCTTCACAAACAAGGTTTACAACGGCAGCCT 414

Qy        364 GTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAG 400
              |     ||| | |     |||  ||  ||  | || |
Db        415 GGACAAGCGGCTGGGCATCACCGCCGGCATCTGCGTG 451


RESULT 59
US-11-000-463-95
; Sequence 95, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Cao, Yi-Cheng
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 785CIP4CN
;  CURRENT APPLICATION NUMBER: US/11/000,463
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: 10/291,265
;  PRIOR FILING DATE: 2002-11-08
;  PRIOR APPLICATION NUMBER: PCT/US01/02623
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 09/617,746
;  PRIOR FILING DATE: 2000-07-17
;  PRIOR APPLICATION NUMBER: 09/631,451
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 09/633,870
;  PRIOR FILING DATE: 2000-09-15
;  NUMBER OF SEQ ID NOS: 944
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
;   LENGTH: 1278
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (50)..(1003)
US-11-000-463-95

  Query Match             3.0%;  Score 39.4;  DB 12;  Length 1278;
  Best Local Similarity   97.6%;  Pred. No. 4.3;
  Matches   40;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         41 ATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAA 81
              |||||||||||||||||||||||||||||||||||||||||
Db          1 ATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGCA 41


RESULT 60
US-11-052-554A-549
; Sequence 549, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 549
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-549

  Query Match             3.0%;  Score 39.4;  DB 12;  Length 1920;
  Best Local Similarity   65.2%;  Pred. No. 5.1;
  Matches   58;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy        118 GGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGC 177
              ||||   |||  | |  | |||| |  ||||||| ||| || |   |||| |||| |||
Db       1573 GGCGGCAGCGGAGGCCTACTGTTCGGTAGCGGCGGGGCCGGCGGCATCGGCGGGGCCGGC 1632

Qy        178 GGCGTCGGCGCCCGAAGCAGCACAGGCGG 206
              ||||||||||      ||| |   |||||
Db       1633 GGCGTCGGCGGGTCGGGCAACGACGGCGG 1661


RESULT 61
US-10-750-185-35012/c
; Sequence 35012, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 35012
;   LENGTH: 2518
```

```
;   TYPE: DNA
;   ORGANISM: Bovine  19866881113848
US-10-750-185-35012

  Query Match             3.0%;  Score 39.4;  DB 8;  Length 2518;
  Best Local Similarity   61.0%;  Pred. No. 5.7;
  Matches   64;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy        109 GGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGC 168
              |||| | || ||| || | || ||||||| | | |||| | ||
Db        207 GGACTCAGCCGCGGCTGCTGGACCCGGGCACCGGGAGGCGGCGGCGGCGGCAGCAGCAGC 148

Qy        169 AGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGA 213
              || |  ||||||| ||||| | |   ||||| | | ||| |||||
Db        147 AGCGGCGGCGGCGGCGGCGGCGGCGACAGCAGAAGAGGAAGAGGA 103


RESULT 62
US-10-750-623-35012/c
; Sequence 35012, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 35012
;   LENGTH: 2518
;   TYPE: DNA
;   ORGANISM: Bovine  19866881113848
US-10-750-623-35012

  Query Match             3.0%;  Score 39.4;  DB 8;  Length 2518;
  Best Local Similarity   61.0%;  Pred. No. 5.7;
  Matches   64;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy        109 GGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGC 168
              |||| | || ||| || | || ||||||| | | |||| | ||
Db        207 GGACTCAGCCGCGGCTGCTGGACCCGGGCACCGGGAGGCGGCGGCGGCGGCAGCAGCAGC 148

Qy        169 AGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGA 213
              || |  ||||||| ||||| | |   ||||| | | ||| |||||
Db        147 AGCGGCGGCGGCGGCGGCGGCGGCGACAGCAGAAGAGGAAGAGGA 103


RESULT 63
US-09-925-065A-869021/c
; Sequence 869021, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 869021
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-869021

  Query Match             3.0%;  Score 39.2;  DB 6;  Length 610;
  Best Local Similarity   54.1%;  Pred. No. 3.5;
  Matches   80;  Conservative    0;  Mismatches   68;  Indels    0;  Gaps    0;

Qy         65 CGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATG 124
              ||||     |||  || ||  | | | || | |||  |      ||||  |  ||| |
Db        368 CGGCCGCCGGGGAGGGAGTGGTGAGGAGGGGGGGCCTGAGCGGGGGCGCGGGCGMCCTCC 309

Qy        125 GCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCG 184
              |  | ||| |  |  | | |||||| | ||||   |||| ||| |||  |||||||  |
Db        308 CCCGCGGGCGGGCGGGCGGGTGCGGCCGGGCCCCTCGGGCGGGCTGGGGCGGCGGCGCGG 249

Qy        185 GCGCCCGAAGCAGCACAGGCGGAGGAGG 212
              |  ||  ||    ||| |   |  ||
Db        248 CGGAGCGCGGCGCTGCAGCCATGGCGGG 221


RESULT 64
US-11-136-527-7671/c
; Sequence 7671, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7671
;   LENGTH: 1400
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-7671

  Query Match             3.0%;  Score 39.2;  DB 12;  Length 1400;
  Best Local Similarity   60.2%;  Pred. No. 5;
  Matches   65;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy        105 GAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGC 164
              |||| | ||  |   |  ||||    | |||  |||  |     |||| | | ||| | |
Db        127 GAGGAGGCCTGGTCACCCTGGCACTGGCCAGGTGTTGGGTCTGCGCGGCGGCAGCGGCGG 68

Qy        165 CGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGG 212
              |||| | |  ||||||| ||||  | |  ||||||  ||||| || ||
Db         67 CGGCGGCGGCGGCGGCGGCGGCAGCAGCGGCAGCAGCGGCGGCGGTGG 20


RESULT 65
US-11-136-527-3575/c
; Sequence 3575, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
```

```
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3575
;   LENGTH: 3050
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-3575

  Query Match             3.0%;  Score 39.2;  DB 12;  Length 3050;
  Best Local Similarity   60.2%;  Pred. No. 7;
  Matches   65;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy        105 GAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGC 164
              |||| | ||  |   | ||||   | |||  |||  |     |||| | | || | |
Db       1777 GAGGAGGCCTGGTCACCCTGGCACTGGCCAGGTGTTGGGTCTGCGCGGCGGCAGCGGCGG 1718

Qy        165 CGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGG 212
              |||| | |  |||||| ||||  | |  ||||||  ||||| || ||
Db       1717 CGGCGGCGGCGGCGGCGGCGGCAGCAGCGGCAGCAGCGGCGGCGGTGG 1670


RESULT 66
US-09-925-065A-765916/c
; Sequence 765916, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 765916
;   LENGTH: 619
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-765916

  Query Match             3.0%;  Score 39;  DB 6;  Length 619;
  Best Local Similarity   49.7%;  Pred. No. 4;
  Matches   99;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy        628 CTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTT 687
              |||   ||  ||  | || |  | | |  | | | |||  |  | | | |||  | | | ||   |
Db        267 CTATTCAACAAATGGTGCTGGAATCTGCTCAGCAACAAAAAGGAATAGACATGAATGAAT 208

Qy        688 CACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTT 747
              | |  ||| ||| ||  ||| ||||    ||   | ||   |  | ||  ||
Db        207 GTCTGAAATATTATAGAGTAAAAGAAGGAGTTAAGAGACAAAGGAATATATACTTTAAAA 148

Qy        748 TCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGA 807
              | ||| ||| |  | | | |  |||   | |  | |  |||| |||  |  || |  ||
Db        147 TTCAATTACTATGAAGTTCAAGAACAACCAAAACTTGTCTGTGGGTTCTGTGGGTACAGA 88

Qy        808 ATTTTGCCTTATGACATCC 826
              | || |  |  || |  ||
Db         87 AATTGGAATAGTGGCTACC 69


RESULT 67
US-09-925-065A-832620/c
; Sequence 832620, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 832620
;   LENGTH: 619
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-832620

  Query Match             3.0%;  Score 39;  DB 6;  Length 619;
  Best Local Similarity   49.7%;  Pred. No. 4;
  Matches   99;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy        628 CTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTT 687
              |||   ||  ||  | || |  | | |  | | | |||  |  | | | |||  | | | ||   |
Db        267 CTATTCAACAAATGGTGCTGGAATCTGCTCAGCAACAAAAAGGAATAGACATGAATGAAT 208

Qy        688 CACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTT 747
              | |  ||| ||| ||  ||| ||||    ||   | ||   |  | ||  ||
Db        207 GTCTGAAATATTATAGAGTAAAAGAAGGAGTTAAGAGACAAAGGAATATATACTTTAAAA 148

Qy        748 TCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGA 807
              | ||| ||| |  | | | |  |||   | |  | |  |||| |||  |  || |  ||
Db        147 TTCAATTACTATGAAGTTCAAGAACAACCAAAACTTGTCTGTGGGTTCTGTGGGTACAGA 88

Qy        808 ATTTTGCCTTATGACATCC 826
              | || |  |  || |  ||
Db         87 AATTGGAATAGTGGCTACC 69


RESULT 68
US-09-925-065A-832621/c
; Sequence 832621, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 832621
;   LENGTH: 619
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-832621

  Query Match             3.0%;  Score 39;  DB 6;  Length 619;
  Best Local Similarity   49.7%;  Pred. No. 4;
  Matches   99;  Conservative    0;  Mismatches  100;  Indels    0;  Gaps    0;

Qy        628 CTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTT 687
              |||   ||  ||  | || |  |  | | ||   |  | | | |||  | | | ||   |
Db        267 CTATTCAACAAATGGTGCTGGAATCTGCTCAGCAACAAAAAGGAATAGACATGAATGAAT 208

Qy        688 CACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTT 747
                | |  ||| ||| ||  ||| ||||      ||   | ||   |  | ||  ||
Db        207 GTCTGAAATATTATAGAGTAAAAGAAGGAGTTAAGAGACAAAGGAATATATACTTTAAAA 148

Qy        748 TCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGA 807
              | ||| ||| |  | | | |  |||   | |  | |   |||| ||| |  || |  ||
Db        147 TTCAATTACTATGAAGTTCAAGAACAACCAAAACTTGTCTGTGGGTTCTGTGGGTACAGA 88

Qy        808 ATTTTGCCTTATGACATCC 826
              | || |  |  || |  ||
Db         87 AATTGGAATAGTGGCTACC 69


RESULT 69
US-11-000-463-231
; Sequence 231, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Cao, Yi-Cheng
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 785CIP4CN
;  CURRENT APPLICATION NUMBER: US/11/000,463
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: 10/291,265
;  PRIOR FILING DATE: 2002-11-08
;  PRIOR APPLICATION NUMBER: PCT/US01/02623
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 09/617,746
;  PRIOR FILING DATE: 2000-07-17
;  PRIOR APPLICATION NUMBER: 09/631,451
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 09/633,870
;  PRIOR FILING DATE: 2000-09-15
;  NUMBER OF SEQ ID NOS: 944
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 231
;   LENGTH: 980
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (133)..(453)
US-11-000-463-231

  Query Match             3.0%;  Score 39;  DB 12;  Length 980;
  Best Local Similarity   71.8%;  Pred. No. 4.9;
  Matches   51;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy         41 ATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTG 100
              |||||||||||||||||||||||||||||||||||||||||   |  |   |    ||||  ||
Db          1 ATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGCTGGCTATCTACATGGAAGGGTAGGGT 60

Qy        101 TTTTGAGGGGA 111
              || ||| || |
Db         61 TTATGACGGCA 71


RESULT 70
US-11-096-568A-11538/c
; Sequence 11538, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 11538
;   LENGTH: 1653
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1653)
;   OTHER INFORMATION: Ceres Seq. ID no. 13655244
US-11-096-568A-11538

  Query Match             3.0%;  Score 39;  DB 9;  Length 1653;
  Best Local Similarity   58.0%;  Pred. No. 6.1;
  Matches   69;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy        114 CTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGC 173
              | | ||||  ||||    |   ||  ||||  ||| ||||  |||| ||  |||   ||
Db        200 CGGGGGCGGCGGCGGGGAGGGTCACCTCCGGCGCGTCGGAAGCCGCCGGCGCGGGGAGGG 141

Qy        174 GGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGG 232
              | || |  ||| | |||  |  |   ||| | ||||||    ||  | |||| || ||
Db        140 CGTCGCCTGCGGAGGCCGCCGTGGGGGAGGAGAAGGAGGGGTCGGCCATGAGGAGGGGG 82


RESULT 71
US-11-052-554A-539
; Sequence 539, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
;  APPLICANT: Sachdeva, et al.
;  TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
;  TITLE OF INVENTION:  PROTEINS OF THERAPEUTIC POTENTIAL
;  FILE REFERENCE: 30853/40359A
;  CURRENT APPLICATION NUMBER: US/11/052,554A
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/589,227
;  PRIOR FILING DATE: 2004-07-20
;  PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 763
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 539
;   LENGTH: 2196
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-539

  Query Match             3.0%;  Score 39;  DB 12;  Length 2196;
  Best Local Similarity   58.0%;  Pred. No. 6.9;
```

Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```
Qy        114 CTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGC 173
              | ||  ||  |||  |  |  | || |  ||||| ||| || | | |||| ||||
Db        456 CGGCAACGGCGGCAGCGCGGGGTTGATCGGCAACGGCGGGGCCGGCGGCGGCGGCGGGGC 515

Qy        174 GGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGG 232
              |||| || ||||| ||   || |||  |||||  |        ||  | |  | ||| |
Db        516 CGGCGCCGCCGGCGGCCTCGGCGGCAACGGCGGGTGGCTGTACGGCAACGGCGGGGCCG 574
```

RESULT 72
US-10-501-035-137
; Sequence 137, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 137
; LENGTH: 2833
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-501-035-137

Query Match 3.0%; Score 39; DB 7; Length 2833;
Best Local Similarity 49.7%; Pred. No. 7.6;
Matches 99; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
Qy         46 GCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTG 105
              ||||  |   || |||  |||||  ||    | |   ||||  |   | || | |    |
Db       1642 GCCCCAGGCTCCCAGAGCTCGGCCGGAACTGGAGCCTCAGCCAGGTCAACGTGAGCCACG 1701

Qy        106 AGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCC 165
               | |    | || ||||   |   |  | | ||  | |   || |  |||| || |  ||
Db       1702 CGCGCCTGCGGCCGCGACACCAGGTCCCCGAAGGCCTGCACCGCCTCAGCCTGCTGCTCC 1761

Qy        166 GGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAG 225
              || |   |||| ||  | |  | ||   ||||   | ||    |||||  |||  |
Db       1762 GGGACTCGGGGCAGCCGCCCCAGCAGCGCGAGCAGCCTCTGAACGTGACCGTGTGCCGCT 1821

Qy        226 GCGGCGGCTTCAGTGCCTG 244
              |||||     | | | |||
Db       1822 GCGGCAAGGACGGCGTCTG 1840
```

RESULT 73
US-11-000-688-780
; Sequence 780, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOULGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; CURRENT FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 780
; LENGTH: 2875
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences:primer
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2875)
; OTHER INFORMATION: cadherin 15, m-cadherin (myotubule)(CDH15)
; OTHER INFORMATION: gene.
US-11-000-688-780

Query Match 3.0%; Score 39; DB 12; Length 2875;
Best Local Similarity 49.7%; Pred. No. 7.7;
Matches 99; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
Qy         46 GCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTG 105
              ||||  |   || |||  |||||  ||    | |   ||||  |   | || | |    |
Db       1642 GCCCCAGGCTCCCAGAGCTCGGCCGGAACTGGAGCCTCAGCCAGGTCAACGTGAGCCACG 1701

Qy        106 AGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCC 165
               | |    | || ||||   |   |  | | ||  | |   || |  |||| || |  ||
Db       1702 CGCGCCTGCGGCCGCGACACCAGGTCCCCGAAGGCCTGCACCGCCTCAGCCTGCTGCTCC 1761

Qy        166 GGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAG 225
              || |   |||| ||  | |  | ||   ||||   | ||    |||||  |||  |
Db       1762 GGGACTCGGGGCAGCCGCCCCAGCAGCGCGAGCAGCCTCTGAACGTGACCGTGTGCCGCT 1821

Qy        226 GCGGCGGCTTCAGTGCCTG 244
              |||||     | | | |||
Db       1822 GCGGCAAGGACGGCGTCTG 1840
```

RESULT 74
US-11-052-554A-522
; Sequence 522, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 522
; LENGTH: 4146
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-522

Query Match 3.0%; Score 39; DB 12; Length 4146;
Best Local Similarity 48.4%; Pred. No. 9;
Matches 108; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```
Qy        104 TGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGG 163
              |||  ||   |  |||||  | || | | ||      |    | |  |  | |  | ||
Db        866 TGAACGGCGGCGACGGCGGCGCCGGCGGTCAAGGCGGCGACGGTGCGGCCGGCGACGCGG 925

Qy        164 CCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAG 223
              | | ||| || ||||||  |||||  |  | || ||  ||||||  |  ||       | ||
Db        926 CTGCCAGCGCCGGCGGCACCGGCGGCAAAGGCGGCCAAGGCGGCGACGGCGGCACCGGAG 985

Qy        224 AGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTC 283
              ||| |  |   |||  || | ||| | ||     || ||  |    |  |   | || |
Db        986 GGGCCGGCGGCGCAGGCCCAGTGCTGTTCGGCCACGGCGGCGCCGGCGGCATGGGCGGCC 1045

Qy        284 CCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCC 326
                  |  | |  |||     |    || | ||  ||| || |
Db       1046 AAGGCGGCACCGGTGGAATGGGCGGCGCCGGCGGAGACGGCAC 1088
```

```
RESULT 75
US-10-933-025-22
; Sequence 22, Application US/10933025
; Publication No. US20050265987A1
; GENERAL INFORMATION:
;  APPLICANT: ROSEN, STEVEN
;  APPLICANT:  HEMMERICH, STEFAN
;  APPLICANT:  TOMITA, MEGUMI
;  TITLE OF INVENTION: Sulfotransferases and methods of use
;  TITLE OF INVENTION:  thereof
;  FILE REFERENCE: UCAL-230CON
;  CURRENT APPLICATION NUMBER: US/10/933,025
;  CURRENT FILING DATE:  2004-09-01
;  PRIOR APPLICATION NUMBER: 10/025,966
;  PRIOR FILING DATE: 2001-12-21
;  PRIOR APPLICATION NUMBER: 60/258,577
;  PRIOR FILING DATE: 2000-12-27
;  PRIOR APPLICATION NUMBER: 60/267,831
;  PRIOR FILING DATE: 2001-09-02
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
;   LENGTH: 268685
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(268685)
;   OTHER INFORMATION: n = A,T,C or G
US-10-933-025-22

  Query Match             3.0%;  Score 39;  DB 8;  Length 268685;
  Best Local Similarity   54.0%;  Pred. No. 53;
  Matches  141;  Conservative    1;  Mismatches  116;  Indels    3;  Gaps    3;

Qy         45 GGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTT 104
              ||| :| | || |  | |   |||| ||||  | | |||||||      | |||
Db      67170 GGCGYTGGGGGGCGGGGAGGAGGCAGGAGGAGGAGCAGCAGCGCTCGCCGGGGGATGCAG 67229

Qy        105 GAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGC 164
              || ||   | ||||||  |||| | ||||||  |   ||  |  ||||| || | | ||
Db      67230 CAGCGGCAGCGGCGGCGGCGGCGGC-GGCAGCGGCAGCGGCGGCGCGGAGTCCCCTGCGC 67288

Qy        165 CGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGA 224
              |    || | ||| | | | ||| | || || ||   ||| |     |  |||||||| |
Db      67289 CCAGCGGCCCGGCCGGG-CTGCGGCAGAGGCGGCGGCGGC-GCCCCGCTCCGGGTGAGGT 67346

Qy        225 GGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCC 284
              || | |||       ||||    || |  |  ||||| | | |       |  || |
Db      67347 TGCCGGGGCCGTGCGCGCTGGTGAGTGGGGCGTGGGGGGCASGTGGGTGGCTGGGGGACG 67406

Qy        285 CCAGCGTCCTGCGGGAGAACG 305
              |  |  ||  ||||| |  ||
Db      67407 CGGGGCTCGGGCGGGGGCTCG 67427


Search completed: March 11, 2006, 16:22:32
Job time : 515 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2006, 15:38:24 ; Search time 1203 Seconds
(without alignments)
9018.640 Million cell updates/sec

Title: US-10-757-745-3
Perfect score: 1312
Sequence: 1 agctattaatgattcgaatt.....tctcaacctttcaggacatc 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1312	100.0	1312	9	US-10-757-745-3	Sequence 3, Appli
	2	668.2	50.9	1920	9	US-10-757-745-1	Sequence 1, Appli
	3	668.2	50.9	1948	5	US-10-037-270-889	Sequence 889, App
	4	668.2	50.9	1948	6	US-10-117-722-889	Sequence 889, App
	5	668.2	50.9	1948	9	US-10-122-851-889	Sequence 889, App
	6	668.2	50.9	3152	3	US-09-981-353-88	Sequence 88, Appl
	7	667	50.8	1296	3	US-09-925-299-170	Sequence 170, App
	8	667	50.8	1296	3	US-09-925-299-170	Sequence 170, App
	9	666.6	50.8	1940	9	US-10-783-271-28	Sequence 28, Appl
	10	493.2	37.6	1088	9	US-10-450-763-22058	Sequence 22058, A
c	11	490	37.3	1079	3	US-09-745-288-63	Sequence 63, Appl
c	12	490	37.3	1079	7	US-10-453-919-63	Sequence 63, Appl
	13	461.6	35.2	674	5	US-10-066-543-1067	Sequence 1067, Ap
	14	436.4	33.3	644	5	US-10-066-543-1070	Sequence 1070, Ap
	15	428.4	32.7	625	5	US-10-066-543-1031	Sequence 1031, Ap
	16	428.4	32.7	633	5	US-10-066-543-1084	Sequence 1084, Ap
	17	418	31.9	625	5	US-10-066-543-548	Sequence 548, App
	18	385.8	29.4	553	5	US-10-066-543-983	Sequence 983, App
	19	290.2	22.1	401	5	US-10-066-543-1363	Sequence 1363, Ap
	20	227.2	17.3	400	7	US-10-242-535A-8082	Sequence 8082, Ap
	21	227.2	17.3	400	7	US-10-085-783A-8082	Sequence 8082, Ap
	22	198.2	15.1	1227	9	US-10-450-763-22059	Sequence 22059, A
c	23	140.8	10.7	391	5	US-10-066-543-826	Sequence 826, App
	24	134.4	10.2	176	3	US-09-878-722-121	Sequence 121, App
c	25	134.4	10.2	176	3	US-09-878-178-50	Sequence 50, Appl
	26	134.4	10.2	176	3	US-09-904-456-121	Sequence 121, App
c	27	134.4	10.2	176	5	US-10-046-935-50	Sequence 50, Appl
c	28	134.4	10.2	176	5	US-10-146-502-50	Sequence 50, Appl
c	29	134.4	10.2	176	5	US-10-060-036-1367	Sequence 1367, Ap
	30	108.4	8.3	606	4	US-09-925-065A-783537	Sequence 783537,
	31	108.4	8.3	634	4	US-09-925-065A-783536	Sequence 783536,
c	32	101.4	7.7	268	3	US-09-876-889-312	Sequence 312, App
	33	74.6	5.7	444	3	US-09-864-761-16762	Sequence 16762, A
	34	74.6	5.7	691	3	US-09-864-761-33229	Sequence 33229, A
	35	65	5.0	499	6	US-10-115-831-61	Sequence 61, Appl
	36	63.8	4.9	690	7	US-10-404-460-149	Sequence 149, App
	37	63.2	4.8	696	7	US-10-404-460-154	Sequence 154, App
	38	63.2	4.8	704	7	US-10-404-460-105	Sequence 105, App
	39	62.2	4.7	1549	5	US-10-093-045-5	Sequence 5, Appli
	40	62.2	4.7	1549	5	US-10-093-246-5	Sequence 5, Appli
	41	61.6	4.7	701	7	US-10-404-460-60	Sequence 60, Appl
c	42	61.4	4.7	682	7	US-10-404-460-20	Sequence 20, Appl
	43	61.4	4.7	1775	5	US-10-098-841-26	Sequence 26, Appl
	44	61.2	4.7	956	2	US-08-816-011-52	Sequence 52, Appl
	45	61.2	4.7	956	8	US-10-870-492-52	Sequence 52, Appl
	46	60.4	4.6	692	7	US-10-404-460-121	Sequence 121, App
	47	60.4	4.6	706	7	US-10-404-460-231	Sequence 231, App
	48	60.2	4.6	657	7	US-10-404-460-209	Sequence 209, App
	49	59	4.5	686	7	US-10-404-460-138	Sequence 138, App
c	50	59	4.5	691	7	US-10-404-460-24	Sequence 24, Appl
	51	59	4.5	695	7	US-10-404-460-169	Sequence 169, App
c	52	58.8	4.5	627	7	US-10-404-460-201	Sequence 201, App
	53	58.8	4.5	688	7	US-10-404-460-214	Sequence 214, App
c	54	58.6	4.5	382	9	US-10-756-149-4637	Sequence 4637, Ap
	55	58.6	4.5	631	7	US-10-404-460-129	Sequence 129, App
	56	58.6	4.5	705	7	US-10-404-460-155	Sequence 155, App
	57	58.6	4.5	846	5	US-10-198-846-9716	Sequence 9716, Ap
	58	58.4	4.5	422	5	US-10-198-846-12993	Sequence 12993, A
c	59	58.2	4.4	519	9	US-10-756-149-4511	Sequence 4511, Ap
	60	58	4.4	633	7	US-10-404-460-264	Sequence 264, App
	61	58	4.4	666	7	US-10-404-460-6	Sequence 6, Appli
	62	57.8	4.4	670	7	US-10-404-460-27	Sequence 27, Appl
	63	57.8	4.4	670	7	US-10-404-460-207	Sequence 207, App
	64	57.8	4.4	671	7	US-10-404-460-93	Sequence 93, Appl
	65	57.8	4.4	677	7	US-10-404-460-196	Sequence 196, App
	66	57.8	4.4	689	7	US-10-404-460-187	Sequence 187, App
c	67	57.6	4.4	315	7	US-10-152-319A-915	Sequence 915, App
	68	57.6	4.4	708	7	US-10-404-460-221	Sequence 221, App
	69	57.6	4.4	728	7	US-10-404-460-137	Sequence 137, App
	70	57.4	4.4	663	7	US-10-404-460-166	Sequence 166, App
	71	57.4	4.4	705	7	US-10-404-460-163	Sequence 163, App
	72	57.2	4.4	422	7	US-10-404-460-150	Sequence 150, App
	73	57.2	4.4	690	7	US-10-404-460-123	Sequence 123, App
	74	57.2	4.4	719	7	US-10-404-460-230	Sequence 230, App
	75	57.2	4.4	3232	8	US-10-357-930-24319	Sequence 24319, A

ALIGNMENTS

RESULT 1
US-10-757-745-3
; Sequence 3, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/10/757,745
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/697,863A
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29

```
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (122)..(1234)
;   OTHER INFORMATION:
US-10-757-745-3

  Query Match           100.0%;  Score 1312;  DB 9;  Length 1312;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1312;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60

Qy         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120

Qy        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180

Qy        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240

Qy        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300

Qy        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360

Qy        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420

Qy        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480

Qy        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540

Qy        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600

Qy        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660

Qy        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720

Qy        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780

Qy        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840

Qy        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900

Qy        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960

Qy        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020

Qy       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080

Qy       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140

Qy       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200

Qy       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260

Qy       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312


RESULT 2
US-10-757-745-1
; Sequence 1, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1627)..(1627)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1108)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1849)..(1849)
;   OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

  Query Match           50.9%;  Score 668.2;  DB 9;  Length 1920;
  Best Local Similarity  77.0%;  Pred. No. 9.1e-180;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
```

```
Db        36 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 95

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        96 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 155

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db       156 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 215

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db       216 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 275

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db       276 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 335

Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db       336 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 395

Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db       396 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 455

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db       456 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 515

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db       516 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 575

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db       576 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 635

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db       636 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 695

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db       696 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 755

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db       756 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 815

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db       816 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 875

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db       876 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 935

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db       936 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 995

Qy      1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       996 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1055

Qy      1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db      1056 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1115

Qy      1242 CCCACTTGCAGCTTT 1256

Db      1116 TTCAAGTGTGGGTTT 1130
```

RESULT 3
US-10-037-270-889
; Sequence 889, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-037-270-889

Query Match 50.9%; Score 668.2; DB 5; Length 1948;
Best Local Similarity 75.6%; Pred. No. 9.2e-180;
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;

```
Qy       131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190

Db        15 GCCAAGAATTCGGCACGAGGGAAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAG 74

Qy       191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250

Db        75 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 134

Qy       251 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 310

Db       135 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 194

Qy       311 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 370

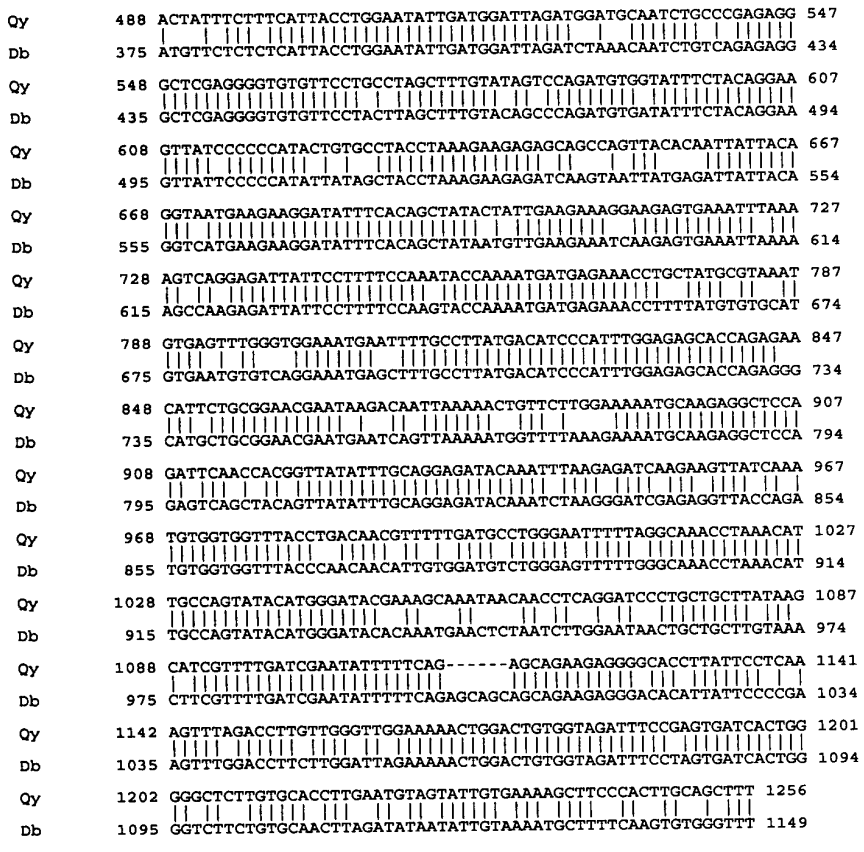
Db       195 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 254

Qy       371 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 430

Db       255 CGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 314

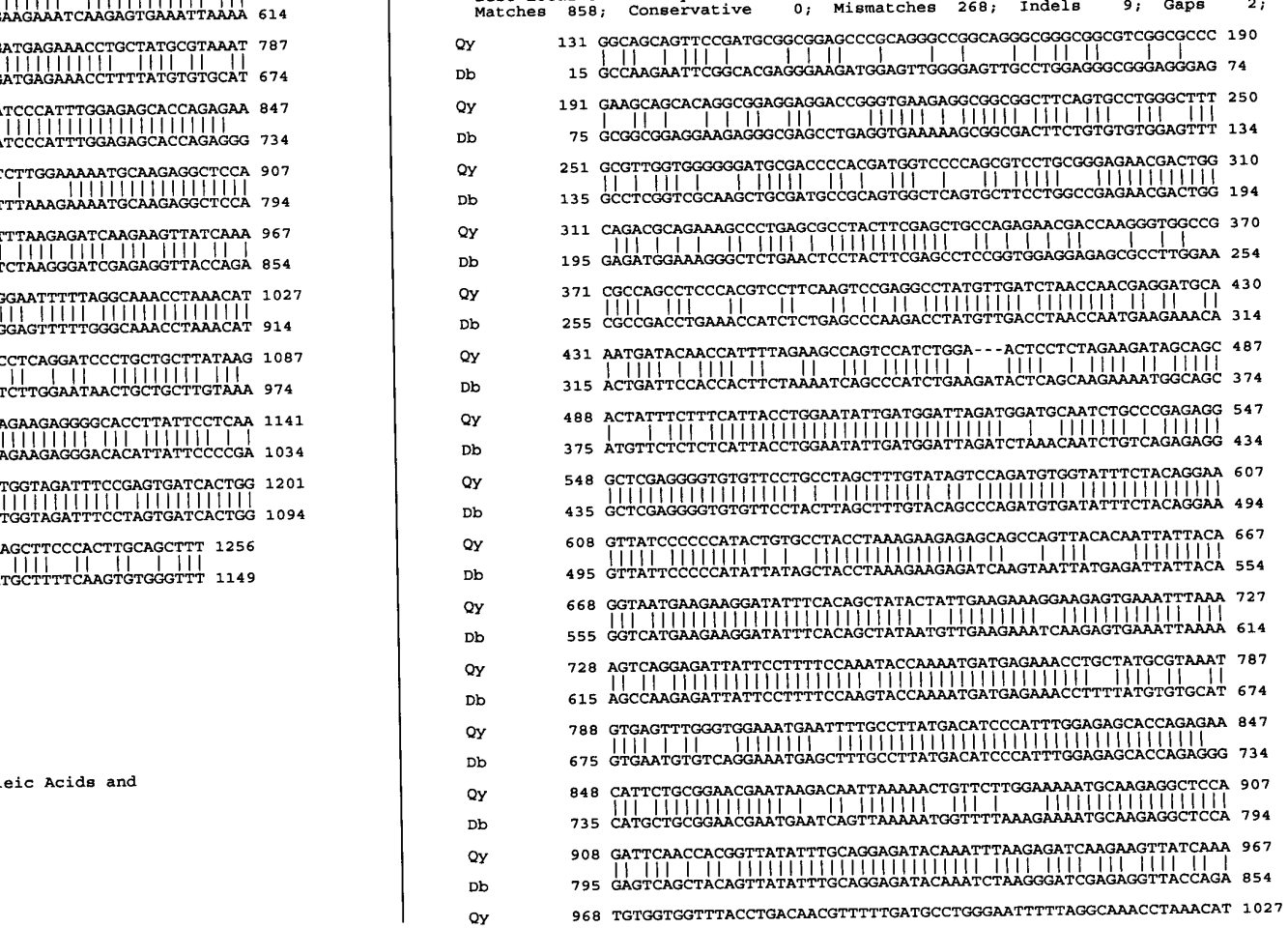
Qy       431 AATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGC 487

Db       315 ACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGC 374
```



RESULT 4
US-10-117-722-889
; Sequence 889, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-117-722-889

Query Match 50.9%; Score 668.2; DB 6; Length 1948;
Best Local Similarity 75.6%; Pred. No. 9.2e-180;
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;



```
Db        855 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 914

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087

Db        915 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 974

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141

Db        975 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 1034

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201

Db       1035 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 1094

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256

Db       1095 GGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 1149
```

RESULT 5
US-10-122-851-889
; Sequence 889, Application US/10122851
; Publication No. US20050239060A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BDV3
; CURRENT APPLICATION NUMBER: US/10/122,851
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-122-851-889

Query Match 50.9%; Score 668.2; DB 9; Length 1948;
Best Local Similarity 75.6%; Pred. No. 9.2e-180;
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;

```
Qy        131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190

Db         15 GCCAAGAATTCGGCACGAGGGAAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAG 74

Qy        191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250

Db         75 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 134

Qy        251 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 310

Db        135 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 194

Qy        311 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 370

Db        195 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 254

Qy        371 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 430

Db        255 CGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 314

Qy        431 AATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGC 487

Db        315 ACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGC 374

Qy        488 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 547

Db        375 ATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGG 434

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607

Db        435 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 494

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667

Db        495 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 554

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727

Db        555 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 614

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787

Db        615 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 674

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847

Db        675 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 734

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907

Db        735 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 794

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967

Db        795 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 854

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027

Db        855 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 914

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087

Db        915 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 974

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141

Db        975 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 1034

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201

Db       1035 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 1094

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256

Db       1095 GGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 1149
```

RESULT 6
US-09-981-353-88
; Sequence 88, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program

```
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020160382A1 232992.1
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

  Query Match            50.9%;  Score 668.2;  DB 3;  Length 3152;
  Best Local Similarity  77.0%;  Pred. No. 1.2e-179;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy       171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db      1268 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 1327

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db      1328 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 1387

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db      1388 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 1447

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db      1448 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 1507

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db      1508 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 1567

Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db      1568 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 1627

Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db      1628 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 1687

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db      1688 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 1747

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db      1748 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 1807

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db      1808 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 1867

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db      1868 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 1927

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db      1928 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1987

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db      1988 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 2047

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db      2048 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 2107

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db      2108 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 2167

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db      2168 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 2227

Qy      1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db      2228 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 2287

Qy      1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db      2288 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 2347

Qy      1242 CCCACTTGCAGCTTT 1256
Db      2348 TTCAAGTGTGGGTTT 2362


RESULT 7
US-09-925-299-170
; Sequence 170, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

  Query Match            50.8%;  Score 667;  DB 3;  Length 1296;
  Best Local Similarity  76.7%;  Pred. No. 1.6e-179;
  Matches  840;  Conservative    3;  Mismatches  243;  Indels    9;  Gaps    2;

Qy       171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db        40 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 99

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db       100 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 159

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db       160 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 219

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db       220 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 279

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db       280 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 339
```

```
Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db       340 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 399

Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db       400 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 459

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db       460 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 519

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db       520 GTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACAGCTATAATGTTGAAGA 579

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db       580 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 639

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db       640 GAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 699

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db       700 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 759

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db       760 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 819

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db       820 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 879

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db       880 TTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 939

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db       940 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 999

Qy      1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db      1000 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1059

Qy      1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db      1060 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1119

Qy      1242 CCCACTTGCAGCTTT 1256
Db      1120 TTCAAGTGTGGGTTT 1134
```

```
RESULT 8
US-09-925-299-170
; Sequence 170, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

  Query Match             50.8%;  Score 667;  DB 3;  Length 1296;
  Best Local Similarity   76.7%;  Pred. No. 1.6e-179;
  Matches  840;  Conservative    3;  Mismatches  243;  Indels    9;  Gaps    2;

Qy       171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db        40 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 99

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db       100 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 159

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db       160 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 219

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db       220 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 279

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db       280 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 339

Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db       340 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 399

Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db       400 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 459

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db       460 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 519

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db       520 GTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACAGCTATAATGTTGAAGA 579

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db       580 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 639

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db       640 GAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 699

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db       700 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 759

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db       760 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 819

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db       820 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 879
```

```
Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
             |||| ||:|||||||||||||||||||||||||||||||||   ||    ||      || ||
Db        880 TTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 939

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
             | ||  |||||||||| ||| | |||||||||||||||||||||||||      ||||||||
Db        940 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 999

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
             |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db       1000 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1059

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
             ||||||| |||||||||||||| ||  |||||| |||   || || |||||| ||| ||||
Db       1060 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1119

Qy       1242 CCCACTTGCAGCTTT 1256
                ||  ||  | |||
Db       1120 TTCAAGTGTGGGTTT 1134
```

RESULT 9
US-10-783-271-28
; Sequence 28, Application US/10783271
; Publication No. US20050186577A1
; GENERAL INFORMATION:
; APPLICANT: Veridex, LLC
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
; FILE REFERENCE: VDX-5003 USNP
; CURRENT APPLICATION NUMBER: US/10/783,271
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: human
US-10-783-271-28

Query Match 50.8%; Score 666.6; DB 9; Length 1940;
Best Local Similarity 76.9%; Pred. No. 2.6e-179;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |  | | ||  |||     |||||| | |||||
Db         42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  ||  |
Db        162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   |||| |||   ||   ||  || || || |||||||||
Db        222 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        282 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 341

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        342 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 401

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||| | |||||||||| || ||||
Db        402 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 461

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  |||||||||||||||| ||
Db        462 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 521

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              |  |||   |||||||||||| |||||||||||||||||||||||||| | |||||||
Db        522 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 581

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        582 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 641

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  ||||||   ||   ||||||||  ||||||||||||||||||
Db        642 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 701

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||  ||  |||||||  ||| |||||
Db        702 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 761

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  ||||| || | |||| ||||||| |
Db        822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||||||||||   ||    || ||
Db        882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||| | |||||||||||||||||||||||      ||||||||
Db        942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db       1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||   || || |||||| ||| ||||
Db       1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

Qy       1242 CCCACTTGCAGCTTT 1256
                ||  ||  | |||
Db       1122 TTCAAGTGTGGGTTT 1136
```

RESULT 10
US-10-450-763-22058
; Sequence 22058, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 22058
; LENGTH: 1088
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1)..(687)

; OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein
; OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
; OTHER INFORMATION: Waterman Score=1199.
US-10-450-763-22058

Query Match 37.6%; Score 493.2; DB 9; Length 1088;
Best Local Similarity 82.9%; Pred. No. 7.7e-130;
Matches 577; Conservative 0; Mismatches 113; Indels 6; Gaps 1;

```
Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||   | |||    |||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  ||||||||||| | |
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||| ||||||||||||||||||||||||||||||||| ||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| ||||||||||||| |  || |||||||  | | |    ||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || |||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||  ||   ||    || ||  | ||  |||||||||| |||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
              | |||||||||||||||||||||||      |||||||||||| ||| ||||||| | |
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
              ||||| |||||| |||| || |||||||||||||||||||||||| ||||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 660

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAA 1237
              || ||  |||||| |||  || || ||||   ||||
Db        661 GGTCTTCTGTGCAACTTAGATATAATATTTGTAAAA 696
```

RESULT 11
US-09-745-288-63/c
; Sequence 63, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446D1
; CURRENT APPLICATION NUMBER: US/09/745,288
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-745-288-63

Query Match 37.3%; Score 490; DB 3; Length 1079;
Best Local Similarity 82.8%; Pred. No. 6.3e-129;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              |||||||||||| | |||||||||| || ||||||||| ||||||||||||||||||| |
Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              ||||||| | |  ||||||||||||||||| ||   | |||    |||||||||| |||
Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              |||||||||||||||||||||||| | |||||||||  |||||||||| ||||| || |
Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              |||||||||||||||||| ||||||||||||||||||||||  |||| ||  |||||| |
Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||   ||||||||  |||||||||||||||||||||||||||||||||||||  ||| |||
Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              ||||||||||  |  || |||||||  ||| |    ||||||||||||||||||||| |||
Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || |||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              |||||||  ||||| || | |||| |||||| |||||||||||||||||||||||||||
Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              ||||||||||||  ||   || ||  | ||  |||||||| ||| | ||||
Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
              ||||||||||||||||||      |||||||||||| ||| ||||||| | |||||| |
Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
              ||||| |||| || |||||||||||||||||||||||| ||||||||||||| ||
Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              |||||| |||  || || |||||| ||| ||||  ||  ||  | |||
Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

RESULT 12
US-10-453-919-63/c
; Sequence 63, Application US/10453919
; Publication No. US20040033230A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C7
; CURRENT APPLICATION NUMBER: US/10/453,919

```
;  CURRENT FILING DATE:  2003-06-03
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-453-919-63

  Query Match             37.3%;  Score 490;  DB 7;  Length 1079;
  Best Local Similarity   82.8%;  Pred. No. 6.3e-129;
  Matches  586;  Conservative    0;  Mismatches  115;  Indels    7;  Gaps    2;

Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              ||||||||||||| | |||||||||| || |||||||| ||||||||||||||||||| |
Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              ||||||| | |  |||||||||||||||| ||   | |||    |||||||||||| |||
Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              |||||||||||||||||||||||| | |||||||||  |||||||||||| ||||| || |
Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              |||||||||||||||||| ||||||||||||||||||||  |||| ||  |||||| |
Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||   ||||||||  |||||||||||||||||||||||||||||||||||||  ||| |||
Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              |||||||||| |  || |||||||  ||| |    ||||||||||||||||||| |||
Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || ||||||||||||||||||||||||| |||| ||| |||| || ||||||||
Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              |||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||||||||
Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              ||||||||||||  ||   ||   || || | ||  |||||||| ||| | |||||
Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
              |||||||||||||||||      |||||||||| ||| ||||||| | ||||||| |
Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
              |||||  |||| || ||||||||||||||||||||||||||||| ||||||||||||| ||
Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              ||||||  |||  || || |||||| ||| |||| ||  ||  | |||
Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373


RESULT 13
US-10-066-543-1067
; Sequence 1067, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1067
;   LENGTH: 674
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11, 377, 640, 654
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1067

  Query Match             35.2%;  Score 461.6;  DB 5;  Length 674;
  Best Local Similarity   83.4%;  Pred. No. 6.5e-121;
  Matches  548;  Conservative    0;  Mismatches  102;  Indels    7;  Gaps    2;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || |||||||| ||||||||||||||
Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 77

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  |||||||||||||||| ||   | |||    |||||||||
Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||||| ||||||||||||||||||||||||  |||| ||  ||
Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              ||||  | ||   ||||||||  ||||||||||||||||||||||||||||||||||||||
Db        258 GTGAAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||| |  || |||||||  ||| |    ||||||||||||||||
Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        908 -GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAA 966
               || ||| | || |||||||||||||||||||||||| |||| |||| ||| |||| ||
Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        967 ATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACA 1026
              ||||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||
Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy       1027 TTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAA 1086
              |||||||||||||||||||||  ||    ||      || ||  | ||  |||||||||
Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

Qy       1087 GCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCA 1140
               | |||||||||||||||||||||||      |||||||||||| ||| ||||||| |
Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCA 1197
              |||||| |||||| |||| ||  |||||||||||| |||||||||||  ||||||||
Db        618 AAGTTTGGACCTTCTTGGATTANAAAAACTGGACTGNGGTAGATTTCCTAGTGATCA 674


RESULT 14
US-10-066-543-1070
```

```
; Sequence 1070, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1070
;   LENGTH: 644
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-066-543-1070

  Query Match             33.3%;  Score 436.4;  DB 5;  Length 644;
  Best Local Similarity   83.3%;  Pred. No. 1e-113;
  Matches  510;  Conservative    0;  Mismatches   96;  Indels    6;  Gaps    1;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||    | |||     |||||||||
Db         92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| |||||||||||||||||||||  |||| ||  ||
Db        212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   ||||||||  ||||||||||||||||||||||||||||||||||||
Db        272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||  |  || |||||||  ||| |    |||||| ||||||||||
Db        332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || |||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||  |||| || | |||| |||||| ||||| ||||||||||||||||
Db        452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              ||||||||||||||||||  ||   ||   || ||  | ||  |||||||| |||
Db        512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
              | |||||||||||||||||||||||      |||||||||||| ||| ||||||| | |
Db        572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy       1142 AGTTTAGACCTT 1153
              ||||| ||||||
Db        632 AGTTTGGACCTT 643


RESULT 15
US-10-066-543-1031
; Sequence 1031, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1031
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 10
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1031

  Query Match             32.7%;  Score 428.4;  DB 5;  Length 625;
  Best Local Similarity   84.1%;  Pred. No. 2e-111;
  Matches  483;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||    | |||     |||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| |||||||||||||||||||||  |||| ||  ||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   ||||||||  ||||||||||||||||||||||||||||||||||||
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||  |  || |||||||  ||| |    |||||| ||||||||||
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || |||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||  |||| || | |||| |||||| ||||| ||||||||||||||||
Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
```

```
                  |||||||||||||||||||  ||   ||    || ||  | ||  |||||||||| |||
Db        522 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 581

Qy       1088 CATCGTTTTGATCGAATATTTTTCAGAGCAGAAG 1121
                | |||||||||||||||||||||||||||||| ||
Db        582 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAG 615


RESULT 16
US-10-066-543-1084
; Sequence 1084, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1084
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1084

  Query Match             32.7%;  Score 428.4;  DB 5;  Length 633;
  Best Local Similarity   84.1%;  Pred. No. 2e-111;
  Matches  483;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||| ||    | |||    |||||||||
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  ||||||||||| |||
Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||||  |||| ||  ||
Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   ||||||||  ||||||||||||||||||||||||||||||||||||
Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||| |  || |||||||  ||| |  ||||||||||||||||||||
Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || ||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              ||||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||
Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||  ||   ||  || ||  | ||  ||||||||| |||
Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy       1088 CATCGTTTTGATCGAATATTTTTCAGAGCAGAAG 1121
              | |||||||||||||||||||||||||||||| ||
Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAG 607


RESULT 17
US-10-066-543-548
; Sequence 548, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 19, 71
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-548

  Query Match             31.9%;  Score 418;  DB 5;  Length 625;
  Best Local Similarity   84.0%;  Pred. No. 1.8e-108;
  Matches  483;  Conservative    0;  Mismatches   91;  Indels    1;  Gaps    1;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || |||||||| ||||||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||| ||    | |||    |||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy        668 -GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAA 726
               ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| ||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy        727 AAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAA 786
              ||| || |||||||||||||||||| |||||||||||||||||||  |||| ||  |
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy        787 TGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGA 846
              ||||| | ||   ||||||||  ||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        847 ACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCC 906
              ||| ||||||||||||| |  || |||||||  ||| |  ||||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385
```

```
Qy        907 AGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAA 966
              ||| ||| | || ||||||||||||||||||||||||| |||| |||| ||| |||| ||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        967 ATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACA 1026
              |||||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy       1027 TTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAA 1086
              |||||||||||||||||||||  ||   ||    || ||  | ||  ||||||||| |||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy       1087 GCATCGTTTTGATCGAATATTTTTCAGAGCAGAAG 1121
              | |||||||||||||||||||||||||||||| ||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAG 600
```

RESULT 18
US-10-066-543-983
; Sequence 983, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 983
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6, 35
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-983

Query Match 29.4%; Score 385.8; DB 5; Length 553;
Best Local Similarity 84.9%; Pred. No. 2.8e-99;
Matches 432; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```
Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | | |||||||||||||||| ||    | |||     |||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||| |||||||||||||||||||||||||||||||||||||
```

```
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||  | || |||||||  ||| |     |||||||||||||||||
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || ||||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||
Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAA 1056
              |||||||||||||||||||  ||   ||
Db        522 TGCCAGTATACATGGGATACACAAATGAA 550
```

RESULT 19
US-10-066-543-1363
; Sequence 1363, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1363
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1363

Query Match 22.1%; Score 290.2; DB 5; Length 401;
Best Local Similarity 84.9%; Pred. No. 5.2e-72;
Matches 325; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```
Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         19 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 78

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||    | |||     |||||||||
Db         79 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 138

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        139 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 198

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        199 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||| |||||||||||||||||||||||||||||||||||||
Db        259 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 318

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
```

```
              ||| |||||||||||||| |  || |||||||  ||| |    ||||||||||||||||||
Db        319 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 378

Qy        908 GATTCAACCACGGTTATATTTGC 930
              || ||| | || ||||||||||
Db        379 GAGTCAGCTACAGTTATATTTGC 401


RESULT 20
US-10-242-535A-8082
; Sequence 8082, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8082
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-8082

  Query Match             17.3%;  Score 227.2;  DB 7;  Length 400;
  Best Local Similarity   79.3%;  Pred. No. 5.5e-54;
  Matches  284;  Conservative    0;  Mismatches   68;  Indels    6;  Gaps    1;

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||| |  |||  | |     |  ||||||||||||||||||||||||||||  ||| |||
Db         42 TGGATCCAAAGAATTCGGCACGAGTGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 101

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              |||||||||| |  || |||||||  ||| |    ||||||||||||||||||||| |||
Db        102 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 161

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || ||||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db        162 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 221

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              |||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||||||||
Db        222 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 281

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              |||||||||||||  ||   ||    || ||  | ||  ||||||||| ||| | |||||
Db        282 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 341

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTT 1146
              ||||||||||||||||||      |||||||||||| ||| ||||||| | ||||||
Db        342 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT 399


RESULT 21
US-10-085-783A-8082
; Sequence 8082, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8082
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-8082

  Query Match             17.3%;  Score 227.2;  DB 7;  Length 400;
  Best Local Similarity   79.3%;  Pred. No. 5.5e-54;
  Matches  284;  Conservative    0;  Mismatches   68;  Indels    6;  Gaps    1;

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||| |  |||  | |     |  ||||||||||||||||||||||||||||  ||| |||
Db         42 TGGATCCAAAGAATTCGGCACGAGTGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 101

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              |||||||||| |  || |||||||  ||| |    ||||||||||||||||||||| |||
Db        102 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 161

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || ||||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db        162 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 221

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              |||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||||||||
Db        222 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 281

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              |||||||||||||  ||   ||    || ||  | ||  ||||||||| ||| | |||||
Db        282 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 341

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTT 1146
              ||||||||||||||||||      |||||||||||| ||| ||||||| | ||||||
Db        342 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT 399


RESULT 22
US-10-450-763-22059
; Sequence 22059, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22059
;   LENGTH: 1227
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (84)..(203)
;   OTHER INFORMATION: 85% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=164.
```

US-10-450-763-22059

Query Match 15.1%; Score 198.2; DB 9; Length 1227;
Best Local Similarity 70.3%; Pred. No. 2.1e-45;
Matches 282; Conservative 0; Mismatches 113; Indels 6; Gaps 1;

```
Qy        851 TCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGAT 910
              | | |||||||||| |  || |||||||  ||| |    ||||||||||||||||||||
Db        554 TATACGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAG 613

Qy        911 TCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGT 970
              ||| | || ||||||||||||||||||||||||| |||| |||| ||| || |   |
Db        614 TCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTGAGTGACAAA 673

Qy        971 GGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGC 1030
              |   ||||        | |                ||     |    |      ||||
Db        674 GTCAGTTTGATGTTTGAGGACAGATTATTGCTATGGAGACTAATTTAGGTTCCTTTTTGC 733

Qy       1031 CAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCAT 1090
              |||||||||||||||||  ||   ||    || ||  | ||  |||||||| ||| | |
Db        734 CAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTT 793

Qy       1091 CGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGT 1144
              ||||||||||||||||||||||      |||||||||||| ||| ||||||| | ||||
Db        794 CGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGT 853

Qy       1145 TTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGG 1204
              || |||||| |||| || ||||||||||||||||||||||||| ||||||||||||||
Db        854 TTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGT 913

Qy       1205 CTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCA 1245
              ||  |||||| |||  || || |||||| ||| ||||  ||
Db        914 CTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCA 954
```

RESULT 23
US-10-066-543-826/c
; Sequence 826, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 826
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 373
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-826

Query Match 10.7%; Score 140.8; DB 5; Length 391;
Best Local Similarity 69.6%; Pred. No. 2.9e-29;
Matches 206; Conservative 0; Mismatches 87; Indels 3; Gaps 1;

```
Qy        262 GGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAA 321
              | | |||||  || |  |||  |   || |||||   |||||||||||| ||| | | |
Db        338 GAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAG 279

Qy        322 AGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCC 381
              || |||| | ||||||||||||  || | | | ||     | | |    ||||  |||
Db        278 GGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGA 219

Qy        382 CACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAAC 441
              ||    ||   || || || |||||||||| |||||||| || ||  ||| |||| | ||
Db        218 AACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCAC 159

Qy        442 CATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTT 498
              || ||   ||  ||| ||||||| |    ||||  | |||| || ||||||    | ||| |
Db        158 CACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCT 99

Qy        499 CATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAG 554
              |||||||||||||||||||||||||||  | |||||||| | ||||||||||||
Db         98 CATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAG 43
```

RESULT 24
US-09-878-722-121
; Sequence 121, Application US/09878722
; Publication No. US20020040127A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Hepler, William T.
; APPLICANT: Clapper, Jonathan
; APPLICANT: Wang, Aijun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.524
; CURRENT APPLICATION NUMBER: US/09/878,722
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 245
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-722-121

Query Match 10.2%; Score 134.4; DB 3; Length 176;
Best Local Similarity 85.2%; Pred. No. 1.2e-27;
Matches 150; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy        579 ATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGA 638
              | || |||||||| |||||||||||||||||| |||||||  | |  ||||||||||
Db          1 ACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGA 60

Qy        639 AGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATAC 698
              ||||| ||   | |||    |||||||||||| ||||||||||||||||||||||||||
Db         61 AGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAA 120

Qy        699 TATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              | |||||||||  |||||||||||| |||||  || ||||||||||||||||||| |
Db        121 TGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176
```

RESULT 25
US-09-878-178-50/c
; Sequence 50, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08

```
;  NUMBER OF SEQ ID NOS: 2237
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-878-178-50

  Query Match             10.2%;  Score 134.4;  DB 3;  Length 176;
  Best Local Similarity   85.2%;  Pred. No. 1.2e-27;
  Matches  150;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy        579 ATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGA 638
              | || ||||||||| |||||||||||||||||||| |||||||| | |  ||||||||||
Db        176 ACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGA 117

Qy        639 AGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATAC 698
              ||||| ||    | |||     |||||||||||| ||||||||||||||||||||||||
Db        116 AGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAA 57

Qy        699 TATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              | |||||||||  |||||||||||| ||||| || ||||||||||||||||||| |
Db         56 TGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1


RESULT 26
US-09-904-456-121
; Sequence 121, Application US/09904456
; Publication No. US20030017167A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.524C1
;  CURRENT APPLICATION NUMBER: US/09/904,456
;  CURRENT FILING DATE:  2001-07-11
;  NUMBER OF SEQ ID NOS: 247
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-904-456-121

  Query Match             10.2%;  Score 134.4;  DB 3;  Length 176;
  Best Local Similarity   85.2%;  Pred. No. 1.2e-27;
  Matches  150;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy        579 ATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGA 638
              | || ||||||||| |||||||||||||||||||| |||||||| | |  ||||||||||
Db          1 ACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGA 60

Qy        639 AGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATAC 698
              ||||| ||    | |||     |||||||||||| ||||||||||||||||||||||||
Db         61 AGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAA 120

Qy        699 TATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              | |||||||||  |||||||||||| ||||| || ||||||||||||||||||| |
Db        121 TGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176


RESULT 27
US-10-046-935-50/c
; Sequence 50, Application US/10046935
; Publication No. US20020156011A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C1
;  CURRENT APPLICATION NUMBER: US/10/046,935
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2239
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-046-935-50

  Query Match             10.2%;  Score 134.4;  DB 5;  Length 176;
  Best Local Similarity   85.2%;  Pred. No. 1.2e-27;
  Matches  150;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy        579 ATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGA 638
              | || ||||||||| |||||||||||||||||||| |||||||| | |  ||||||||||
Db        176 ACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGA 117

Qy        639 AGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATAC 698
              ||||| ||    | |||     |||||||||||| ||||||||||||||||||||||||
Db        116 AGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAA 57

Qy        699 TATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              | |||||||||  |||||||||||| ||||| || ||||||||||||||||||| |
Db         56 TGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1


RESULT 28
US-10-146-502-50/c
; Sequence 50, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C2
;  CURRENT APPLICATION NUMBER: US/10/146,502
;  CURRENT FILING DATE:  2002-05-14
;  NUMBER OF SEQ ID NOS: 2241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-146-502-50

  Query Match             10.2%;  Score 134.4;  DB 5;  Length 176;
  Best Local Similarity   85.2%;  Pred. No. 1.2e-27;
  Matches  150;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy        579 ATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGA 638
              | || ||||||||| |||||||||||||||||||| |||||||| | |  ||||||||||
Db        176 ACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGA 117

Qy        639 AGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATAC 698
              ||||| ||    | |||     |||||||||||| ||||||||||||||||||||||||
Db        116 AGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAA 57

Qy        699 TATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              | |||||||||  |||||||||||| ||||| || ||||||||||||||||||| |
Db         56 TGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1


RESULT 29
US-10-060-036-1367/c
; Sequence 1367, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Benson, Darin R.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Persing, David H.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF PANCREATIC CANCER
;  FILE REFERENCE: 210121.566
;  CURRENT APPLICATION NUMBER: US/10/060,036
;  CURRENT FILING DATE:  2002-01-30
;  NUMBER OF SEQ ID NOS: 4560
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1367
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-060-036-1367

  Query Match             10.2%;  Score 134.4;  DB 5;  Length 176;
  Best Local Similarity   85.2%;  Pred. No. 1.2e-27;
  Matches  150;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy        579 ATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGA 638
              | || |||||||||  |||||||||||||||||||| ||||||||  | |  ||||||||||
Db        176 ACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGA 117

Qy        639 AGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATAC 698
              |||||  ||    | |||    |||||||||||| ||||||||||||||||||||||||||
Db        116 AGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAA 57

Qy        699 TATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              | |||||||||  |||||||||||| ||||| || ||||||||||||||||||  |
Db         56 TGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1


RESULT 30
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783537

  Query Match             8.3%;  Score 108.4;  DB 4;  Length 606;
  Best Local Similarity   78.6%;  Pred. No. 6.8e-20;
  Matches  143;  Conservative    0;  Mismatches   36;  Indels    3;  Gaps    1;

Qy        401 GCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCA 460
              |||  |||||| |||||||| || ||  ||| |||| | |||| ||    ||  ||| |||
Db        408 GCCAGTGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCA 467

Qy        461 TCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGAT 517
              ||||  |    ||||  | |||| || ||||||    | ||| ||||||||||||||||||
Db        468 TCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGAT 527

Qy        518 GGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTG 577
              |||||||||  |  ||||||| | ||||||||||||||||||||||||| | ||||||||
Db        528 GGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTG 587

Qy        578 TA 579
              ||
Db        588 TA 589


RESULT 31
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783536

  Query Match             8.3%;  Score 108.4;  DB 4;  Length 634;
  Best Local Similarity   78.6%;  Pred. No. 7e-20;
  Matches  143;  Conservative    0;  Mismatches   36;  Indels    3;  Gaps    1;

Qy        401 GCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCA 460
              |||  |||||| |||||||| || ||  ||| |||| | |||| ||    ||  ||| |||
Db        259 GCCAGTGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCA 318

Qy        461 TCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGAT 517
              ||||  |    ||||  | |||| || ||||||    | ||| ||||||||||||||||||
Db        319 TCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGAT 378

Qy        518 GGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTG 577
              |||||||||  |  ||||||| | ||||||||||||||||||||||||| | ||||||||
Db        379 GGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTG 438

Qy        578 TA 579
              ||
Db        439 TA 440


RESULT 32
US-09-876-889-312/c
; Sequence 312, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
;  APPLICANT: Benson, Darin R.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
```

```
;  TITLE OF INVENTION:  CANCER THERAPY AND DIAGNOSIS
;  FILE REFERENCE: 210121.466C3
;  CURRENT APPLICATION NUMBER: US/09/876,889
;  CURRENT FILING DATE:  2001-06-06
;  NUMBER OF SEQ ID NOS: 353
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 312
;   LENGTH: 268
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-876-889-312

  Query Match             7.7%;  Score 101.4;  DB 3;  Length 268;
  Best Local Similarity   81.8%;  Pred. No. 4.3e-18;
  Matches  117;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1114 AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGA 1173
             ||||||||||||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||
Db        262 AGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGA 203

Qy       1174 CTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTG 1233
             |||||||||||||||| |||||||||||||| ||  |||||| |||  || || ||||||
Db        202 CTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTA 143

Qy       1234 AAAAGCTTCCCACTTGCAGCTTT 1256
              ||| |||| ||  ||  | |||
Db        142 AAATGCTTTTCAAGTGTGGGTTT 120


RESULT 33
US-09-864-761-16762
; Sequence 16762, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16762
;   LENGTH: 444
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
US-09-864-761-16762

  Query Match             5.7%;  Score 74.6;  DB 3;  Length 444;
  Best Local Similarity   71.3%;  Pred. No. 2.7e-10;
  Matches  112;  Conservative    0;  Mismatches   44;  Indels    1;  Gaps    1;

Qy        577 GTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAA 636
              ||| || ||||||||| ||||||||||||||||||| |||||||| | |  |||||||||
Db        253 GTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAA 312

Qy        637 GAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTAT 696
              ||||||| ||   | |||    ||||||||||||||  |     | |||| ||  | ||
Db        313 GAAGAGATCAAGTAATTATGAGATTATTACAGGTAACAATTTTTGTTATTCAACTTGTCT 372

Qy        697 ACT-ATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 732
              |  ||||| ||  ||  |||||||   | || |||
Db        373 TTTCTTTGAAAAACAGATCAGTGAAAGAGATAAATCA 409


RESULT 34
US-09-864-761-33229
; Sequence 33229, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
```

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33229
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
;   OTHER INFORMATION: SWISSPROT HIT: P18004, EVALUE 7.00e-01
;   OTHER INFORMATION: NT HIT: AF201687.1, EVALUE 3.00e-45
;   OTHER INFORMATION: EST_HUMAN HIT: AV717253.1, EVALUE 5.00e-45
US-09-864-761-33229

  Query Match             5.7%;  Score 74.6;  DB 3;  Length 691;
  Best Local Similarity   71.3%;  Pred. No. 3.4e-10;
  Matches  112;  Conservative    0;  Mismatches   44;  Indels    1;  Gaps    1;

Qy        577 GTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAA 636
              ||| || |||||||||| |||||||||||||||||||| |||||||| | |  |||||||||
Db        115 GTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAA 174

Qy        637 GAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTAT 696
              ||||||| ||    | |||     |||||||||||||| |      | ||||  ||   | |
Db        175 GAAGAGATCAAGTAATTATGAGATTATTACAGGTAACAATTTTTGTTATTCAACTTGTCT 234

Qy        697 ACT-ATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 732
                |  ||||| ||  ||  |||||||   | || |||
Db        235 TTTCTTTGAAAAACAGATCAGTGAAAGAGATAAATCA 271


RESULT 35
US-10-115-831-61
; Sequence 61, Application US/10115831
; Publication No. US20030219743A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 792CIP2ADIV
;  CURRENT APPLICATION NUMBER: US/10/115,831
;  CURRENT FILING DATE:  2002-04-02
;  PRIOR APPLICATION NUMBER: 09/667,298
;  PRIOR FILING DATE: 2000-09-22
;  PRIOR APPLICATION NUMBER: 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 61
;   LENGTH: 499
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (93)..(344)
US-10-115-831-61

  Query Match             5.0%;  Score 65;  DB 6;  Length 499;
  Best Local Similarity   67.2%;  Pred. No. 1.6e-07;
  Matches   92;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy         22 ATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAA 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db          1 ATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGTGTTCA 60

Qy         82 GCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTC 141
              | | |      || |||   |    | |   ||| |||     | |||||      ||
Db         61 GTTGTGGCGGTTGCTGGTCAGTAACAGCCAAGATGCTGCGGAATCTGCTGGCTCTTCGTC 120

Qy        142 CGATGCGGCGGAGCCCG 158
               |||  ||| |||  ||
Db        121 AGATTGGGCAGAGGACG 137


RESULT 36
US-10-404-460-149
; Sequence 149, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 149
;   LENGTH: 690
;   TYPE:- DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
US-10-404-460-149

  Query Match             4.9%;  Score 63.8;  DB 7;  Length 690;
  Best Local Similarity   61.8%;  Pred. No. 4.2e-07;
  Matches  118;  Conservative    0;  Mismatches   72;  Indels    1;  Gaps    1;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGAT-GGCGTCTGGCAG 135
              | |    ||| |  | |  |  |     ||  |||| | ||  |  |  || | |
Db         77 CGAACCTCGCGGGGCGGGAAGCCGCGAGATGGACACCCCTCCGCTCTCAGACTCGGACTC 136

Qy        136 CAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGC 195
              | | || |||| |     | || |||  |  |    || ||  ||||   | | || |
Db        137 CGGGTCGGATGAGTGCCTGGCCTCAGATCAAGAGTTGCAGGATGCGTTTTCCCGCGGACT 196

Qy        196 AGCACAGGCGG 206
              | || | |
Db        197 CCTAAAGCCAG 207


RESULT 37
US-10-404-460-154
; Sequence 154, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
```

```
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 154
;    LENGTH: 696
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: predictive genes for liver necrosis
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: 604, 613, 678
;    OTHER INFORMATION: n = A,T,C or G
US-10-404-460-154

  Query Match             4.8%;  Score 63.2;  DB 7;  Length 696;
  Best Local Similarity   84.5%;  Pred. No. 6.3e-07;
  Matches   71;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              | ||||| || |    | || ||
Db         76 GAGAAGCTGCCTTCTCATTGGCTG 99


RESULT 38
US-10-404-460-105
; Sequence 105, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 105
;    LENGTH: 704
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: predictive genes for liver necrosis
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: 630, 650, 673, 694
;    OTHER INFORMATION: n = A,T,C or G
US-10-404-460-105

  Query Match             4.8%;  Score 63.2;  DB 7;  Length 704;
  Best Local Similarity   84.5%;  Pred. No. 6.4e-07;
  Matches   71;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAT 77

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              ||| |   ||| || |||   |||
Db         78 GGGTATGTGCGGGATGAGACAGTG 101


RESULT 39
US-10-093-045-5
; Sequence 5, Application US/10093045
; Publication No. US20020151708A1
; GENERAL INFORMATION:
;  APPLICANT: Guerriero, Vincent
;  APPLICANT: Raynes, Deborah A
;  TITLE OF INVENTION: DNA Encoding Proteins That Inhibit Hsp70 Function
;  FILE REFERENCE: HspBP DNA and Protein Sequences
;  CURRENT APPLICATION NUMBER: US/10/093,045
;  CURRENT FILING DATE: 2002-03-07
;  EARLIER APPLICATION NUMBER: 09/444,336
;  EARLIER FILING DATE: 1999-11-19
;  EARLIER APPLICATION NUMBER: 60/109,351
;  EARLIER FILING DATE: 1998-11-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;    LENGTH: 1549
;    TYPE: DNA
;    ORGANISM: Mus musculus
US-10-093-045-5

  Query Match             4.7%;  Score 62.2;  DB 5;  Length 1549;
  Best Local Similarity   95.5%;  Pred. No. 1.9e-06;
  Matches   64;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGG 78
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          2 TTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCCGG 61

Qy         79 GAAGCAG 85
              |||||||
Db         62 CAAGCAG 68


RESULT 40
US-10-093-246-5
; Sequence 5, Application US/10093246
; Publication No. US20030023071A1
; GENERAL INFORMATION:
;  APPLICANT: Guerriero, Vincent
;  APPLICANT:  Raynes, Deborah A
;  TITLE OF INVENTION: DNA Encoding Proteins That Inhibit Hsp70 Function
;  FILE REFERENCE: HspBP DNA and Protein Sequences
;  CURRENT APPLICATION NUMBER: US/10/093,246
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: 09/444,336
;  PRIOR FILING DATE: 1999-11-19
;  PRIOR APPLICATION NUMBER: 60/109,351
;  PRIOR FILING DATE: 1998-11-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;    LENGTH: 1549
;    TYPE: DNA
;    ORGANISM: Mus musculus
US-10-093-246-5

  Query Match             4.7%;  Score 62.2;  DB 5;  Length 1549;
  Best Local Similarity   95.5%;  Pred. No. 1.9e-06;
  Matches   64;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGG 78
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          2 TTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCCGG 61

Qy         79 GAAGCAG 85
              |||||||
Db         62 CAAGCAG 68
```

```
RESULT 41
US-10-404-460-60
; Sequence 60, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60
;   LENGTH: 701
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 606,701
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-60

  Query Match             4.7%;  Score 61.6;  DB 7;  Length 701;
  Best Local Similarity   56.4%;  Pred. No. 1.8e-06;
  Matches  115;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
              || |   |      || |   | |   |   |   |||  |   | | |    |
Db         77 GGAACCATGGCCAGCCCGCTGCGCTCCTTGATGCTACTGCTGGCCGTCCTGGCCGTGGCC 136

Qy        137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
              | |||   |  |   |||  |   ||        | | ||  ||  || | | || ||
Db        137 TGGGCCGGAACCTCCAGGCCACCCCCGCGATTGTTGGGAGCTCCGCAGGAGGCAGATGCC 196

Qy        197 GCACAGGCGGAGGAGGACCGGGTG 220
                 ||| ||  | | | || | |
Db        197 AGCGAGGAGGGCGTGCAGCGAGCG 220


RESULT 42
US-10-404-460-20/c
; Sequence 20, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;   LENGTH: 682
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 675,677
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-20

  Query Match             4.7%;  Score 61.4;  DB 7;  Length 682;
  Best Local Similarity   81.6%;  Pred. No. 2e-06;
  Matches   71;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          7 TAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCG 66
              | || |  | | || ||||||||||||||||||||||||||||||||||||||||||||
Db        672 TGATTACGCAATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCG 613

Qy         67 GCACGAGGGCGGGAAGCAGCGTGAAGA 93
              ||||||||| | | |  |   || |||
Db        612 GCACGAGGGTGAGGAAGAAGATGGAGA 586


RESULT 43
US-10-098-841-26
; Sequence 26, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Xu, Chongjun
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Qian, Xiaohong B.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2
;  CURRENT APPLICATION NUMBER: US/10/098,841
;  CURRENT FILING DATE:  2002-03-13
;  PRIOR APPLICATION NUMBER: 09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 331
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 26
;   LENGTH: 1775
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (397)..(1176)
US-10-098-841-26

  Query Match             4.7%;  Score 61.4;  DB 5;  Length 1775;
  Best Local Similarity   91.5%;  Pred. No. 3.5e-06;
  Matches   65;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          8 AATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGG 67
              | ||| |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          4 ACTGACTCGAATTTTTACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGG 63
```

```
Qy          68 CACGAGGGCGG 78
               ||||||   ||
Db          64 CACGAGATTGG 74


RESULT 44
US-08-816-011-52
; Sequence 52, Application US/08816011
; Publication No. US20030165806A1
; GENERAL INFORMATION:
;    APPLICANT:  Price, Laura A.
;    APPLICANT:  Pausch, Mark H.
;    TITLE OF INVENTION:  Potassium Channels, Nucleotide Sequences
;    TITLE OF INVENTION:  Encoding Them, and Methods of Using Same
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  American Home Products Corporation
;      STREET:  One Campus Drive
;      CITY:  Parsippany
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/816,011
;      FILING DATE:  11-MAR-1997
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Matthews, Gale F.
;      REGISTRATION NUMBER:  32,269
;      REFERENCE/DOCKET NUMBER:  32,421-C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-683-2134
;      TELEFAX:  201-683-4117
;  INFORMATION FOR SEQ ID NO:  52:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  956 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-816-011-52

  Query Match             4.7%;  Score 61.2;  DB 2;  Length 956;
  Best Local Similarity   95.5%;  Pred. No. 2.8e-06;
  Matches   63;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy           9 ATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 68
               ||||| ||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGATACGATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 60

Qy          69 ACGAGG 74
               ||||||
Db          61 ACGAGG 66


RESULT 45
US-10-870-492-52
; Sequence 52, Application US/10870492
; Publication No. US20050032165A1
; GENERAL INFORMATION:
;  APPLICANT: PAUSCH, MARK H.
;  TITLE OF INVENTION: POTASSIUM CHANNELS, NUCLEOTIDE SEQUENCES ENCODING THEM,
;  TITLE OF INVENTION:  AND METHODS OF USING SAME
;  FILE REFERENCE: 01142.0114 SEQUENCE LISTING
;  CURRENT APPLICATION NUMBER: US/10/870,492
;  CURRENT FILING DATE:  2004-06-18
;  PRIOR APPLICATION NUMBER: US/09/503,849
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 08/816,011
;  PRIOR FILING DATE: 1997-03-11
;  PRIOR APPLICATION NUMBER: PCT/US95/14364
;  PRIOR FILING DATE: 1995-10-25
;  PRIOR APPLICATION NUMBER: 07/332,312
;  PRIOR FILING DATE: 1994-10-31
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
;   LENGTH: 956
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-870-492-52

  Query Match             4.7%;  Score 61.2;  DB 8;  Length 956;
  Best Local Similarity   95.5%;  Pred. No. 2.8e-06;
  Matches   63;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy           9 ATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 68
               ||||| ||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGATACGATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 60

Qy          69 ACGAGG 74
               ||||||
Db          61 ACGAGG 66


RESULT 46
US-10-404-460-121
; Sequence 121, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
;   LENGTH: 692
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
US-10-404-460-121

  Query Match             4.6%;  Score 60.4;  DB 7;  Length 692;
  Best Local Similarity   91.4%;  Pred. No. 4e-06;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
               | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 77

Qy          77 GGGAAGCAGC 86
               ||  ||||
Db          78 TAGAGCCAGC 87


RESULT 47
US-10-404-460-231
; Sequence 231, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
```

```
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 231
;   LENGTH: 706
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
US-10-404-460-231

  Query Match             4.6%;  Score 60.4;  DB 7;  Length 706;
  Best Local Similarity   91.4%;  Pred. No. 4e-06;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGG 76

Qy         77 GGGAAGCAGC 86
                ||||  |||
Db         77 AAGAAGAAGC 86


RESULT 48
US-10-404-460-209
; Sequence 209, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 209
;   LENGTH: 657
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 567, 657
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-209

  Query Match             4.6%;  Score 60.2;  DB 7;  Length 657;
  Best Local Similarity   89.0%;  Pred. No. 4.4e-06;
  Matches   65;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 77

Qy         77 GGGAAGCAGCGTG 89
              | | ||  | | |
Db         78 GCGCAGGGGAGCG 90


RESULT 49
US-10-404-460-138
; Sequence 138, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 138
;   LENGTH: 686
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 522, 559, 589, 650, 655
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-138

  Query Match             4.5%;  Score 59;  DB 7;  Length 686;
  Best Local Similarity   78.0%;  Pred. No. 1e-05;
  Matches   71;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || |||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCC 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAG 107
              | |   ||   |  |    ||  |||| ||
Db         77 TGCAGTGAGAAAGTGGCAAAGGGCTTTTCAG 107


RESULT 50
US-10-404-460-24/c
; Sequence 24, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1, 5, 21, 31, 34, 51, 60, 63, 81, 112, 118, 208, 209, 210,
;   LOCATION: 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222,
```

```
;   LOCATION: 223, 224, 225, 226
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-24

  Query Match             4.5%;  Score 59;  DB 7;  Length 691;
  Best Local Similarity   92.5%;  Pred. No. 1e-05;
  Matches   62;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 616

Qy         77 GGGAAGC 83
              | |||| |
Db        615 GAGAACC 609


RESULT 51
US-10-404-460-169
; Sequence 169, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169
;   LENGTH: 695
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 619, 661
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-169

  Query Match             4.5%;  Score 59;  DB 7;  Length 695;
  Best Local Similarity   92.5%;  Pred. No. 1e-05;
  Matches   62;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCT 76

Qy         77 GGGAAGC 83
              | |||||
Db         77 GAGAAGC 83


RESULT 52
US-10-404-460-201/c
; Sequence 201, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 201
;   LENGTH: 627
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 22, 117, 150, 208, 231, 269, 275, 307, 432, 624, 625, 626
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-201

  Query Match             4.5%;  Score 58.8;  DB 7;  Length 627;
  Best Local Similarity   90.0%;  Pred. No. 1.1e-05;
  Matches   63;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || |||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        605 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCC 546

Qy         77 GGGAAGCAGC 86
                 |||   |||
Db        545 TGGAGCGAGC 536


RESULT 53
US-10-404-460-214
; Sequence 214, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 214
;   LENGTH: 688
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
US-10-404-460-214

  Query Match             4.5%;  Score 58.8;  DB 7;  Length 688;
  Best Local Similarity   90.0%;  Pred. No. 1.1e-05;
  Matches   63;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || |||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         15 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCC 74

Qy         77 GGGAAGCAGC 86
                 |||   |||
Db         75 TGGAGCGAGC 84


RESULT 54
US-10-756-149-4637/c
; Sequence 4637, Application US/10756149
; Publication No. US20050181375A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
;  TITLE OF INVENTION:  METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
;  FILE REFERENCE: file
;  CURRENT APPLICATION NUMBER: US/10/756,149
;  CURRENT FILING DATE:  2004-01-12
;  NUMBER OF SEQ ID NOS: 5818
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4637
;   LENGTH: 382
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-756-149-4637

  Query Match             4.5%;  Score 58.6;  DB 9;  Length 382;
  Best Local Similarity   87.7%;  Pred. No. 9.3e-06;
  Matches   64;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 289

Qy         77 GGGAAGCAGCGTG 89
              |  |   |||| |
Db        288 GACAGTGAGCGGG 276


RESULT 55
US-10-404-460-129
; Sequence 129, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 129
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 2, 588, 626
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-129

  Query Match             4.5%;  Score 58.6;  DB 7;  Length 631;
  Best Local Similarity   93.8%;  Pred. No. 1.2e-05;
  Matches   61;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 81

Qy         77 GGGAA 81
              ||| |
Db         82 CGGCA 86


RESULT 56
US-10-404-460-155
; Sequence 155, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155
;   LENGTH: 705
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 617, 620
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-155

  Query Match             4.5%;  Score 58.6;  DB 7;  Length 705;
  Best Local Similarity   75.3%;  Pred. No. 1.3e-05;
  Matches   73;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACC 113
              |||||   |    |||||       |  || || |
Db         76 TTTAAGCATATTAGTCAGCGGAGAAGCTTCGGCGAGC 112


RESULT 57
US-10-198-846-9716
; Sequence 9716, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9716
;   LENGTH: 846
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1, 2, 845, 846
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9716

  Query Match             4.5%;  Score 58.6;  DB 5;  Length 846;
  Best Local Similarity   87.7%;  Pred. No. 1.5e-05;
```

```
  Matches   64;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         11 GATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCAC 70
              ||| |||||||||||||||||||||||| | ||||||||||||||||||||||||||||
Db         11 GATACGAATTTATACGACTCACTATAGGGGAATTGGCCCTCGAGGCCAAGAATTCGGCAC 70

Qy         71 GAGGGCGGGAAGC 83
              |||   |  |  ||
Db         71 GAGATCATGGTGC 83


RESULT 58
US-10-198-846-12993
; Sequence 12993, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12993
;   LENGTH: 422
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1, 422
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12993

  Query Match             4.5%;  Score 58.4;  DB 5;  Length 422;
  Best Local Similarity   81.0%;  Pred. No. 1.1e-05;
  Matches   68;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || |||||||||||||| |||||||||||||||||||||||||||||||||||||
Db         19 ATTTAATACGACTCACTATATGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 78

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              |    | |    || | || ||||
Db         79 GCAGTGAACTCAGAGGTGCTGGTG 102


RESULT 59
US-10-756-149-4511/c
; Sequence 4511, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
;  TITLE OF INVENTION:  METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
;  FILE REFERENCE: file
;  CURRENT APPLICATION NUMBER: US/10/756,149
;  CURRENT FILING DATE:  2004-01-12
;  NUMBER OF SEQ ID NOS: 5818
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4511
;   LENGTH: 519
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-756-149-4511

  Query Match             4.4%;  Score 58.2;  DB 9;  Length 519;
  Best Local Similarity   95.2%;  Pred. No. 1.4e-05;
  Matches   60;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 382

Qy         77 GGG 79
              |||
Db        381 GGG 379


RESULT 60
US-10-404-460-264
; Sequence 264, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 264
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
US-10-404-460-264

  Query Match             4.4%;  Score 58;  DB 7;  Length 633;
  Best Local Similarity   92.4%;  Pred. No. 1.8e-05;
  Matches   61;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 77

Qy         77 GGGAAG 82
              ||||
Db         78 TTGAAG 83


RESULT 61
US-10-404-460-6
; Sequence 6, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 666
;   TYPE: DNA
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 659
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-6

  Query Match             4.4%;  Score 58;  DB 7;  Length 666;
  Best Local Similarity   86.5%;  Pred. No. 1.9e-05;
  Matches   64;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAG 77

Qy         77 GGGAAGCAGCGTGA 90
              || | || || |
Db         78 CTGAGGAAGAGTAA 91


RESULT 62
US-10-404-460-27
; Sequence 27, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
;   LENGTH: 670
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 369, 379, 581, 590, 639, 668
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-27

  Query Match             4.4%;  Score 57.8;  DB 7;  Length 670;
  Best Local Similarity   89.9%;  Pred. No. 2.2e-05;
  Matches   62;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 78

Qy         77 GGGAAGCAG 85
              |  ||  ||
Db         79 GACAAAGAG 87


RESULT 63
US-10-404-460-207
; Sequence 207, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 207
;   LENGTH: 670
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 5, 670
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-207

  Query Match             4.4%;  Score 57.8;  DB 7;  Length 670;
  Best Local Similarity   84.4%;  Pred. No. 2.2e-05;
  Matches   65;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         15 CGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 CAATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 77

Qy         75 GCGGGAAGCAGCGTGAA 91
              |    |  |||  |  ||
Db         78 GGAAAATCCAGGGACAA 94


RESULT 64
US-10-404-460-93
; Sequence 93, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 93
;   LENGTH: 671
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 657
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-93

  Query Match             4.4%;  Score 57.8;  DB 7;  Length 671;
  Best Local Similarity   84.4%;  Pred. No. 2.2e-05;
  Matches   65;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGG 76
```

```
Qy          77 GGGAAGCAGCGTGAAGA 93
               | ||   | |  ||| |
Db          77 AGCACTAACCAGGAAAA 93


RESULT 65
US-10-404-460-196
; Sequence 196, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 196
;   LENGTH: 677
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 612, 669
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-196

  Query Match             4.4%;  Score 57.8;  DB 7;  Length 677;
  Best Local Similarity   89.9%;  Pred. No. 2.2e-05;
  Matches   62;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
               | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          20 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCT 79

Qy          77 GGGAAGCAG 85
               | ||| | |
Db          80 GAGAATCTG 88


RESULT 66
US-10-404-460-187
; Sequence 187, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 187
;   LENGTH: 689
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3, 5
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-187

  Query Match             4.4%;  Score 57.8;  DB 7;  Length 689;
  Best Local Similarity   89.9%;  Pred. No. 2.2e-05;
  Matches   62;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
               | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGG 77

Qy          77 GGGAAGCAG 85
               | ||| | |
Db          78 AGCATGCTG 86


RESULT 67
US-10-152-319A-915/c
; Sequence 915, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Higgs, Brandon
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Elashoff, Michael
;  TITLE OF INVENTION: Molecular Toxicology Modeling
;  FILE REFERENCE: 44921-5089-US
;  CURRENT APPLICATION NUMBER: US/10/152,319A
;  CURRENT FILING DATE:  2002-05-22
;  PRIOR APPLICATION NUMBER: US 60/292,335
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/297,523
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,925
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,810
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/303,807
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/303,808
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/315,047
;  PRIOR FILING DATE: 2001-08-28
;  PRIOR APPLICATION NUMBER: US 60/324,928
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US 60/330,867
;  PRIOR FILING DATE: 2001-11-01
;  PRIOR APPLICATION NUMBER: US 60/330,462
;  PRIOR FILING DATE: 2001-10-22
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2221
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 915
;   LENGTH: 315
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. AI104035
US-10-152-319A-915

  Query Match             4.4%;  Score 57.6;  DB 7;  Length 315;
  Best Local Similarity   82.5%;  Pred. No. 1.6e-05;
  Matches   66;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
               | || ||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         291 ATTTAATACGACTCCCTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 232

Qy          77 GGGAAGCAGCGTGAAGAGCG 96
```

```
                |||| |  |  ||| |
Db        231 GGGAGTGGTCCCGGCGAGTG 212
```

RESULT 68
US-10-404-460-221
; Sequence 221, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
; APPLICANT: PHASE-1 Molecular Toxicology, Inc.
; APPLICANT: Kier, Larry
; APPLICANT: Nolan, Timothy D.
; APPLICANT: Sankar, Usha
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Liver Necrosis Predictive Genes
; FILE REFERENCE: 2874-022
; CURRENT APPLICATION NUMBER: US/10/404,460
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/369,287
; PRIOR FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 221
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: predictive genes for liver necrosis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 606, 651, 681
; OTHER INFORMATION: n = A,T,C or G
US-10-404-460-221

Query Match 4.4%; Score 57.6; DB 7; Length 708;
Best Local Similarity 82.5%; Pred. No. 2.6e-05;
Matches 66; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAT 77

Qy         77 GGGAAGCAGCGTGAAGAGCG 96
                | |||    | |  ||| |
Db         78 AGAAAGTGAAGGGCTGAGGG 97
```

RESULT 69
US-10-404-460-137
; Sequence 137, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
; APPLICANT: PHASE-1 Molecular Toxicology, Inc.
; APPLICANT: Kier, Larry
; APPLICANT: Nolan, Timothy D.
; APPLICANT: Sankar, Usha
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Liver Necrosis Predictive Genes
; FILE REFERENCE: 2874-022
; CURRENT APPLICATION NUMBER: US/10/404,460
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/369,287
; PRIOR FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 137
; LENGTH: 728
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: predictive genes for liver necrosis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 23, 646, 648, 655, 696, 707
; OTHER INFORMATION: n = A,T,C or G
US-10-404-460-137

Query Match 4.4%; Score 57.6; DB 7; Length 728;
Best Local Similarity 86.3%; Pred. No. 2.6e-05;
Matches 63; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          5 ATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATT 64
              | | ||  |||  |   |||||||||||||||||||||||||||||||||||||||||
Db          7 AGTGATTTTTCCCAAGNATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATT 66

Qy         65 CGGCACGAGGGCG 77
              ||||||||||| |
Db         67 CGGCACGAGGGAG 79
```

RESULT 70
US-10-404-460-166
; Sequence 166, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
; APPLICANT: PHASE-1 Molecular Toxicology, Inc.
; APPLICANT: Kier, Larry
; APPLICANT: Nolan, Timothy D.
; APPLICANT: Sankar, Usha
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Liver Necrosis Predictive Genes
; FILE REFERENCE: 2874-022
; CURRENT APPLICATION NUMBER: US/10/404,460
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/369,287
; PRIOR FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 266
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 166
; LENGTH: 663
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: predictive genes for liver necrosis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 635, 641, 647, 651, 660, 661, 663
; OTHER INFORMATION: n = A,T,C or G
US-10-404-460-166

Query Match 4.4%; Score 57.4; DB 7; Length 663;
Best Local Similarity 76.9%; Pred. No. 2.8e-05;
Matches 70; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 78

Qy         77 GGGAAGCAGCGTGAAGAGCGGGGTGTTTTGAG 107
                 |    |  | ||| | | ||  |  ||
Db         79 ATTTAAGCTCCAGCAGACCAGCTGCCTGAAG 109
```

RESULT 71
US-10-404-460-163
; Sequence 163, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
; APPLICANT: PHASE-1 Molecular Toxicology, Inc.
; APPLICANT: Kier, Larry
; APPLICANT: Nolan, Timothy D.
; APPLICANT: Sankar, Usha
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Liver Necrosis Predictive Genes
; FILE REFERENCE: 2874-022
; CURRENT APPLICATION NUMBER: US/10/404,460

```
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 163
;   LENGTH: 705
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 636, 656, 692
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-163

  Query Match             4.4%;  Score 57.4;  DB 7;  Length 705;
  Best Local Similarity   85.3%;  Pred. No. 2.9e-05;
  Matches   64;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 78

Qy         77 GGGAAGCAGCGTGAA 91
              ||   ||   | ||
Db         79 GGTGGCCAAGTTCAA 93


RESULT 72
US-10-404-460-150
; Sequence 150, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 150
;   LENGTH: 422
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 12, 14, 109, 159, 172, 173, 174, 247, 305, 307, 319, 325,
;   LOCATION: 341, 349, 358, 369, 375, 380, 382, 388, 394, 398, 402, 408,
;   LOCATION: 410, 414, 416, 420
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-150

  Query Match             4.4%;  Score 57.2;  DB 7;  Length 422;
  Best Local Similarity   72.6%;  Pred. No. 2.5e-05;
  Matches   85;  Conservative    0;  Mismatches   31;  Indels    1;  Gaps    1;

Qy          3 CTATTAATGATTCGAATTT-ATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGA 61
              ||    |||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db          9 CTCNANATGATTCGAATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGA 68

Qy         62 ATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCG 118
              ||||||||||||||      | |||  |    | | |    ||       ||||||
Db         69 ATTCGGCACGAGGGACATCTGTAGCTGGGGAGTCAGTTAANGTGGTCTCTTCCTGCG 125


RESULT 73
US-10-404-460-123
; Sequence 123, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 123
;   LENGTH: 690
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 615, 639
;   OTHER INFORMATION: n = A,T,C or G
US-10-404-460-123

  Query Match             4.4%;  Score 57.2;  DB 7;  Length 690;
  Best Local Similarity   95.2%;  Pred. No. 3.3e-05;
  Matches   59;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCC 76

Qy         77 GG 78
              ||
Db         77 GG 78


RESULT 74
US-10-404-460-230
; Sequence 230, Application US/10404460
; Publication No. US20040076974A1
; GENERAL INFORMATION:
;  APPLICANT: PHASE-1 Molecular Toxicology, Inc.
;  APPLICANT:  Kier, Larry
;  APPLICANT:  Nolan, Timothy D.
;  APPLICANT:  Sankar, Usha
;  APPLICANT:  Derbel, Maher
;  TITLE OF INVENTION: Liver Necrosis Predictive Genes
;  FILE REFERENCE: 2874-022
;  CURRENT APPLICATION NUMBER: US/10/404,460
;  CURRENT FILING DATE:  2003-04-01
;  PRIOR APPLICATION NUMBER: US 60/369,287
;  PRIOR FILING DATE: 2002-04-01
;  NUMBER OF SEQ ID NOS: 266
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 230
;   LENGTH: 719
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: predictive genes for liver necrosis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 588, 597, 643, 660, 666, 700, 703, 713
;   OTHER INFORMATION: n = A,T,C or G
```

US-10-404-460-230

Query Match 4.4%; Score 57.2; DB 7; Length 719;
Best Local Similarity 88.6%; Pred. No. 3.3e-05;
Matches 62; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy          17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
               | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 77

Qy          77 GGGAAGCAGC 86
                 |   |  ||
Db          78 AGTTTGATGC 87
```

RESULT 75
US-10-357-930-24319
; Sequence 24319, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel. Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24319
; LENGTH: 3232
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 3230, 3231, 3232
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24319

Query Match 4.4%; Score 57.2; DB 8; Length 3232;
Best Local Similarity 83.3%; Pred. No. 7.9e-05;
Matches 65; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy          17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
               | || |||||||||||||||||||||||||||||||||||||||||||||| || | |
Db          33 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACCAGAGT 92

Qy          77 GGGAAGCAGCGTGAAGAG 94
                 ||  ||| | | |  | |
Db          93 GGAGAGCGGGGAGTTGTG 110
```

Search completed: March 11, 2006, 16:54:37
Job time : 1210 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2006, 12:48:52 ; Search time 268 Seconds
(without alignments)
8702.093 Million cell updates/sec

Title: US-10-757-745-3
Perfect score: 1312
Sequence: 1 agctattaatgattcgaatt.....tctcaacctttcaggacatc 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1312	100.0	1312	3	US-09-697-863A-3	Sequence 3, Appli
	2	668.2	50.9	1920	3	US-09-697-863A-1	Sequence 1, Appli
	3	668.2	50.9	1948	3	US-09-620-312D-889	Sequence 889, App
c	4	490	37.3	1079	3	US-09-118-554-63	Sequence 63, Appl
c	5	490	37.3	1079	3	US-09-118-627-63	Sequence 63, Appl
c	6	490	37.3	1079	3	US-09-602-877A-63	Sequence 63, Appl
c	7	101.4	7.7	268	3	US-09-397-787-312	Sequence 312, App
	8	62.2	4.7	1549	3	US-09-444-336-5	Sequence 5, Appli
	9	62.2	4.7	1549	3	US-10-093-045-5	Sequence 5, Appli
	10	62.2	4.7	1549	3	US-10-093-246-5	Sequence 5, Appli
	11	56.2	4.3	768	3	US-09-530-219B-4	Sequence 4, Appli
	12	55.8	4.3	565	3	US-09-712-016-56	Sequence 56, Appl
	13	55.8	4.3	2885	3	US-09-232-200-36	Sequence 36, Appl
	14	55.8	4.3	2885	3	US-09-232-200-56	Sequence 56, Appl
	15	55.8	4.3	2885	3	US-09-232-197-36	Sequence 36, Appl
	16	55.8	4.3	2885	3	US-09-232-197-56	Sequence 56, Appl
	17	55.8	4.3	2885	3	US-09-232-201-36	Sequence 36, Appl
	18	55.8	4.3	2885	3	US-09-232-201-56	Sequence 56, Appl
	19	55.8	4.3	2885	3	US-09-232-195-36	Sequence 36, Appl
	20	55.8	4.3	2885	3	US-09-232-195-56	Sequence 56, Appl
	21	54.8	4.2	1055	3	US-09-215-131-3	Sequence 3, Appli
	22	54.8	4.2	1055	3	US-09-222-734-3	Sequence 3, Appli
c	23	53.4	4.1	55	2	US-08-715-941-11	Sequence 11, Appl
c	24	53.4	4.1	55	2	US-08-715-941-14	Sequence 14, Appl
	25	52.2	4.0	1979	3	US-09-620-312D-838	Sequence 838, App
c	26	49.4	3.8	7218	2	US-08-232-463-14	Sequence 14, Appl
	27	49.4	3.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	28	49.2	3.8	2846	3	US-09-613-182-5	Sequence 5, Appli
	29	48.8	3.7	1370	3	US-09-620-312D-680	Sequence 680, App
c	30	47.2	3.6	906	3	US-09-206-166-4	Sequence 4, Appli
	31	47.2	3.6	1141	3	US-09-806-708B-22	Sequence 22, Appl
c	32	47.2	3.6	1411	3	US-09-620-312D-535	Sequence 535, App
c	33	47.2	3.6	2672	3	US-09-206-166-9	Sequence 9, Appli
	34	46.8	3.6	296	2	US-08-715-941-4	Sequence 4, Appli
	35	46.4	3.5	234	3	US-08-906-156A-21	Sequence 21, Appl
c	36	46.2	3.5	426	3	US-09-252-991A-7105	Sequence 7105, Ap
	37	46.2	3.5	1479	3	US-09-252-991A-7048	Sequence 7048, Ap
c	38	46.2	3.5	4236	3	US-09-252-991A-7057	Sequence 7057, Ap
	39	46.2	3.5	10023	3	US-09-252-991A-6997	Sequence 6997, Ap
	40	46	3.5	1794	3	US-09-620-312D-988	Sequence 988, App
	41	46	3.5	2241	3	US-09-023-942A-9	Sequence 9, Appli
	42	45.4	3.5	600	3	US-09-342-353-3	Sequence 3, Appli
	43	45.4	3.5	641	3	US-09-620-312D-954	Sequence 954, App
	44	45.4	3.5	1752	3	US-09-902-540-3466	Sequence 3466, Ap
	45	45.4	3.5	4427	3	US-09-342-353-2	Sequence 2, Appli
	46	45.4	3.5	12277	3	US-09-902-540-1018	Sequence 1018, Ap
	47	45.2	3.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	48	44.6	3.4	1028	3	US-08-118-200-1	Sequence 1, Appli
	49	44.6	3.4	1028	3	US-08-458-745-1	Sequence 1, Appli
c	50	44.6	3.4	2339	3	US-09-495-050A-26	Sequence 26, Appl
	51	44.2	3.4	6611	2	US-08-402-282-2	Sequence 2, Appli
	52	44.2	3.4	6611	2	US-08-508-004-2	Sequence 2, Appli
	53	44.2	3.4	6611	2	US-08-402-066-2	Sequence 2, Appli
	54	44.2	3.4	6611	2	US-08-402-068-2	Sequence 2, Appli
c	55	42.8	3.3	700	3	US-09-735-271-843	Sequence 843, App
	56	42.8	3.3	1745	3	US-09-620-312D-711	Sequence 711, App
	57	42.6	3.2	4362	2	US-08-455-073A-1	Sequence 1, Appli
c	58	42.4	3.2	2641	3	US-09-799-451-224	Sequence 224, App
c	59	42.2	3.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	60	42.2	3.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	61	42	3.2	1845	3	US-09-902-540-7042	Sequence 7042, Ap
	62	42	3.2	3471	3	US-09-902-540-8385	Sequence 8385, Ap
	63	42	3.2	4087	3	US-09-902-540-625	Sequence 625, App
	64	42	3.2	7186	3	US-09-902-540-879	Sequence 879, App
	65	42	3.2	8865	3	US-09-949-016-4238	Sequence 4238, Ap
	66	42	3.2	239527	3	US-09-949-016-15980	Sequence 15980, A
	67	41.8	3.2	1493	3	US-09-620-312D-728	Sequence 728, App
	68	41.8	3.2	14382	3	US-09-902-540-1145	Sequence 1145, Ap
	69	41.6	3.2	985	3	US-09-056-556-182	Sequence 182, App
	70	41.6	3.2	985	3	US-09-072-596-177	Sequence 177, App
	71	41.6	3.2	985	3	US-09-072-967-182	Sequence 182, App
	72	41.6	3.2	985	3	US-10-193-002-177	Sequence 177, App
	73	41.6	3.2	985	3	US-10-084-843-182	Sequence 182, App
	74	41.6	3.2	10178	3	US-09-902-540-977	Sequence 977, App
c	75	41.4	3.2	2755	3	US-09-902-540-5035	Sequence 5035, Ap

ALIGNMENTS

RESULT 1
US-09-697-863A-3
; Sequence 3, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/09/697,863A
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3

```
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (122)..(1234)
;   OTHER INFORMATION:
US-09-697-863A-3

  Query Match             100.0%;  Score 1312;  DB 3;  Length 1312;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1312;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60

Qy         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120

Qy        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180

Qy        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240

Qy        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300

Qy        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360

Qy        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420

Qy        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480

Qy        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540

Qy        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600

Qy        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660

Qy        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720

Qy        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780

Qy        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840

Qy        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900

Qy        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960

Qy        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020

Qy       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080

Qy       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140

Qy       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200

Qy       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260

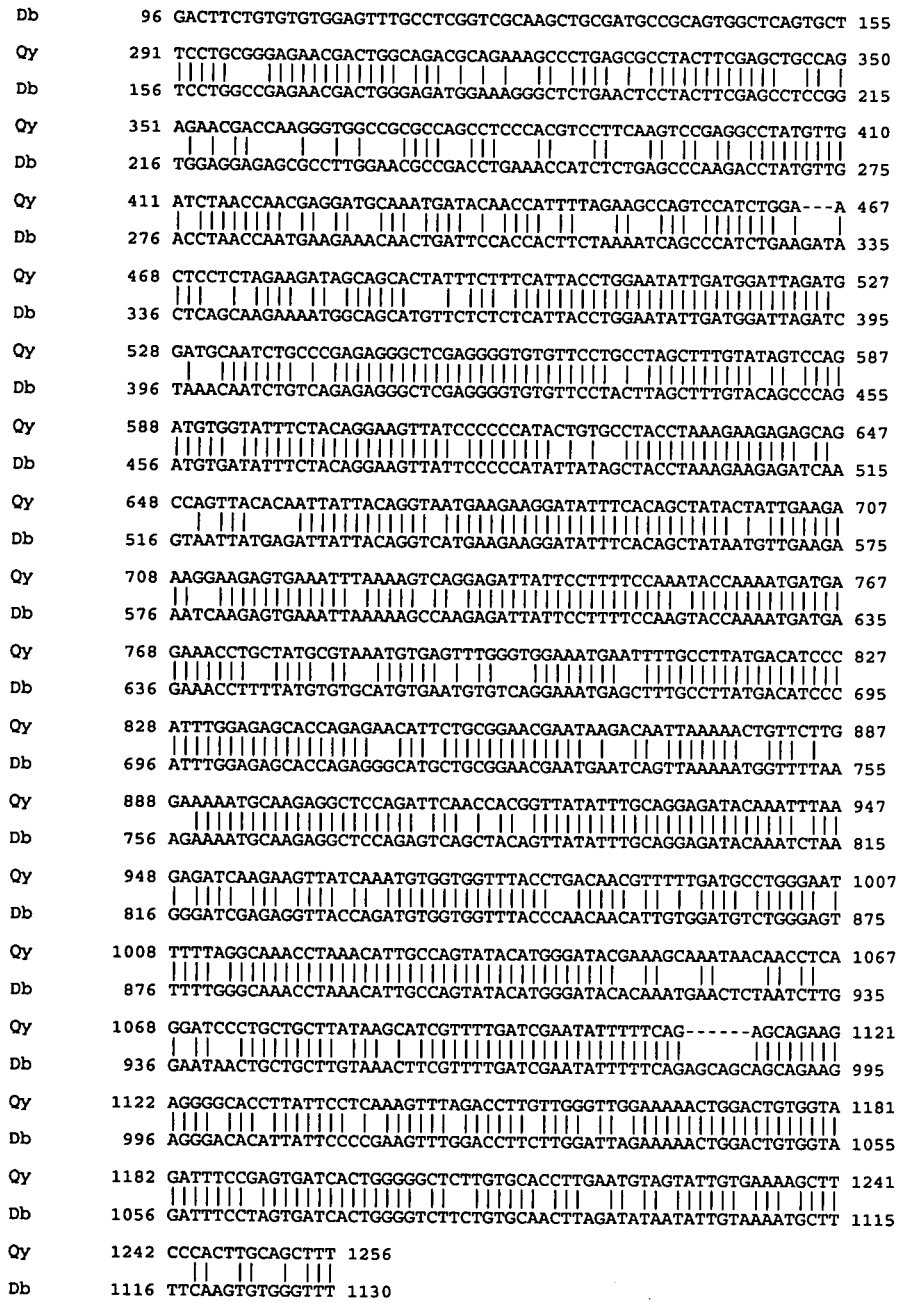
Qy       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312


RESULT 2
US-09-697-863A-1
; Sequence 1, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/09/697,863A
;  CURRENT FILING DATE:  2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1627)..(1627)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1108)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1849)..(1849)
;   OTHER INFORMATION: N stands for any nucleotide.
US-09-697-863A-1

  Query Match             50.9%;  Score 668.2;  DB 3;  Length 1920;
  Best Local Similarity   77.0%;  Pred. No. 4.7e-174;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||  | | |  | || | | | ||  |||      ||||| | |||||
Db         36 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 95

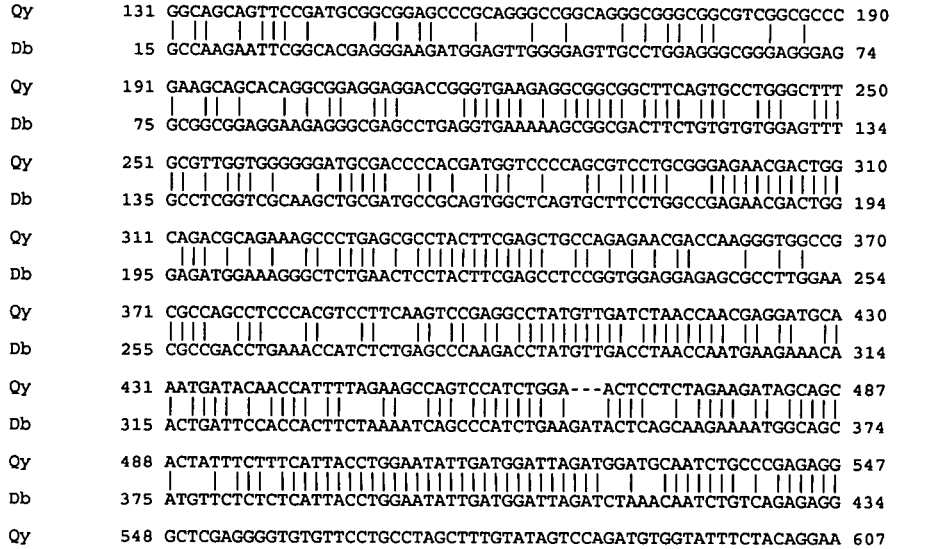
Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
```

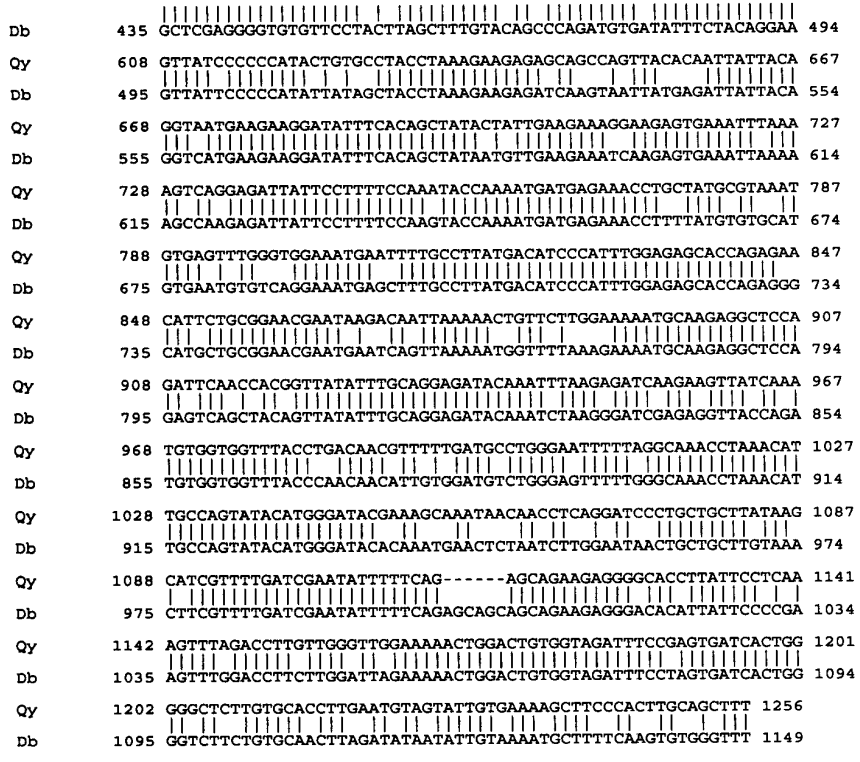


RESULT 3
US-09-620-312D-889

; Sequence 889, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Peiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-09-620-312D-889

Query Match 50.9%; Score 668.2; DB 3; Length 1948;
Best Local Similarity 75.6%; Pred. No. 4.7e-174;
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;





RESULT 4
US-09-118-554-63/c
; Sequence 63, Application US/09118554A
; Patent No. 6365348
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND
; TITLE OF INVENTION: METHODS FOR THEIR USE
; FILE REFERENCE: 210121.450C1
; CURRENT APPLICATION NUMBER: US/09/118,554A
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 08/988,255
; EARLIER FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-118-554-63

Query Match 37.3%; Score 490; DB 3; Length 1079;
Best Local Similarity 82.8%; Pred. No. 5.7e-125;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;



RESULT 5
US-09-118-627-63/c
; Sequence 63, Application US/09118627A
; Patent No. 6379951
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER
; TITLE OF INVENTION: AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C1
; CURRENT APPLICATION NUMBER: US/09/118,627A
; CURRENT FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-118-627-63

Query Match 37.3%; Score 490; DB 3; Length 1079;
Best Local Similarity 82.8%; Pred. No. 5.7e-125;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy       555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
Db      1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy       615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
Db      1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy       675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
Db       959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy       735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
Db       899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy       795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
Db       839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy       855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
Db       779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy       915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
Db       719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy       975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
Db       659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy      1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
Db       599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy      1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
Db       539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy      1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
Db       480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy      1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
Db       420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

RESULT 6
US-09-602-877A-63/c
; Sequence 63, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-63

Query Match 37.3%; Score 490; DB 3; Length 1079;
Best Local Similarity 82.8%; Pred. No. 5.7e-125;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy       555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
Db      1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy       615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
Db      1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy       675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
Db       959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy       735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
Db       899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy       795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
Db       839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy       855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
Db       779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy       915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
Db       719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy       975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
Db       659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy      1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
Db       599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy      1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
Db       539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy      1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
Db       480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy      1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
Db       420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

RESULT 7
US-09-397-787-312/c
; Sequence 312, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-312

```
  Query Match             7.7%;  Score 101.4;  DB 3;  Length 268;
  Best Local Similarity   81.8%;  Pred. No. 5.9e-18;
  Matches  117;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1114 AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGA 1173
             ||||||||||||| ||| ||||||| | |||||| |||||| |||| || ||||||||||
Db        262 AGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGA 203

Qy       1174 CTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTG 1233
             |||||||||||||||| |||||||||||||| ||  |||||| |||  || || ||||||
Db        202 CTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTA 143

Qy       1234 AAAAGCTTCCCACTTGCAGCTTT 1256
              ||| ||||  ||  ||  | |||
Db        142 AAATGCTTTTCAAGTGTGGGTTT 120


RESULT 8
US-09-444-336-5
; Sequence 5, Application US/09444336
; Patent No. 6410713
; GENERAL INFORMATION:
;  APPLICANT: Guerriero, Vincent
;  APPLICANT: Raynes, Deborah A
;  TITLE OF INVENTION: DNA Encoding Proteins That Inhibit Hsp70 Function
;  FILE REFERENCE: HspBP DNA and Protein Sequences
;  CURRENT APPLICATION NUMBER: US/09/444,336
;  CURRENT FILING DATE: 1999-11-19
;  EARLIER APPLICATION NUMBER: 60/109,351
;  EARLIER FILING DATE: 1998-11-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1549
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-444-336-5

  Query Match             4.7%;  Score 62.2;  DB 3;  Length 1549;
  Best Local Similarity   95.5%;  Pred. No. 9.5e-07;
  Matches   64;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGG 78
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          2 TTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCCGG 61

Qy         79 GAAGCAG 85
              |||||||
Db         62 CAAGCAG 68


RESULT 9
US-10-093-045-5
; Sequence 5, Application US/10093045
; Patent No. 6903202
; GENERAL INFORMATION:
;  APPLICANT: Guerriero, Vincent
;  APPLICANT: Raynes, Deborah A
;  TITLE OF INVENTION: DNA Encoding Proteins That Inhibit Hsp70 Function
;  FILE REFERENCE: HspBP DNA and Protein Sequences
;  CURRENT APPLICATION NUMBER: US/10/093,045
;  CURRENT FILING DATE: 2002-03-07
;  EARLIER APPLICATION NUMBER: 09/444,336
;  EARLIER FILING DATE: 1999-11-19
;  EARLIER APPLICATION NUMBER: 60/109,351
;  EARLIER FILING DATE: 1998-11-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1549
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-093-045-5

  Query Match             4.7%;  Score 62.2;  DB 3;  Length 1549;
  Best Local Similarity   95.5%;  Pred. No. 9.5e-07;
  Matches   64;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGG 78
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          2 TTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCCGG 61

Qy         79 GAAGCAG 85
              |||||||
Db         62 CAAGCAG 68


RESULT 10
US-10-093-246-5
; Sequence 5, Application US/10093246
; Patent No. 6908985
; GENERAL INFORMATION:
;  APPLICANT: Guerriero, Vincent
;  APPLICANT:  Raynes, Deborah A
;  TITLE OF INVENTION: DNA Encoding Proteins That Inhibit Hsp70 Function
;  FILE REFERENCE: HspBP DNA and Protein Sequences
;  CURRENT APPLICATION NUMBER: US/10/093,246
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: 09/444,336
;  PRIOR FILING DATE: 1999-11-19
;  PRIOR APPLICATION NUMBER: 60/109,351
;  PRIOR FILING DATE: 1998-11-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1549
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-093-246-5

  Query Match             4.7%;  Score 62.2;  DB 3;  Length 1549;
  Best Local Similarity   95.5%;  Pred. No. 9.5e-07;
  Matches   64;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGG 78
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          2 TTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCCGG 61

Qy         79 GAAGCAG 85
              |||||||
Db         62 CAAGCAG 68


RESULT 11
US-09-530-219B-4
; Sequence 4, Application US/09530219B
; Patent No. 6844422
; GENERAL INFORMATION:
;  APPLICANT: NIEHRS, CHRISTOF
;  APPLICANT:  GLINKA, ANDREI
;  TITLE OF INVENTION: AN INHIBITOR PROTEIN OF THE WNT SIGNAL PATH
;  FILE REFERENCE: RABG/40168
;  CURRENT APPLICATION NUMBER: US/09/530,219B
;  CURRENT FILING DATE:  2000-07-27
;  PRIOR APPLICATION NUMBER: PCT/DE98/03155
;  PRIOR FILING DATE: 1998-10-27
;  PRIOR APPLICATION NUMBER: DE 197 47 418.7
;  PRIOR FILING DATE: 1997-10-27
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 768
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-530-219B-4
```

```
  Query Match             4.3%;  Score 56.2;  DB 3;  Length 768;
  Best Local Similarity   95.1%;  Pred. No. 3e-05;
  Matches   58;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         22 ATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAA 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||  || |
Db          1 ATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGTTGGGA 60

Qy         82 G 82
              |
Db         61 G 61


RESULT 12
US-09-712-016-56
; Sequence 56, Application US/09712016
; Patent No. 6468790
; GENERAL INFORMATION:
;  APPLICANT: Giese, Klaus
;  TITLE OF INVENTION: METASTATIC BREAST AND COLON CANCER
;  TITLE OF INVENTION:  REGULATED GENES
;  FILE REFERENCE: 200130.460
;  CURRENT APPLICATION NUMBER: US/09/712,016
;  CURRENT FILING DATE:  2000-11-13
;  PRIOR APPLICATION NUMBER: US/09/417,615
;  PRIOR FILING DATE: PRIORFILING DATE: 1999-10-13
;  NUMBER OF SEQ ID NOS: 85
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 56
;   LENGTH: 565
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-712-016-56

  Query Match             4.3%;  Score 55.8;  DB 3;  Length 565;
  Best Local Similarity   96.6%;  Pred. No. 3.4e-05;
  Matches   57;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 97


RESULT 13
US-09-232-200-36
; Sequence 36, Application US/09232200A
; Patent No. 6288213
; GENERAL INFORMATION:
;  APPLICANT: Stahl, Andreas
;  APPLICANT: Hirsch, David J.
;  APPLICANT: Lodish, Harvey F.
;  APPLICANT: Gimeno, Ruth E.
;  APPLICANT: Tartaglia, Louis A.
;  TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
;  FILE REFERENCE: WHI97-21p3MB
;  CURRENT APPLICATION NUMBER: US/09/232,200A
;  CURRENT FILING DATE: 1999-01-14
;  EARLIER APPLICATION NUMBER: 60/071,374
;  EARLIER FILING DATE: 1998-01-15
;  EARLIER APPLICATION NUMBER: 60/093,491
;  EARLIER FILING DATE: 1998-07-20
;  EARLIER APPLICATION NUMBER: 60/110,941
;  EARLIER FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 36
;   LENGTH: 2885
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-232-200-36

  Query Match             4.3%;  Score 55.8;  DB 3;  Length 2885;
  Best Local Similarity   96.6%;  Pred. No. 7.7e-05;
  Matches   57;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208


RESULT 14
US-09-232-200-56
; Sequence 56, Application US/09232200A
; Patent No. 6288213
; GENERAL INFORMATION:
;  APPLICANT: Stahl, Andreas
;  APPLICANT: Hirsch, David J.
;  APPLICANT: Lodish, Harvey F.
;  APPLICANT: Gimeno, Ruth E.
;  APPLICANT: Tartaglia, Louis A.
;  TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
;  FILE REFERENCE: WHI97-21p3MB
;  CURRENT APPLICATION NUMBER: US/09/232,200A
;  CURRENT FILING DATE: 1999-01-14
;  EARLIER APPLICATION NUMBER: 60/071,374
;  EARLIER FILING DATE: 1998-01-15
;  EARLIER APPLICATION NUMBER: 60/093,491
;  EARLIER FILING DATE: 1998-07-20
;  EARLIER APPLICATION NUMBER: 60/110,941
;  EARLIER FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 56
;   LENGTH: 2885
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-232-200-56

  Query Match             4.3%;  Score 55.8;  DB 3;  Length 2885;
  Best Local Similarity   96.6%;  Pred. No. 7.7e-05;
  Matches   57;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208


RESULT 15
US-09-232-197-36
; Sequence 36, Application US/09232197A
; Patent No. 6300096
; GENERAL INFORMATION:
;  APPLICANT: Stahl, Andreas
;  APPLICANT: Hirsch, David J.
;  APPLICANT: Lodish, Harvey F.
;  APPLICANT: Gimeno, Ruth E.
;  APPLICANT: Tartaglia, Louis A.
;  TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
;  FILE REFERENCE: WHI97-21p3MA
;  CURRENT APPLICATION NUMBER: US/09/232,197A
;  CURRENT FILING DATE: 1999-01-14
;  EARLIER APPLICATION NUMBER: 60/071,374
;  EARLIER FILING DATE: 1998-01-15
;  EARLIER APPLICATION NUMBER: 60/093,491
;  EARLIER FILING DATE: 1998-07-20
;  EARLIER APPLICATION NUMBER: 60/110,941
;  EARLIER FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 36
;   LENGTH: 2885
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-232-197-36

  Query Match             4.3%;  Score 55.8;  DB 3;  Length 2885;
```

```
  Best Local Similarity   96.6%;  Pred. No. 7.7e-05;
  Matches   57;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208


RESULT 16
US-09-232-197-56
; Sequence 56, Application US/09232197A
; Patent No. 6300096
; GENERAL INFORMATION:
;  APPLICANT: Stahl, Andreas
;  APPLICANT: Hirsch, David J.
;  APPLICANT: Lodish, Harvey F.
;  APPLICANT: Gimeno, Ruth E.
;  APPLICANT: Tartaglia, Louis A.
;  TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
;  FILE REFERENCE: WHI97-21p3MA
;  CURRENT APPLICATION NUMBER: US/09/232,197A
;  CURRENT FILING DATE: 1999-01-14
;  EARLIER APPLICATION NUMBER: 60/071,374
;  EARLIER FILING DATE: 1998-01-15
;  EARLIER APPLICATION NUMBER: 60/093,491
;  EARLIER FILING DATE: 1998-07-20
;  EARLIER APPLICATION NUMBER: 60/110,941
;  EARLIER FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 56
;   LENGTH: 2885
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-232-197-56

  Query Match             4.3%;  Score 55.8;  DB 3;  Length 2885;
  Best Local Similarity   96.6%;  Pred. No. 7.7e-05;
  Matches   57;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208


RESULT 17
US-09-232-201-36
; Sequence 36, Application US/09232201A
; Patent No. 6348321
; GENERAL INFORMATION:
;  APPLICANT: Stahl, Andreas
;  APPLICANT: Hirsch, David J.
;  APPLICANT: Lodish, Harvey F.
;  APPLICANT: Gimeno, Ruth E.
;  APPLICANT: Tartaglia, Louis A.
;  TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
;  FILE REFERENCE: WHI97-21p3MC
;  CURRENT APPLICATION NUMBER: US/09/232,201A
;  CURRENT FILING DATE: 1999-01-14
;  EARLIER APPLICATION NUMBER: 60/071,374
;  EARLIER FILING DATE: 1998-01-15
;  EARLIER APPLICATION NUMBER: 60/093,491
;  EARLIER FILING DATE: 1998-07-20
;  EARLIER APPLICATION NUMBER: 60/110,941
;  EARLIER FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 36
;   LENGTH: 2885
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-232-201-36

  Query Match             4.3%;  Score 55.8;  DB 3;  Length 2885;
  Best Local Similarity   96.6%;  Pred. No. 7.7e-05;
  Matches   57;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208


RESULT 18
US-09-232-201-56
; Sequence 56, Application US/09232201A
; Patent No. 6348321
; GENERAL INFORMATION:
;  APPLICANT: Stahl, Andreas
;  APPLICANT: Hirsch, David J.
;  APPLICANT: Lodish, Harvey F.
;  APPLICANT: Gimeno, Ruth E.
;  APPLICANT: Tartaglia, Louis A.
;  TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
;  FILE REFERENCE: WHI97-21p3MC
;  CURRENT APPLICATION NUMBER: US/09/232,201A
;  CURRENT FILING DATE: 1999-01-14
;  EARLIER APPLICATION NUMBER: 60/071,374
;  EARLIER FILING DATE: 1998-01-15
;  EARLIER APPLICATION NUMBER: 60/093,491
;  EARLIER FILING DATE: 1998-07-20
;  EARLIER APPLICATION NUMBER: 60/110,941
;  EARLIER FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 56
;   LENGTH: 2885
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-232-201-56

  Query Match             4.3%;  Score 55.8;  DB 3;  Length 2885;
  Best Local Similarity   96.6%;  Pred. No. 7.7e-05;
  Matches   57;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208


RESULT 19
US-09-232-195-36
; Sequence 36, Application US/09232195A
; Patent No. 6657049
; GENERAL INFORMATION:
;  APPLICANT: Stahl, Andreas
;  APPLICANT: Hirsch, David J.
;  APPLICANT: Lodish, Harvey F.
;  APPLICANT: Gimeno, Ruth E.
;  APPLICANT: Tartaglia, Louis A.
;  TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
;  FILE REFERENCE: WHI97-21p3MD
;  CURRENT APPLICATION NUMBER: US/09/232,195A
;  CURRENT FILING DATE: 1999-01-04
;  EARLIER APPLICATION NUMBER: 60/071,374
;  EARLIER FILING DATE: 1998-01-15
;  EARLIER APPLICATION NUMBER: 60/093,491
;  EARLIER FILING DATE: 1998-07-20
;  EARLIER APPLICATION NUMBER: 60/110,941
;  EARLIER FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 36
;   LENGTH: 2885
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-232-195-36
```

```
  Query Match             4.3%;  Score 55.8;  DB 3;  Length 2885;
  Best Local Similarity   96.6%;  Pred. No. 7.7e-05;
  Matches   57;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208


RESULT 20
US-09-232-195-56
; Sequence 56, Application US/09232195A
; Patent No. 6657049
; GENERAL INFORMATION:
;  APPLICANT: Stahl, Andreas
;  APPLICANT: Hirsch, David J.
;  APPLICANT: Lodish, Harvey F.
;  APPLICANT: Gimeno, Ruth E.
;  APPLICANT: Tartaglia, Louis A.
;  TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
;  FILE REFERENCE: WHI97-21p3MD
;  CURRENT APPLICATION NUMBER: US/09/232,195A
;  CURRENT FILING DATE: 1999-01-04
;  EARLIER APPLICATION NUMBER: 60/071,374
;  EARLIER FILING DATE: 1998-01-15
;  EARLIER APPLICATION NUMBER: 60/093,491
;  EARLIER FILING DATE: 1998-07-20
;  EARLIER APPLICATION NUMBER: 60/110,941
;  EARLIER FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 56
;   LENGTH: 2885
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-232-195-56

  Query Match             4.3%;  Score 55.8;  DB 3;  Length 2885;
  Best Local Similarity   96.6%;  Pred. No. 7.7e-05;
  Matches   57;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208


RESULT 21
US-09-215-131-3
; Sequence 3, Application US/09215131
; Patent No. 6030834
; GENERAL INFORMATION:
;  APPLICANT: Chu, Keting
;  APPLICANT: Pot, David
;  TITLE OF INVENTION: IKK Beta Regulates Transcription Factors
;  FILE REFERENCE: 1449.002
;  CURRENT APPLICATION NUMBER: US/09/215,131
;  CURRENT FILING DATE: 1998-12-18
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 1055
;   TYPE: DNA
;   ORGANISM: human
US-09-215-131-3

  Query Match             4.2%;  Score 54.8;  DB 3;  Length 1055;
  Best Local Similarity   96.6%;  Pred. No. 8.7e-05;
  Matches   56;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 73


RESULT 22
US-09-222-734-3
; Sequence 3, Application US/09222734A
; Patent No. 6077701
; GENERAL INFORMATION:
;  APPLICANT: Chu, Keting
;  APPLICANT: Pot, David
;  TITLE OF INVENTION: IKK-beta Regulates Transcription Factors
;  FILE REFERENCE: 12441.78080
;  CURRENT APPLICATION NUMBER: US/09/222,734A
;  CURRENT FILING DATE: 1998-12-29
;  EARLIER APPLICATION NUMBER: 09/215,131
;  EARLIER FILING DATE: 1998-12-18
;  EARLIER APPLICATION NUMBER: 60/068,954
;  EARLIER FILING DATE: 1997-12-30
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1055
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-222-734-3

  Query Match             4.2%;  Score 54.8;  DB 3;  Length 1055;
  Best Local Similarity   96.6%;  Pred. No. 8.7e-05;
  Matches   56;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 73


RESULT 23
US-08-715-941-11/c
; Sequence 11, Application US/08715941
; Patent No. 5846721
;  GENERAL INFORMATION:
;    APPLICANT:  Soares, Marcelo B.
;    APPLICANT:  de Fatima Bonaldo, Maria
;    TITLE OF INVENTION:  AN EFFICIENT AND SIMPLER METHOD TO
;    TITLE OF INVENTION:  CONSTRUCT NORMALIZED cDNA LIBRARIES WITH IMPROVED
;    TITLE OF INVENTION:  REPRESENTATIONS OF FULL-LENGTH cDNAs.
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/715,941
;      FILING DATE:  19-SEP-1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/51083
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0526
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  55 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-715-941-11

  Query Match             4.1%;  Score 53.4;  DB 2;  Length 55;
  Best Local Similarity   98.2%;  Pred. No. 4.7e-05;
  Matches   54;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         20 TTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 TAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 1


RESULT 24
US-08-715-941-14/c
; Sequence 14, Application US/08715941
; Patent No. 5846721
;  GENERAL INFORMATION:
;    APPLICANT:  Soares, Marcelo B.
;    APPLICANT:  de Fatima Bonaldo, Maria
;    TITLE OF INVENTION:  AN EFFICIENT AND SIMPLER METHOD TO
;    TITLE OF INVENTION:  CONSTRUCT NORMALIZED cDNA LIBRARIES WITH IMPROVED
;    TITLE OF INVENTION:  REPRESENTATIONS OF FULL-LENGTH cDNAs.
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/715,941
;      FILING DATE:  19-SEP-1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/51083
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0526
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  55 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-715-941-14

  Query Match             4.1%;  Score 53.4;  DB 2;  Length 55;
  Best Local Similarity   98.2%;  Pred. No. 4.7e-05;
  Matches   54;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         20 TTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 TAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 1


RESULT 25
US-09-620-312D-838
; Sequence 838, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 838
;   LENGTH: 1979
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (175)..(1053)
US-09-620-312D-838

  Query Match             4.0%;  Score 52.2;  DB 3;  Length 1979;
  Best Local Similarity   87.7%;  Pred. No. 0.00062;
  Matches   57;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         24 ACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGC 83
              |||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db          5 ACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGTGAGGTGAG 64

Qy         84 AGCGT 88
              |||||
Db         65 AGAGT 69


RESULT 26
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             3.8%;  Score 49.4;  DB 2;  Length 7218;
  Best Local Similarity   4.1%;  Pred. No. 0.0071;
  Matches   14;  Conservative  193;  Mismatches  134;  Indels    0;  Gaps    0;

Qy          52 GAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGA 111
               :::: ::::: :: : ::::: :::::: :: : :::::: ::: :    :::::::
Db        1384 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1325

Qy         112 CCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGG 171
               : :: :: :: :   :: :: ::    :: : :: ::::   : ::::  :: :::
Db        1324 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1265

Qy         172 GCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCG 231
               : ::: :: :  :: :   :::: :: : ::: ::::::::  ::: :::::: :: :
Db        1264 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1205

Qy         232 GCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGT 291
               :    :: :   :::    : :  :: :::::::: : ::    : :: ::     :: :
Db        1204 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1145

Qy         292 CCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGA 351
               : ::::::: ::  :: ::: : ::::::    ::: :   :    :::  :  :::
Db        1144 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1085

Qy         352 GAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCA 392
               ::: ::  ::::: ::      ||  ||||| || ||| ||
Db        1084 RRRRRRRRRRRRRRRRRRATCGCAAGCTCCCTCGACCTGCA 1044


RESULT 27
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             3.8%;  Score 49.4;  DB 3;  Length 4411529;
  Best Local Similarity   53.3%;  Pred. No. 0.18;
  Matches  104;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy          71 GAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCT 130
               | || |  |  |  ||  |  | ||||||| |    | |   |  ||| ||  ||| ||
Db     1212240 GGGGCCACGGGGGTGCCGGCGGGGCGGGTGGTGCCGGCGTGGCTGGCGCCGGCGGCTTCG 1212299

Qy         131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
               |  |  | || |   |||||| ||| || | | |||| | || |||||| |||||||
Db     1212300 AGGGGACGATCGGTGCCGGCGGTGCCGGCGGTGTCGGCGGTGCCGGCGGTGTCGGCGGTG 1212359

Qy         191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
               || |  | ||||| |        || || |  || | |  |    |||  |
Db     1212360 CCGGCGGTGCCGGCGGGTGGCTGTACGGCGACGCCGGTGCCGGTGGGGATGGTGGTGTCG 1212419

Qy         251 GCGTTGGTGGGGGGA 265
               ||| ||  || || |
Db     1212420 GCGGTGCCGGCGGCA 1212434


RESULT 28
US-09-613-182-5
; Sequence 5, Application US/09613182
; Patent No. 6294653
; GENERAL INFORMATION:
;  APPLICANT: Mayfield, Stephen
;  TITLE OF INVENTION: RNA BINDING PROTEIN AND BINDING SITE USEFUL FOR
;  TITLE OF INVENTION:  EXPRESSION OF RECOMBINANT MOLECULES
;  FILE REFERENCE: SCR2177S
;  CURRENT APPLICATION NUMBER: US/09/613,182
;  CURRENT FILING DATE:  2000-07-10
;  PRIOR APPLICATION NUMBER: 09/341,550
;  PRIOR FILING DATE: 1999-07-13
;  PRIOR APPLICATION NUMBER: PCT/US98/00840
;  PRIOR FILING DATE: 1998-01-16
;  PRIOR APPLICATION NUMBER: 60/035,955
;  PRIOR FILING DATE: 1997-01-17
;  PRIOR APPLICATION NUMBER: 60/069,400
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 2846
;   TYPE: DNA
;   ORGANISM: Chlamydomonas reinhardtii
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (197)..(2065)
US-09-613-182-5

  Query Match             3.8%;  Score 49.2;  DB 3;  Length 2846;
  Best Local Similarity   53.0%;  Pred. No. 0.0051;
  Matches  105;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy         110 GACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCA 169
               | ||| || ||   |||| ||| |  |    | ||  ||| || |||||  |  | |
Db        1757 GCCCCCGCTGCCGCGGCGCCTGCCGCCGCGGCGGAGCCGGAGGCCCCCGCCGCCCAGCAG 1816

Qy         170 GGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGG 229
                || | |  || || ||||  |  || ||   | ||||| || |   | ||| ||   |
Db        1817 CCGCTGACCGCCTCCGCGCTGGCCGCCGCCGCGCCGGAGCAGCAGAAGATGATGATCGGC 1876

Qy         230 CGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGC 289
                ||  | || || |    | |||  | ||  |    |    |    | | || || ||
Db        1877 GAGCGCCTGTACCCGCAGGTGGCGGAGCTGCAGCCCGACCTGGCTGGCAAGATCACCGGC 1936

Qy         290 GTCCTGCGGGAGAACGAC 307
               | |||| |||||  |||
Db        1937 ATGCTGCTGGAGATGGAC 1954
```

```
RESULT 29
US-09-620-312D-680
; Sequence 680, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 680
;   LENGTH: 1370
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (122)..(1054)
US-09-620-312D-680

  Query Match             3.7%;  Score 48.8;  DB 3;  Length 1370;
  Best Local Similarity   77.6%;  Pred. No. 0.0045;
  Matches   59;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         26 GACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAG 85
              ||||||||||||||||||||||||||||||||| |||||||||||||||||   |  | |
Db         11 GACTCACTATAGGGAATTTGGCCCTCGAGGCAAAGAATTCGGCACGAGGTGAGCGGTCTG 70

Qy         86 CGTGAAGAGCGGGTGT 101
              ||    |  || || |
Db         71 CGGCGCGCACGTGTTT 86


RESULT 30
US-09-206-166-4/c
; Sequence 4, Application US/09206166A
; Patent No. 6492112
; GENERAL INFORMATION:
;  APPLICANT: Reinhard, Christoph
;  APPLICANT: Kassam, Altaf
;  APPLICANT: Pot, David
;  TITLE OF INVENTION: MITOGEN-ACTIVATED PROTEIN KINASE KINASE
;  TITLE OF INVENTION: 7 (MKK7)
;  FILE REFERENCE: 200130.436 / 1400.002
;  CURRENT APPLICATION NUMBER: US/09/206,166A
;  CURRENT FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 906
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-206-166-4

  Query Match             3.6%;  Score 47.2;  DB 3;  Length 906;
  Best Local Similarity   72.1%;  Pred. No. 0.01;
  Matches   75;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy         31 ACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGA 90
              |||||||||||||||||||||||||||||||||||||||||||||| || |     | ||  ||
Db        906 ACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGA-GGTGTTTGTCTGCCGGA 848

Qy         91 AGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCA 134
                |||| |    |  ||  | | |||||| ||||| | || |
Db        847 CTGACGGGCGGCCGGGCGGTGCGCGGCGGCGGTGGCGGCGGGGA 804


RESULT 31
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
;  FILE REFERENCE: 4810-58741
;  CURRENT APPLICATION NUMBER: US/09/806,708B
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/147,133
;  PRIOR FILING DATE: 1999-08-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 1141
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (1)..(1141)
;   OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

  Query Match             3.6%;  Score 47.2;  DB 3;  Length 1141;
  Best Local Similarity   11.1%;  Pred. No. 0.011;
  Matches   69;  Conservative  210;  Mismatches  342;  Indels    0;  Gaps    0;

Qy        408 TTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAA 467
              :::| : :              :: |::: ::: ::::|:::::  | :: ::   ::::: :
Db        338 KWSAAMGVYWNNNNNNNWTYKKARHBARWDWVWHSAWKKWHANAAHYSRKKWTBYKRKTM 397

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              :     |  : :::| ::::::::: |:        :::|      :  : :::    :
Db        398 VNNNNGTTMWKRMWAWYWKMDMDWBGTYNNNNNGGRTYYGWTKNKKMWTYYKWKANNCKW 457

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              :|:::    :     ::::         :: |  : : :   ::  : ::  ::|
Db        458 RAWDHKTCTHNNTTWWKMKTYWNNCYWKSMTNGKSHRBAAAVYTWYMWWWRRYAHANNNN 517

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              :: : :  |:: : : ::::  :|  :  :::    ::   :    :::: : :::| | |
Db        518 WDYWWKACTWYKYBVCSKWWNNYAAWYTKSSWNYTSRYYRWKTNNSWRWRSDTRSMGRAN 577

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              :|: ::: : : |:::     : :::|  |:  ::: ::   |: ::|:   ::  ||
Db        578 NYARABHYGYKWNTRWWBWSHTWBHBRAGAAHYWMBMMYBAKCHCMKAWYKAKKYAGAGG 637

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
                          |  |::    |::::|    :::::  ||   :  :|  ::|
Db        638 SNNNNNNNNNNNNNNNNATCARDDYYAASRWYAMANAKWYYYKBAANNAYYTHANNWWGC 697

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
                |  |: |:::       ||          :   ||: |    :  :::::  :  :
Db        698 WNNATDTRRTMWKNNNNNNAGTWKNNNNNNAKNASAAKNYAAAAVKAAKKHWRWANKWAM 757
```

```
Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              : :: :| |:    :  ||   :| |      |::|   | : :
Db        758 RGWHADAAABTTDKRNNGAYTKYTTTNNNNTYRGVVTNTAARDGWANNNNNNNNNNNNNN 817

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |: :::   :| |   |           :|    ::: ||:: :::|
Db        818 GWSDMWVTWWAYANYGTNNNNNNNNNNAYAWWTNKWYYTTDDRWRBAYTNNNNNNRMAYY 877

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              || |:  |  :::| : :: ::  :    ||     | :         ::  :::  |:
Db        878 GAYADDYAYYMSDTCDAWMKWDATKMNNATTYNRGTAWRTNNNNNNMTMKTKYYBHAAWN 937

Qy       1008 TTTTAGGCAAACCTAAACATT 1028
                |::  |: :::  ||:
Db        938 NNNNNGKMCTAHTWWVCKATK 958


RESULT 32
US-09-620-312D-535/c
; Sequence 535, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 535
;   LENGTH: 1411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (158)..(1333)
US-09-620-312D-535

  Query Match             3.6%;  Score 47.2;  DB 3;  Length 1411;
  Best Local Similarity   86.7%;  Pred. No. 0.013;
  Matches   52;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         30 CACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTG 89
              |||||||||||||||||||||||||||||||||||||||||| |     |||  ||||
Db       1411 CACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGCGGTTTTAGCTCCGTG 1352


RESULT 33
US-09-206-166-9/c
; Sequence 9, Application US/09206166A
; Patent No. 6492112
; GENERAL INFORMATION:
;  APPLICANT: Reinhard, Christoph
;  APPLICANT: Kassam, Altaf
;  APPLICANT: Pot, David
;  TITLE OF INVENTION: MITOGEN-ACTIVATED PROTEIN KINASE KINASE
;  TITLE OF INVENTION: 7 (MKK7)
;  FILE REFERENCE: 200130.436 / 1400.002
;  CURRENT APPLICATION NUMBER: US/09/206,166A
;  CURRENT FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 2672
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2672)
;   OTHER INFORMATION: n = A,T,C or G
US-09-206-166-9

  Query Match             3.6%;  Score 47.2;  DB 3;  Length 2672;
  Best Local Similarity   72.1%;  Pred. No. 0.017;
  Matches   75;  Conservative    0;  Mismatches   28;  Indels    1;  Gaps    1;

Qy         31 ACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGA 90
              |||||||||||||||||||||||||||||||||||||||||| || |     | ||  ||
Db       2672 ACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGA-GGTGTTTGTCTGCCGGA 2614

Qy         91 AGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCA 134
                |||| |    |  ||  | | |||||| ||||| | || |
Db       2613 CTGACGGGCGGCCGGGCGGTGCGCGGCGGCGGTGGCGGCGGGGA 2570


RESULT 34
US-08-715-941-4
; Sequence 4, Application US/08715941
; Patent No. 5846721
;  GENERAL INFORMATION:
;    APPLICANT:  Soares, Marcelo B.
;    APPLICANT:  de Fatima Bonaldo, Maria
;    TITLE OF INVENTION:  AN EFFICIENT AND SIMPLER METHOD TO
;    TITLE OF INVENTION:  CONSTRUCT NORMALIZED cDNA LIBRARIES WITH IMPROVED
;    TITLE OF INVENTION:  REPRESENTATIONS OF FULL-LENGTH cDNAs.
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/715,941
;      FILING DATE:  19-SEP-1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  0575/51083
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0526
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  296 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  DNA (genomic)
US-08-715-941-4

  Query Match             3.6%;  Score 46.8;  DB 2;  Length 296;
  Best Local Similarity   87.9%;  Pred. No. 0.0074;
  Matches   51;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || ||||||||||||||||||||||||||||||||||||||||||||||||   || | |
Db        100 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCCCGACTACG 157


RESULT 35
US-08-906-156A-21
; Sequence 21, Application US/08906156A
; Patent No. 6287854
;  GENERAL INFORMATION:
;    APPLICANT:  SPURR, NIGEL K
;    APPLICANT:  GRAY, IAN C
;    APPLICANT:  STEWART, LORNA M
;    TITLE OF INVENTION:  DIAGNOSIS OF SUSCEPTIBILITY TO CANCER
;    TITLE OF INVENTION:  AND TREATMENT THEREOF
;    NUMBER OF SEQUENCES:  94
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 NORTH GLEBE ROAD
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/906,156A
;      FILING DATE:  05-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/042,655
;      FILING DATE:  02-APR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/033,147
;      FILING DATE:  13-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/005,840
;      FILING DATE:  23-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/96GB/02588
;      FILING DATE:  22-OCT-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SADOFF, B.J.
;      REGISTRATION NUMBER:  36,663
;      REFERENCE/DOCKET NUMBER:  1090-14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  234 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-906-156A-21

  Query Match             3.5%;  Score 46.4;  DB 3;  Length 234;
  Best Local Similarity   89.3%;  Pred. No. 0.0085;
  Matches   50;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              || ||||||||||||||||||||||||||||||||||||||||||||||||   || | |
Db          2 TTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCCCGACTACG 57


RESULT 36
US-09-252-991A-7105/c
; Sequence 7105, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 7105
;   LENGTH: 426
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7105

  Query Match             3.5%;  Score 46.2;  DB 3;  Length 426;
  Best Local Similarity   49.0%;  Pred. No. 0.013;
  Matches  123;  Conservative    0;  Mismatches  128;  Indels    0;  Gaps    0;

Qy        154 GCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGA 213
              | |||| || | | | |    ||| ||  ||||| |    |        |  ||   |
Db        401 GGCCGCCGGCCTGCCGGCCGAGGCACCGATGGCGCACCTGCCGCGCGGTGTCGACCTGCC 342

Qy        214 CCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCC 273
              | | |  | || || || ||  || ||| ||    | | |    |   ||| | |
Db        341 GCAGTCGGCGGCGCAGCAGCGTCTCTGGCTGACCTGGCAGATCGATCCGCAGAGCGCCGC 282

Qy        274 CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGC 333
              |   |  ||||| ||| ||||||   |  || |  || |    |    || ||| ||||
Db        281 CTACAACATCCCCGGCGGCCTGCGCCTGCGCGGCGAGCTGGACGAAGCGGCGCTGCGCGC 222

Qy        334 CTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAA 393
              | |||| ||| ||  |   |   ||| |    || ||||  || | |      | |
Db        221 CAGCTTCCAGCGGCTGGTCGAACGCCACGAAGCGCTGCGCACGCGTTTCCTCGAGCGCGA 162

Qy        394 GTCCGAGGCCT 404
              || |||||
Db        161 CGGCGCGGCCT 151


RESULT 37
US-09-252-991A-7048
; Sequence 7048, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 7048
;   LENGTH: 1479
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7048
```

```
  Query Match             3.5%;  Score 46.2;  DB 3;  Length 1479;
  Best Local Similarity   49.0%;  Pred. No. 0.024;
  Matches  123;  Conservative    0;  Mismatches  128;  Indels    0;  Gaps    0;

Qy        154 GCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGA 213
              |||||| || | | | |     |||  ||  ||||| |     |        |  ||    |
Db        694 GGCCGCCGGCCTGCCGGCCGAGGCACCGATGGCGCACCTGCCGCGCGGTGTCGACCTGCC 753

Qy        214 CCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCC 273
              | | |  |  || || || ||  || |||  ||     | | |     |    ||| | |
Db        754 GCAGTCGGCGGCGCAGCAGCGTCTCTGGCTGACCTGGCAGATCGATCCGCAGAGCGCCGC 813

Qy        274 CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGC 333
              |    |   ||||| ||| |||||||    |  || |  || |     |     || ||| ||||
Db        814 CTACAACATCCCCGGCGGCCTGCGCCTGCGCGGCGAGCTGGACGAAGCGGCGCTGCGCGC 873

Qy        334 CTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAA 393
              |  |||| ||| ||  |    |   ||| |    || ||||  || |  |         | |
Db        874 CAGCTTCCAGCGGCTGGTCGAACGCCACGAAGCGCTGCGCACGCGTTTCCTCGAGCGCGA 933

Qy        394 GTCCGAGGCCT 404
                || |||||
Db        934 CGGCGCGGCCT 944
```

RESULT 38
US-09-252-991A-7057/c
; Sequence 7057, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7057
; LENGTH: 4236
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7057

```
  Query Match             3.5%;  Score 46.2;  DB 3;  Length 4236;
  Best Local Similarity   49.0%;  Pred. No. 0.042;
  Matches  123;  Conservative    0;  Mismatches  128;  Indels    0;  Gaps    0;

Qy        154 GCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGA 213
              |||||| || | | | |     |||  ||  ||||| |     |        |  ||    |
Db       2800 GGCCGCCGGCCTGCCGGCCGAGGCACCGATGGCGCACCTGCCGCGCGGTGTCGACCTGCC 2741

Qy        214 CCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCC 273
              | | |  |  || || || ||  || |||  ||     | | |     |    ||| | |
Db       2740 GCAGTCGGCGGCGCAGCAGCGTCTCTGGCTGACCTGGCAGATCGATCCGCAGAGCGCCGC 2681

Qy        274 CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGC 333
              |    |   ||||| ||| |||||||    |  || |  || |     |     || ||| ||||
Db       2680 CTACAACATCCCCGGCGGCCTGCGCCTGCGCGGCGAGCTGGACGAAGCGGCGCTGCGCGC 2621

Qy        334 CTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAA 393
              |  |||| ||| ||  |    |   ||| |    || ||||  || |  |         | |
Db       2620 CAGCTTCCAGCGGCTGGTCGAACGCCACGAAGCGCTGCGCACGCGTTTCCTCGAGCGCGA 2561

Qy        394 GTCCGAGGCCT 404
                || |||||
Db       2560 CGGCGCGGCCT 2550
```

RESULT 39

US-09-252-991A-6997
; Sequence 6997, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6997
; LENGTH: 10023
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6997

```
  Query Match             3.5%;  Score 46.2;  DB 3;  Length 10023;
  Best Local Similarity   49.0%;  Pred. No. 0.064;
  Matches  123;  Conservative    0;  Mismatches  128;  Indels    0;  Gaps    0;

Qy        154 GCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGA 213
              | |||| || | | | |     |||  ||  ||||| |     |        |  ||    |
Db       1569 GGCCGCCGGCCTGCCGGCCGAGGCACCGATGGCGCACCTGCCGCGCGGTGTCGACCTGCC 1628

Qy        214 CCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCC 273
              | |  |  | || || || ||  || |||  ||     | | |     |    ||| | |
Db       1629 GCAGTCGGCGGCGCAGCAGCGTCTCTGGCTGACCTGGCAGATCGATCCGCAGAGCGCCGC 1688

Qy        274 CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGC 333
              |    |   ||||| ||| |||||||    |  || |  || |     |     || ||| ||||
Db       1689 CTACAACATCCCCGGCGGCCTGCGCCTGCGCGGCGAGCTGGACGAAGCGGCGCTGCGCGC 1748

Qy        334 CTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAA 393
              |  |||| ||| ||  |    |   ||| |    || ||||  || |  |         | |
Db       1749 CAGCTTCCAGCGGCTGGTCGAACGCCACGAAGCGCTGCGCACGCGTTTCCTCGAGCGCGA 1808

Qy        394 GTCCGAGGCCT 404
                || |||||
Db       1809 CGGCGCGGCCT 1819
```

RESULT 40
US-09-620-312D-988
; Sequence 988, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19

```
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 988
;   LENGTH: 1794
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (164)..(1261)
US-09-620-312D-988

  Query Match             3.5%;  Score 46;  DB 3;  Length 1794;
  Best Local Similarity   100.0%;  Pred. No. 0.031;
  Matches   46;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         29 TCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              ||||||||||||||||||||||||||||||||||||||||||||
Db         14 TCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 59


RESULT 41
US-09-023-942A-9
; Sequence 9, Application US/09023942A
; Patent No. 6479274
;  GENERAL INFORMATION:
;    APPLICANT:  (US only) ANTALIS Toni Marie and HOOPER John David
;    TITLE OF INVENTION:  NOVEL MOLECULES
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 GARDEN CITY PLAZA
;      CITY:  GARDEN CITY
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/023,942A
;      FILING DATE:  13-FEB-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PO5101/97
;      FILING DATE:  13-FEB-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PP0422/97
;      FILING DATE:  18-NOV-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  International PCT Application
;      FILING DATE:  13-FEB-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DIGIGLIO,  FRANK S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  11168
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742 4343
;      TELEFAX:  (516) 742 4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2241 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  166..1773
US-09-023-942A-9

  Query Match             3.5%;  Score 46;  DB 3;  Length 2241;
  Best Local Similarity   73.0%;  Pred. No. 0.034;
  Matches   73;  Conservative    0;  Mismatches   25;  Indels    2;  Gaps    1;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || |||||||||||||||||||||||||||||||||  ||||||||||||||||
Db          1 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGG--AAGAATTCGGCACGAGGCT 58

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTG 116
              | |  |||  |||| |     |   |     ||| ||| |
Db         59 GCGGCGCACTGTGAGGGAGTCGCTGTGATCCGGGGCCCCG 98


RESULT 42
US-09-342-353-3
; Sequence 3, Application US/09342353
; Patent No. 6248569
; GENERAL INFORMATION:
;  APPLICANT: Dunn, John
;  TITLE OF INVENTION: METHOD FOR INTRODUCING UNIDIRECTIONAL NESTED DELETIONS
;  FILE REFERENCE: CIP OF U.S. Application 08/966,958
;  CURRENT APPLICATION NUMBER: US/09/342,353
;  CURRENT FILING DATE: 1999-06-29
;  EARLIER APPLICATION NUMBER: 08/966,958
;  EARLIER FILING DATE: 1997-11-10
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: polylinker of
;   OTHER INFORMATION: vector
US-09-342-353-3

  Query Match             3.5%;  Score 45.4;  DB 3;  Length 600;
  Best Local Similarity   89.1%;  Pred. No. 0.026;
  Matches   49;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         20 TTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCCCGACTACG 306


RESULT 43
US-09-620-312D-954
; Sequence 954, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
```

```
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 954
;   LENGTH: 641
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (103)..(360)
US-09-620-312D-954

  Query Match             3.5%;  Score 45.4;  DB 3;  Length 641;
  Best Local Similarity   89.1%;  Pred. No. 0.027;
  Matches   49;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         33 TATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCG 87
              ||||||||||||||||||||||||||||||||||||||||||||    |   |||||
Db          1 TATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGGACGGGCCAGCG 55


RESULT 44
US-09-902-540-3466
; Sequence 3466, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 3466
;   LENGTH: 1752
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-3466

  Query Match             3.5%;  Score 45.4;  DB 3;  Length 1752;
  Best Local Similarity   50.7%;  Pred. No. 0.044;
  Matches  109;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy        142 CGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACA 201
              ||  ||||  |  ||  |  ||  |    ||||  |     |||  |     |  |||  ||| |  |  |
Db        936 CGCGGCGGTGAAGGCGGCGGACGCGGCGGACGCGGCCGGCAAGACGCGCGAGGTGGTGCT 995

Qy        202 GGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGG 261
                ||  | ||  ||  ||  |  |||  ||||     |  |  | ||  |||||  ||  |  | | |
Db        996 CACGTGGAAGAACACGGAGCAGAAGCGGACCATCGACTTCCGGGGCTACGCCTATGAGCG 1055

Qy        262 GGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAA 321
              |||    |  |  |  |    | ||  |       | ||  |  |||||  | ||  | |||
Db       1056 GGGCGCCTCCGACGTGTCCGGCCAGACGTGGATTCGCTACGACGACACGAAGCCCCAGGT 1115

Qy        322 AGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACG 356
                  |     ||||||||||||  |     ||  ||
Db       1116 CTGGAAGGTGCCCTACTTCGAGGAGATTCAGCCCG 1150


RESULT 45
US-09-342-353-2
; Sequence 2, Application US/09342353
; Patent No. 6248569
; GENERAL INFORMATION:
;  APPLICANT: Dunn, John
;  TITLE OF INVENTION: METHOD FOR INTRODUCING UNIDIRECTIONAL NESTED DELETIONS
;  FILE REFERENCE: CIP OF U.S. Application 08/966,958
;  CURRENT APPLICATION NUMBER: US/09/342,353
;  CURRENT FILING DATE: 1999-06-29
;  EARLIER APPLICATION NUMBER: 08/966,958
;  EARLIER FILING DATE: 1997-11-10
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 4427
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: vector
;   OTHER INFORMATION: sequence
US-09-342-353-2

  Query Match             3.5%;  Score 45.4;  DB 3;  Length 4427;
  Best Local Similarity   89.1%;  Pred. No. 0.071;
  Matches   49;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         20 TTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | ||||||||||||||||||||||||||||||||||||||||||||||||||   || | |
Db        984 TAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCCCGACTACG 1038


RESULT 46
US-09-902-540-1018
; Sequence 1018, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1018
;   LENGTH: 12277
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-1018

  Query Match             3.5%;  Score 45.4;  DB 3;  Length 12277;
  Best Local Similarity   50.7%;  Pred. No. 0.12;
  Matches  109;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy        142 CGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACA 201
              ||  ||||  |  ||  |  ||  |    ||||  |     |||  |     |  |||  ||| |  |  |
Db       8611 CGCGGCGGTGAAGGCGGCGGACGCGGCGGACGCGGCCGGCAAGACGCGCGAGGTGGTGCT 8670

Qy        202 GGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGG 261
                ||  | ||  ||  ||  |  |||  ||||     |  |  | ||  |||||  ||  |  | | |
Db       8671 CACGTGGAAGAACACGGAGCAGAAGCGGACCATCGACTTCCGGGGCTACGCCTATGAGCG 8730

Qy        262 GGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAA 321
              |||    |  |  |  |    | ||  |       | ||  |  |||||  | ||  | |||
Db       8731 GGGCGCCTCCGACGTGTCCGGCCAGACGTGGATTCGCTACGACGACACGAAGCCCCAGGT 8790

Qy        322 AGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACG 356
                  |     ||||||||||||  |     ||  ||
Db       8791 CTGGAAGGTGCCCTACTTCGAGGAGATTCAGCCCG 8825


RESULT 47
```

```
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             3.4%;  Score 45.2;  DB 3;  Length 4403765;
  Best Local Similarity   54.1%;  Pred. No. 2.7;
  Matches   92;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy        114 CTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGC 173
              | ||| ||  || | | |    ||   | ||   |||| | | || | |||||| | ||
Db     150275 CGGCGACGGCGGGGCCGGCGGACACGGCGGACAAGGCGCCGGCGGCGGCGCCGGCGGCGC 150334

Qy        174 GGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGC 233
              ||||||| ||| | ||   || || | ||| | |||     ||||   |  | || |
Db     150335 CGGCGGCGACGGGGGCCAGGGCGGCGCCGGCCGCGGACTGTGGGGTACTGGCGGCGCCGG 150394

Qy        234 TTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTC 283
                  |    ||||   ||| ||||   |||   | ||| |    ||||
Db     150395 CGGACACGGCGGGCAAGGCGGTGGTACCGGGGGCCCACCGCTGCCCGGTC 150444


RESULT 48
US-08-118-200-1
; Sequence 1, Application US/08118200
; Patent No. 6197500
;  GENERAL INFORMATION:
;    APPLICANT:  SUTHERLAND, Grant R
;    APPLICANT:  RICHARDS, Robert I
;    APPLICANT:  SCHLESSINGER, David
;    APPLICANT:  NAGARAJA, Ramaiah
;    APPLICANT:  KREMER, Eric J
;    APPLICANT:  YU, Sui
;    APPLICANT:  BAKER, Elizabeth
;    APPLICANT:  MULLEY, John C
;    APPLICANT:  MANDEL, Jean-Louis
;    APPLICANT:  PRITCHARD, Melanie April
;    APPLICANT:  LYNCH, Michael
;    TITLE OF INVENTION:  DNA SEQUENCES RELATED TO ISOLATED
;    TITLE OF INVENTION:  FRAGILE X SYNDROME
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,200
;      FILING DATE:  09-SEP-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/802,650
;      FILING DATE:  05-DEC-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/672,232
;      FILING DATE:  20-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/638,518
;      FILING DATE:  04-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/966,517
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crane-Feury, Sharon E
;      REGISTRATION NUMBER:  36,113
;      REFERENCE/DOCKET NUMBER:  020160-164
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1028 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-118-200-1

  Query Match             3.4%;  Score 44.6;  DB 3;  Length 1028;
  Best Local Similarity   54.1%;  Pred. No. 0.056;
  Matches  112;  Conservative    0;  Mismatches   94;  Indels    1;  Gaps    1;

Qy        104 TGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGG 163
              ||  ||| | |  | | |  |||| | ||| |  |     |  |||| | | || | |
Db        315 TGGAGGGCCGCCTCTGAGCGGGCGGCGGGCCGACGGCGAGCGCGGGCGGCGGCGGCGCG 374

Qy        164 CCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGAC-CGGGTGAA 222
              |||| | |  ||||||| ||||| | |  || |    |||| || ||   ||| |
Db        375 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGTGGCGGCGGCGGCGGCGGCGGCG 434

Qy        223 GAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGT 282
              |||||||||| | | | | | |      | | | ||    | | | ||   ||
Db        435 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCCCGGAGCCACCTCTTCGGG 494

Qy        283 CCCCAGCGTCCTGCGGGAGAACGACTG 309
                |  ||  || |||  || | | |||
Db        495 GGCGGGCTCCCGGCGCTAGCAGGGCTG 521


RESULT 49
US-08-458-745-1
; Sequence 1, Application US/08458745
; Patent No. 6242576
;  GENERAL INFORMATION:
;    APPLICANT:  SUTHERLAND, Grant R
;    APPLICANT:  RICHARDS, Robert I
;    APPLICANT:  SCHLESSINGER, David
;    APPLICANT:  NAGARAJA, Ramaiah
;    APPLICANT:  KREMER, Eric J
;    APPLICANT:  YU, Sui
;    APPLICANT:  BAKER, Elizabeth
;    APPLICANT:  MULLEY, John C
;    APPLICANT:  MANDEL, Jean-Louis
;    APPLICANT:  PRITCHARD, Melanie April
;    APPLICANT:  LYNCH, Michael
;    TITLE OF INVENTION:  DNA SEQUENCES RELATED TO ISOLATED
;    TITLE OF INVENTION:  FRAGILE X SYNDROME
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS
;       STREET:  P.O. Box 1404
;       CITY:  Alexandria
;       STATE:  Virginia
;       COUNTRY:  United States
;       ZIP:  22313-1404
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/458,745
;       FILING DATE:  02-JUN-1995
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/118,200
;       FILING DATE:  09-SEP-1993
;       APPLICATION NUMBER:  US 07/802,650
;       FILING DATE:  05-DEC-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/672,232
;       FILING DATE:  20-MAR-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/638,518
;       FILING DATE:  04-JAN-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/966,517
;       FILING DATE:  23-DEC-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Crane-Feury, Sharon E
;       REGISTRATION NUMBER:  36,113
;       REFERENCE/DOCKET NUMBER:  020160-164
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 836-6620
;       TELEFAX:  (703) 836-2021
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1028 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-458-745-1

  Query Match             3.4%;  Score 44.6;  DB 3;  Length 1028;
  Best Local Similarity   54.1%;  Pred. No. 0.056;
  Matches  112;  Conservative    0;  Mismatches   94;  Indels    1;  Gaps    1;

Qy        104 TGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGG 163
              ||  ||| | |  | | |  |||| | ||| |  |     |     ||||| | | || | |
Db        315 TGGAGGGCCGCCTCTGAGCGGGCGGCGGGCCGACGGCGAGCGCGGGCGGCGGCGGCGGCG 374

Qy        164 CCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGAC-CGGGTGAA 222
              |||| | |  |||||||| ||||| | |  || |     ||||| || ||   ||| |
Db        375 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGTGGCGGCGGCGGCGGCGGCGGCGCG 434

Qy        223 GAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGT 282
              | |||||||||  | | | | | |     |  |  | ||    | | | || |   ||
Db        435 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCCCGGAGCCACCTCTTCGGG 494

Qy        283 CCCCAGCGTCCTGCGGGAGAACGACTG 309
                |  ||  || |||  || | | |||
Db        495 GGCGGGCTCCCGGCGCTAGCAGGGCTG 521


RESULT 50
US-09-495-050A-26/c
; Sequence 26, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
;  APPLICANT: Roopa, Reddy
;  APPLICANT:  Guegler, Karl, J.
;  APPLICANT:  Au-Young, Janice
;  TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED
;  FILE REFERENCE: PA-0013 US
;  CURRENT APPLICATION NUMBER: US/09/495,050A
;  CURRENT FILING DATE:  2000-01-31
;  PRIOR APPLICATION NUMBER: 60/118,318
;  PRIOR FILING DATE: February 1, 1999
;  NUMBER OF SEQ ID NOS: 305
;  SOFTWARE: PERL Program
; SEQ ID NO 26
;   LENGTH: 2339
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6492505 632097CB1
US-09-495-050A-26

  Query Match             3.4%;  Score 44.6;  DB 3;  Length 2339;
  Best Local Similarity   92.1%;  Pred. No. 0.085;
  Matches   58;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy         18 ATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCG 77
              |||||  |||||||||||||||||||||||||||||| |||||||||||||||||  |
Db       2339 ATTTAATCGACTCACTATAGGGAATTTGGCCCTCGAGG-CAAGAATTCGGCACGAGGCTG 2281

Qy         78 GGA 80
              |||
Db       2280 GGA 2278


RESULT 51
US-08-402-282-2
; Sequence 2, Application US/08402282
; Patent No. 5476768
;  GENERAL INFORMATION:
;    APPLICANT:  Pearson, Robert E.
;    APPLICANT:  Dickson, Julie A.
;    APPLICANT:  Hamilton, Paul T.
;    APPLICANT:  Little, Michael C.
;    APPLICANT:  Beyer Jr., Wayne F.
;    TITLE OF INVENTION:  MYCOBACTERIOPHAGE SPECIFIC FOR THE
;    TITLE OF INVENTION:  MYCOBACTERIUM TUBERCULOSIS COMPLEX
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard J. Rodrick, Becton Dickinson and
;      ADDRESSEE:  Company
;      STREET:  1 Becton Drive
;      CITY:  Franklin Lakes
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07417
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/402,282
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fugit, Donna R.
;      REGISTRATION NUMBER:  32,135
;      REFERENCE/DOCKET NUMBER:  P-3283
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6611 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
```

```
;       NAME/KEY:  misc_feature
;       LOCATION:  390..1538
;       OTHER INFORMATION:  /function= "coding sequence"
;       OTHER INFORMATION:  /product= "DNA polymerase"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  2107..3132
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  3138..3359
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  4690..5028
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  5028..5375
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  5375..5653
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  5653..5910
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  6078..6491
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
US-08-402-282-2

  Query Match             3.4%;  Score 44.2;  DB 2;  Length 6611;
  Best Local Similarity   49.4%;  Pred. No. 0.19;
  Matches  115;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy         69 ACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTGAGGGGACCCTGCGGCGATGGCGT 128
              |||||  ||||  | |   ||| | | |  |  || ||   | |||||| || |
Db        505 ACGAGTCCGGGCTGAAAGTTGAGGCGTTCGACTACGAAGTGGCGGCCGCGGCGACGGTGG 564

Qy        129 CTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGC 188
               |||| ||      |  | ||  |       |   |  | | || |||  |  | |||||
Db        565 ACGGCACCACGGTGGCCGAGGGCGGCGAAACGCTGGTGTCGGGCCGGCTGTTGGCGGCGA 624

Qy        189 CCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCT 248
               ||     ||   | || ||  ||  | |||||||     ||           | ||
Db        625 TCGCTAAGGCGTTGCCGAAGCGGGTGCCGGTGAAGTTTACGCACGACGGTGCGCGGGCTG 684

Qy        249 TTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
              | | |  || |||||     ||   |||| ||  | | |  | || |||||||
Db        685 TGGTGCAGGCGGGGGCCGCGGAGTTCACGCTGCCCACGATGGACCCGCGGGAG 737


RESULT 52
US-08-508-004-2
; Sequence 2, Application US/08508004
; Patent No. 5582969
;  GENERAL INFORMATION:
;    APPLICANT:  Pearson, Robert E.
;    APPLICANT:  Dickson, Julie A.
;    APPLICANT:  Hamilton, Paul T.
;    APPLICANT:  Little, Michael C.
;    APPLICANT:  Beyer Jr., Wayne F.
;    TITLE OF INVENTION:  MYCOBACTERIOPHAGE SPECIFIC FOR THE
;    TITLE OF INVENTION:  MYCOBACTERIUM TUBERCULOSIS COMPLEX
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard J. Rodrick, Becton Dickinson and
;      ADDRESSEE:  Company
;      STREET:  1 Becton Drive
;      CITY:  Franklin Lakes
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07417
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/508,004
;      FILING DATE:  27-JUL-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/402,282
;      FILING DATE:  10-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fugit, Donna R.
;      REGISTRATION NUMBER:  32,135
;      REFERENCE/DOCKET NUMBER:  P-3283
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6611 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  390..1538
;      OTHER INFORMATION:  /function= "coding sequence"
;      OTHER INFORMATION:  /product= "DNA polymerase"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  2107..3132
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  3138..3359
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  4690..5028
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  5028..5375
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  5375..5653
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  5653..5910
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  6078..6491
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
```

```
US-08-508-004-2

  Query Match             3.4%;  Score 44.2;  DB 2;  Length 6611;
  Best Local Similarity   49.4%;  Pred. No. 0.19;
  Matches  115;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy         69 ACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGT 128
              ||||| |||| | |   ||| | |   | |     || ||   | ||||||| || |
Db        505 ACGAGTCCGGGCTGAAAGTTGAGGCGTTCGACTACGAAGTGGCGGCCGCGGCGACGGTGG 564

Qy        129 CTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGC 188
               |||| ||      | | ||  |     |   | | | || |||  |  | |||||
Db        565 ACGGCACCACGGTGGCCGAGGGCGGCGAAACGCTGGTGTCGGGCCGGCTGTTGGCGGCGA 624

Qy        189 CCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCT 248
              ||     ||   | || ||  ||  | |||||||     ||           | ||
Db        625 TCGCTAAGGCGTTGCCGAAGCGGGTGCCGGTGAAGTTTACGCACGACGGTGCGCGGGCTG 684

Qy        249 TTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
              | | |  || |||||     ||   |||| ||  | | |  | || |||||||
Db        685 TGGTGCAGGCGGGGGCCGCGGAGTTCACGCTGCCCACGATGGACCCGCGGGAG 737


RESULT 53
US-08-402-066-2
; Sequence 2, Application US/08402066
; Patent No. 5612182
;  GENERAL INFORMATION:
;    APPLICANT:  Pearson, Robert E.
;    APPLICANT:  Dickson, Julie A.
;    APPLICANT:  Hamilton, Paul T.
;    APPLICANT:  Little, Michael C.
;    APPLICANT:  Beyer Jr., Wayne F.
;    TITLE OF INVENTION:  MYCOBACTERIOPHAGE SPECIFIC FOR THE
;    TITLE OF INVENTION:  MYCOBACTERIUM TUBERCULOSIS COMPLEX
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard J. Rodrick, Becton Dickinson and
;      ADDRESSEE:  Company
;      STREET:  1 Becton Drive
;      CITY:  Franklin Lakes
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07417
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/402,066
;      FILING DATE:
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fugit, Donna R.
;      REGISTRATION NUMBER:  32,135
;      REFERENCE/DOCKET NUMBER:  P-3283
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6611 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  390..1538
;      OTHER INFORMATION:  /function= "coding sequence"
;      OTHER INFORMATION:  /product= "DNA polymerase"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  2107..3132
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  3138..3359
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  4690..5028
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  5028..5375
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  5375..5653
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  5653..5910
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  6078..6491
;      OTHER INFORMATION:  /function= "potential open reading
;      OTHER INFORMATION:  frame"
US-08-402-066-2

  Query Match             3.4%;  Score 44.2;  DB 2;  Length 6611;
  Best Local Similarity   49.4%;  Pred. No. 0.19;
  Matches  115;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy         69 ACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGT 128
              ||||| |||| | |   ||| | |   | |     || ||   | ||||||| || |
Db        505 ACGAGTCCGGGCTGAAAGTTGAGGCGTTCGACTACGAAGTGGCGGCCGCGGCGACGGTGG 564

Qy        129 CTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGC 188
               |||| ||      | | ||  |     |   | | | || |||  |  | |||||
Db        565 ACGGCACCACGGTGGCCGAGGGCGGCGAAACGCTGGTGTCGGGCCGGCTGTTGGCGGCGA 624

Qy        189 CCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCT 248
              ||     ||   | || ||  ||  | |||||||     ||           | ||
Db        625 TCGCTAAGGCGTTGCCGAAGCGGGTGCCGGTGAAGTTTACGCACGACGGTGCGCGGGCTG 684

Qy        249 TTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
              | | |  || |||||     ||   |||| ||  | | |  | || |||||||
Db        685 TGGTGCAGGCGGGGGCCGCGGAGTTCACGCTGCCCACGATGGACCCGCGGGAG 737


RESULT 54
US-08-402-068-2
; Sequence 2, Application US/08402068
; Patent No. 5633159
;  GENERAL INFORMATION:
;    APPLICANT:  Pearson, Robert E.
;    APPLICANT:  Dickson, Julie A.
;    APPLICANT:  Hamilton, Paul T.
;    APPLICANT:  Little, Michael C.
;    APPLICANT:  Beyer Jr., Wayne F.
;    TITLE OF INVENTION:  MYCOBACTERIOPHAGE SPECIFIC FOR THE
;    TITLE OF INVENTION:  MYCOBACTERIUM TUBERCULOSIS COMPLEX
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard J. Rodrick, Becton Dickinson and
;      ADDRESSEE:  Company
;      STREET:  1 Becton Drive
;      CITY:  Franklin Lakes
;      STATE:  NJ
```

```
;       COUNTRY:  US
;       ZIP:  07417
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/402,068
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fugit, Donna R.
;       REGISTRATION NUMBER:  32,135
;       REFERENCE/DOCKET NUMBER:  P-3283
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6611 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  390..1538
;       OTHER INFORMATION:  /function= "coding sequence"
;       OTHER INFORMATION:  /product= "DNA polymerase"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  2107..3132
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  3138..3359
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  4690..5028
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  5028..5375
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  5375..5653
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  5653..5910
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  6078..6491
;       OTHER INFORMATION:  /function= "potential open reading
;       OTHER INFORMATION:  frame"
US-08-402-068-2

  Query Match             3.4%;  Score 44.2;  DB 2;  Length 6611;
  Best Local Similarity   49.4%;  Pred. No. 0.19;
  Matches  115;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy         69 ACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGT 128
              ||||| |||| | |  ||| | |  | |   || ||  | ||||||| || |
Db        505 ACGAGTCCGGGCTGAAAGTTGAGGCGTTCGACTACGAAGTGGCGGCCGCGGCGACGGTGG 564

Qy        129 CTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGC 188
```

```
              |||| ||     | | ||  |     |  |  | | || |||  |  | |||||
Db        565 ACGGCACCACGGTGGCCGAGGGCGGCGAAACGCTGGTGTCGGGCCGGCTGTTGGCGGCGA 624

Qy        189 CCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCT 248
              ||   ||   | || ||  ||  | |||||||     ||          | ||
Db        625 TCGCTAAGGCGTTGCCGAAGCGGGTGCCGGTGAAGTTTACGCACGACGGTGCGCGGGCTG 684

Qy        249 TTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
              | | |  || |||||     ||   |||| ||  | | |  | || |||||||
Db        685 TGGTGCAGGCGGGGGCCGCGGAGTTCACGCTGCCCACGATGGACCCGCGGGAG 737


RESULT 55
US-09-735-271-843/c
; Sequence 843, Application US/09735271
; Patent No. 6869762
; GENERAL INFORMATION:
;  APPLICANT: Daly, Mark J.
;  APPLICANT:  Hudson, Thomas J.
;  APPLICANT:  Lander, Eric S.
;  APPLICANT:  Rioux, John
;  APPLICANT:  Siminovitch, Kathy
;  TITLE OF INVENTION: IBD-RELATED POLYMORPHISMS
;  FILE REFERENCE: 2825.1025-002
;  CURRENT APPLICATION NUMBER: US/09/735,271
;  CURRENT FILING DATE:  2000-12-11
;  PRIOR APPLICATION NUMBER: US 60/170,257
;  PRIOR FILING DATE: 1999-12-10
;  PRIOR APPLICATION NUMBER: US 60/196,046
;  PRIOR FILING DATE: 2000-04-10
;  NUMBER OF SEQ ID NOS: 2058
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 843
;   LENGTH: 700
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(700)
;   OTHER INFORMATION: n = A,T,C or G
US-09-735-271-843

  Query Match             3.3%;  Score 42.8;  DB 3;  Length 700;
  Best Local Similarity   50.4%;  Pred. No. 0.15;
  Matches  130;  Conservative    0;  Mismatches  127;  Indels    1;  Gaps    1;

Qy         46 GCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTG 105
              | || || |||||||          ||||  ||  ||  || | |     |  || |||
Db        414 GGCCGCGGGGCCAATCCAGGCGCACAGGGAGCGGGGGCGGGGGCTGGGCCGAGTGGGGAC 355

Qy        106 AGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCC 165
              |||||     | | ||  ||||     || ||        |  |||||||  | | || |  ||
Db        354 AGGGGCGGGCGGGCCGGGGGCGGGGCGCGGCGTCCAGGCGGCGGCGGCGGCGGCTCC 295

Qy        166 GGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCA-CAGGCGGAGGAGGACCGGGTGAAGA 224
                | | |||   | |  || | |  ||  || |  |||  |    |    ||||
Db        294 TCAGAGTCCGGCTCAGGCTCCGGCTGCGGCTCCAGCCCGCGATGCCCCATTCCGTGACCC 235

Qy        225 GGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCC 284
              |||  ||| | |  |  ||||||     |||||||  ||  | |||  ||    |  ||
Db        234 TGCGCGGGCCTTCGCCCTGGGGCTTCCGCCTGGTGGGCGGCCGGGACTTCAGCGCGCCCC 175

Qy        285 CCAGCGTCCTGCGGGAGA 302
              || | ||   |||| ||
Db        174 TCACCATCTCACGGGTGA 157


RESULT 56
US-09-620-312D-711
; Sequence 711, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
```

```
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 711
;   LENGTH: 1745
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (124)..(1506)
US-09-620-312D-711

  Query Match             3.3%;  Score 42.8;  DB 3;  Length 1745;
  Best Local Similarity   70.5%;  Pred. No. 0.23;
  Matches   86;  Conservative    0;  Mismatches   32;  Indels    4;  Gaps    2;

Qy         20 TTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGG 79
              | ||||| |||||||||||||||||||||||||||||| |||||||||||||||||| |
Db          1 TAATACGGCTCACTATAGGGAATTTGGCCCTCGAGGCC-AGAATTCGGCACGAGGGCTGT 59

Qy         80 AAGCAGCGTGAAGAG---CGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
                |||   || |   | |  | |||  |||  |    | ||||   ||| | ||| ||
Db         60 TGTGCGCGCCCAGGGGTCCCGCGGGTTTTCGGGCGCAGGGTGGCGCCCGCGGCAGGCGGC 119

Qy        137 AG 138
                |
Db        120 GG 121


RESULT 57
US-08-455-073A-1
; Sequence 1, Application US/08455073A
; Patent No. 5876949
;  GENERAL INFORMATION:
;    APPLICANT:  Gideon Dreyfuss
;    APPLICANT:  Mikiko C. Siomi
;    APPLICANT:  Yan Zhang
;    TITLE OF INVENTION:  Fragile X Related Proteins, Compositions And Methods
;    TITLE OF INVENTION:  Of Making And Using The Same
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz & No. 5876949ris
;      STREET:  One Liberty Place, 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,073A
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  UPN-2201
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4362 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  220..2118
US-08-455-073A-1

  Query Match             3.2%;  Score 42.6;  DB 2;  Length 4362;
  Best Local Similarity   54.9%;  Pred. No. 0.42;
  Matches   84;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy         84 AGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCG 143
              |||| |    |||| ||   |  ||  ||       | |||||  ||||||    | |
Db          7 AGCGCGGGCGGCGGCGGTGACGGAGGCGCCGCTGCCAGGGGGCGTGCGGCAGCGCGGCGG 66

Qy        144 ATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGG 203
              |||||||  | | || | | |||  | |  ||||||| |||||  | |  || |    ||
Db         67 CGGCGGCGGCGGCGGCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGAGGCGG 126

Qy        204 CGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTC 236
              ||| || ||    || |  |  |    ||| ||
Db        127 CGGCGGCGGCGGCGGCGGCGGCGGCTGGGCCTC 159


RESULT 58
US-09-799-451-224/c
; Sequence 224, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Ghosh, Reena
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6783969el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 803
;  CURRENT APPLICATION NUMBER: US/09/799,451
;  CURRENT FILING DATE:  2001-03-05
;  NUMBER OF SEQ ID NOS: 948
;  SOFTWARE: pt_FL_genes Version 2.0
```

```
; SEQ ID NO 224
;   LENGTH: 2641
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (22)..(1587)
US-09-799-451-224

  Query Match             3.2%;  Score 42.4;  DB 3;  Length 2641;
  Best Local Similarity   53.7%;  Pred. No. 0.37;
  Matches   88;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy        113 CCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGG 172
              || |||||    ||||||| | |  | ||| | |||||   || || || |  | |
Db        230 CCCGCGGCCTGCGCGTCTGTCCCGGGGCCCGGCGGGGCGGGCGGCGTCGGCCCAGGGGCG 171

Qy        173 CGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGG 232
              |||||||  |||||  | |  || ||   |  |  || || | |||||  |  |  | |
Db        170 GCGGCGGCGGCGGCGGCGGCGGCGGCCGCGTTGCCGGGGGCCGGGGTGGCGGCGGCGTTG 111

Qy        233 CTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCAC 276
              | | |||  |||  |  |    |||   ||      ||| ||
Db        110 GCGCTGGGCCCGGGACTGCCCCCCGTGCCCGGGCCGACCCCGAC 67


RESULT 59
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             3.2%;  Score 42.2;  DB 3;  Length 4403765;
  Best Local Similarity   48.9%;  Pred. No. 17;
  Matches  113;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy        122 ATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCG 181
              |||||  | |       |  | | | ||||||  | | ||   ||||||  | || |||||||
Db     969830 ATGGCAACGGCGGCAATGGCGGGTCCGGCGGGGTCGGCCAAGCCGGCGGAGCCGGCGGCA 969771

Qy        182 TCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGC 241
              || ||     | |||||  |||||     ||   ||  | || ||| |  ||      |
Db     969770 GCGCCGGGTTGATCGGCATCGGCGGGACCGGGGGGGCCGGAGGGGCCGGCGCGGTCGGCG 969711

Qy        242 CTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
              |    ||  || ||| ||  |     | | |   ||  ||     | |  | |||
Db     969710 GCGTCGGCGGCAACGGGGGGTGGTTATACGGCAATGGCGGGGCCGGCGGGCTCGGGGGGA 969651

Qy        302 AACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAG 352
              |  |||| | ||   |  || |  | ||  |   ||| |||| | |
Db     969650 CCGGGGTGGCCGGCGTCAATGGCGGTATGGGCGCAGCAGGAGGTGCCGGCG 969600


RESULT 60
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             3.2%;  Score 42.2;  DB 3;  Length 4411529;
  Best Local Similarity   48.9%;  Pred. No. 18;
  Matches  113;  Conservative    0;  Mismatches  118;  Indels    0;  Gaps    0;

Qy        122 ATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCG 181
              |||||  | |       |  | | | ||||||  | | ||   ||||||  | || |||||||
Db     969827 ATGGCAACGGCGGCAATGGCGGGTCCGGCGGGGTCGGCCAAGCCGGCGGAGCCGGCGGCA 969768

Qy        182 TCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGC 241
              || ||     | | |||  |||||     ||   ||  | || ||| |  ||      |
Db     969767 GCGCCGGGTTGATCGGCATCGGCGGGACCGGGGGGGCCGGAGGGGCCGGCGCGGTCGGCG 969708

Qy        242 CTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
              |    ||  || ||| ||  |     | | |   ||  ||     | |  | |||
Db     969707 GCGTCGGCGGCAACGGGGGGTGGTTATACGGCAATGGCGGGGCCGGCGGGCTCGGGGGGA 969648

Qy        302 AACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAG 352
              |  |||| | ||   |  || |  | ||  |   ||| |||| | |
Db     969647 CCGGGGTGGCCGGCGTCAATGGCGGTATGGGCGCAGCAGGAGGTGCCGGCG 969597


RESULT 61
US-09-902-540-7042
; Sequence 7042, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 7042
;   LENGTH: 1845
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-7042

  Query Match             3.2%;  Score 42;  DB 3;  Length 1845;
  Best Local Similarity   45.7%;  Pred. No. 0.4;
  Matches  147;  Conservative    0;  Mismatches  175;  Indels    0;  Gaps    0;

Qy         82 GCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTC 141
              |||| | || |||  |   |   |  ||  | |||  ||| ||    | |  ||
Db       1263 GCAGGGCGAGGAGACGTACTCCGGCTTCGAGGCGGCGCAGATTGCCGAGGCCTTCAACCA 1322
```

```
Qy        142 CGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACA 201
              ||||||||  |      | |||   || ||  |||  |||||   |    |
Db       1323 ACTGTCGGCGGAGAGCCGCAACCAGGCGCTGCAGGGCGCGGTGGCGCGGTACTTCGGCGG 1382

Qy        202 GGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGG 261
              |||  |  |     | |||||  |  |    ||     |||||    | || |||
Db       1383 TGCGTCGCTGTCGAGCGTGAAGCTGGAGAACCAGGAGCAGGTGGGCGCGCCCTTCGTGCT 1442

Qy        262 GGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAA 321
              | | |||   |||  ||  | |  ||  | |  ||||  || |  || || |  |
Db       1443 CCGCTACGAGTTCACCGTGCCGCGCTTCGGGCGGATGGAGGGCGGCCAGCGGATGGCGCT 1502

Qy        322 AGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCC 381
              |  |    | ||| |  || || |   |              | |  || | ||  ||  |
Db       1503 GGGGCCCCTCACCTTCCCCGCGCAGTTGGGCCGGCGGTTCGTGCAGCTGAGCACGCGCCG 1562

Qy        382 CACGTCCTTCAAGTCCGAGGCC 403
              |||  |  || |   |||  ||
Db       1563 CACCCCGCTCTACATCGACACC 1584


RESULT 62
US-09-902-540-8385
; Sequence 8385, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8385
;   LENGTH: 3471
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-8385

  Query Match             3.2%;  Score 42;  DB 3;  Length 3471;
  Best Local Similarity   51.6%;  Pred. No. 0.54;
  Matches   96;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy        143 GATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAG 202
              || | || |||  |||| |  ||||| | |  |  |||  || ||   |  | |||  ||
Db       2548 GAGGAGGAGGAAGCCGCTGCTCCGGCCGAGGTGAAGGCCGCGCCGGATGCGGAAGCCGAG 2607

Qy        203 GCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGG 262
              | | | |  |||| || | ||  |  | |     ||  |  ||. | | |     |||
Db       2608 GGGAACGTCGACCTGGCGGAGGAGTTGAAGGACGAGCTCGAGGCCGTGGGCGGCGTGCCC 2667

Qy        263 GGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAA 322
              | | ||    | |   ||| |    || |||| | ||      |  |  |||| || ||
Db       2668 GAAGCCGCGGACCCCCAGGTGCAGGACGCCCTGGAGAAGCTGCGCGCGGGGACGAAGGAA 2727

Qy        323 GCCCTG 328
              | ||||
Db       2728 GGCCTG 2733


RESULT 63
US-09-902-540-625
; Sequence 625, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 625
;   LENGTH: 4087
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-625

  Query Match             3.2%;  Score 42;  DB 3;  Length 4087;
  Best Local Similarity   45.7%;  Pred. No. 0.59;
  Matches  147;  Conservative    0;  Mismatches  175;  Indels    0;  Gaps    0;

Qy         82 GCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTC 141
              |||| | || |||  |   |   |    ||  | |||  ||| ||    | |  ||
Db       1385 GCAGGGCGAGGAGACGTACTCCGGCTTCGAGGCGGCGCAGATTGCCGAGGCCTTCAACCA 1444

Qy        142 CGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACA 201
              ||||||||  |      | |||   || ||  |||  |||||   |    |
Db       1445 ACTGTCGGCGGAGAGCCGCAACCAGGCGCTGCAGGGCGCGGTGGCGCGGTACTTCGGCGG 1504

Qy        202 GGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGG 261
              |||  |  |     | ||||||  |  |    ||     |||||    | || |||
Db       1505 TGCGTCGCTGTCGAGCGTGAAGCTGGAGAACCAGGAGCAGGTGGGCGCGCCCTTCGTGCT 1564

Qy        262 GGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAA 321
              | | |||   |||  ||  | |  ||  | |  ||||  || |  || || |  |
Db       1565 CCGCTACGAGTTCACCGTGCCGCGCTTCGGGCGGATGGAGGGCGGCCAGCGGATGGCGCT 1624

Qy        322 AGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCC 381
              |  |    | ||| |  || || |   |              | |  || | ||  ||  |
Db       1625 GGGGCCCCTCACCTTCCCCGCGCAGTTGGGCCGGCGGTTCGTGCAGCTGAGCACGCGCCG 1684

Qy        382 CACGTCCTTCAAGTCCGAGGCC 403
              |||  |  || |   |||  ||
Db       1685 CACCCCGCTCTACATCGACACC 1706


RESULT 64
US-09-902-540-879
; Sequence 879, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 879
;   LENGTH: 7186
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-879

  Query Match             3.2%;  Score 42;  DB 3;  Length 7186;
  Best Local Similarity   51.6%;  Pred. No. 0.79;
  Matches   96;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy        143 GATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAG 202
              || | || |||  |||| |  ||||| | |  |  |||  || ||   |  | |||  ||
Db       2549 GAGGAGGAGGAAGCCGCTGCTCCGGCCGAGGTGAAGGCCGCGCCGGATGCGGAAGCCGAG 2608
```

```
Qy        203 GCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGG 262
              | | | |  |||| || | ||  |  | |   || | || | | |     |||
Db       2609 GGGAACGTCGACCTGGCGGAGGAGTTGAAGGACGAGCTCGAGGCCGTGGGCGGCGTGCCC 2668

Qy        263 GGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAA 322
              | |  ||     | |  ||| |    || ||||  | ||    |  |  |||| || ||
Db       2669 GAAGCCGCGGACCCCCAGGTGCAGGACGCCCTGGAGAAGCTGCGCGCGGGGACGAAGGAA 2728

Qy        323 GCCCTG 328
              | |||||
Db       2729 GGCCTG 2734
```

RESULT 65
US-09-949-016-4238
; Sequence 4238, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4238
; LENGTH: 8865
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4238

```
  Query Match             3.2%;  Score 42;  DB 3;  Length 8865;
  Best Local Similarity   54.7%;  Pred. No. 0.87;
  Matches  105;  Conservative    0;  Mismatches   85;  Indels    2;  Gaps    1;

Qy        126 CGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGG 185
              ||  |  |||  |  ||   |||| || | | || | | ||||||  | |||| ||| |||
Db       6699 CGAGTACCAGAGGAACCACAGCGGGGGCGGCGGCGGCGGCAGCGGGGGCAGCGGCGG 6758

Qy        186 CGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
              |   |  |||||  |||||  || |  ||  |||  | ||| |||||  |||||  | |
Db       6759 GGGTGGGGGCAGCGGCGGCGGGGGCCCCCAGTG--GCGGCAGCGGCCACAGTGGCGGC 6816

Qy        246 GCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACG 305
              |      | ||  | || |   | | | |  |  ||  ||   | ||| |         |
Db       6817 CCCAGCAGCTGCGGCGGCGCCCCCAGCACGAGCAGGAGCCGGCCCTCCCGGATCCCCCAG 6876

Qy        306 ACTGGCAGACGC 317
              ||| | ||| |
Db       6877 CCTGTCCGACAC 6888
```

RESULT 66
US-09-949-016-15980
; Sequence 15980, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15980
; LENGTH: 239527
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(239527)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15980

```
  Query Match             3.2%;  Score 42;  DB 3;  Length 239527;
  Best Local Similarity   54.7%;  Pred. No. 4.6;
  Matches  105;  Conservative    0;  Mismatches   85;  Indels    2;  Gaps    1;

Qy        126 CGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGG 185
              ||  |  |||  |  ||   |||| || | | || | | ||||||  | |||| ||| |||
Db     216328 CGAGTACCAGAGGAACCACAGCGGGGGCGGCGGCGGCGGCAGCGGGGGCAGCGGCGG 216387

Qy        186 CGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
              |   |  |||||  |||||  || |  ||  |||  | ||| |||||  |||||  | |
Db     216388 GGGTGGGGGCAGCGGCGGCGGCGGGGCCCCCAGTG--GCGGCAGCGGCCACAGTGGCGGC 216445

Qy        246 GCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACG 305
              |      | ||  | || |   | | | |  |  ||  ||   | ||| |         |
Db     216446 CCCAGCAGCTGCGGCGGCGCCCCCAGCACGAGCAGGAGCCGGCCCTCCCGGATCCCCCAG 216505

Qy        306 ACTGGCAGACGC 317
              ||| | ||| |
Db     216506 CCTGTCCGACAC 216517
```

RESULT 67
US-09-620-312D-728
; Sequence 728, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 728
; LENGTH: 1493
; TYPE: DNA
; ORGANISM: Homo sapiens

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (397)..(1341)
US-09-620-312D-728

  Query Match             3.2%;  Score 41.8;  DB 3;  Length 1493;
  Best Local Similarity   95.6%;  Pred. No. 0.4;
  Matches   43;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         34 ATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGG 78
              |||||||||||||||||||||||||||||||||||||||||  ||
Db          1 ATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAAGG 45


RESULT 68
US-09-902-540-1145
; Sequence 1145, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1145
;   LENGTH: 14382
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-1145

  Query Match             3.2%;  Score 41.8;  DB 3;  Length 14382;
  Best Local Similarity   48.9%;  Pred. No. 1.3;
  Matches  112;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy          61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120
               ||| |||   | |||| ||||   |    |    | || |  ||  |||||  | |||
Db        6094 AATGCGGGTGGTGGGCCGGAACTTGTTGCCCGGCGCGTTGGTGGGAACGGACCGGGTGGC 6153

Qy         121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180
               |    | |   |  | |   |   | | |   ||||||| | || |||  |  |||  ||
Db        6154 GTGCTTGCCCTTCGTCGGCATCACCGTGCCTTCGCCCGCGGCGCAGGCGAGCAGGGGCGC 6213

Qy         181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240
               | |||||  ||  || || | || || || ||       |||  |||| ||  | |
Db        6214 GGCGGCGAACGCGGCGGCCCCGGTGGCGGCGGCGGAAAGGAATTGGCGTCGAGTCCGCAT 6273

Qy         241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGC 289
                   |||   |  |  |||   |||  | | ||   |  | || |
Db        6274 GTCTTCCTTCAGGAAGTCCGGGCCGGGCGGGCTCGCGGGTGCGCGCATC 6322


RESULT 69
US-09-056-556-182
; Sequence 182, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND            TREATM
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  182:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  985 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-182

  Query Match             3.2%;  Score 41.6;  DB 3;  Length 985;
  Best Local Similarity   51.0%;  Pred. No. 0.37;
  Matches   98;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy         108 GGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGG 167
               ||| | || ||||  || || | | ||       || | |||| |  | | |||| |||||
Db         345 GGGCCTCTCCGGCTTTGACGGCGGCCAAGGCGGCCAAGGCGGGGCCGGCGGCAGCGCCGG 404

Qy         168 CAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGC 227
               |   |  |||  |  ||| ||| | |   |   |||   || ||    || |  || | |
Db         405 CGCCGGCGGCATCAACGGGGCCGGCGGGGCCGGCGGCAACGGCGGCGACGGCGGGGACGG 464

Qy         228 GGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCA 287
               ||  |    |  | || ||  |||      || ||| | |       ||| |||  ||
Db         465 CGCAACCGGTGCCGCAGGTCTCGGCGACAACGGCGGGGTCGGCGGTGACGGTGGGGCCGG 524

Qy         288 GCGTCCTGCGGG 299
               |  | || ||
Db         525 TGGCGCCGCCGG 536


RESULT 70
US-09-072-596-177
; Sequence 177, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  177:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  985 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-177

  Query Match             3.2%;  Score 41.6;  DB 3;  Length 985;
  Best Local Similarity   51.0%;  Pred. No. 0.37;
  Matches   98;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        108 GGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGG 167
              ||| | || |||| || || | | ||      || | |||| | | | |||| |||||
Db        345 GGGCCTCTCCGGCTTTGACGGCGGCCAAGGCGGCCAAGGCGGGGCCGGCGGCAGCGCCGG 404

Qy        168 CAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGC 227
              |   | ||| |  ||| ||| | |   |  |||  || ||    || |  || |
Db        405 CGCCGGCGGCATCAACGGGGCCGGCGGGGCCGGCGGCAACGGCGGCGACGGCGGGGACGG 464

Qy        228 GGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCA 287
              ||  |   |  | || ||  |||     || ||| |  |     ||| |||  ||
Db        465 CGCAACCGGTGCCGCAGGTCTCGGCGACAACGGCGGGGTCGGCGGTGACGGTGGGGCCGG 524

Qy        288 GCGTCCTGCGGG 299
                |  | || ||
Db        525 TGGCGCCGCCGG 536


RESULT 71
US-09-072-967-182
; Sequence 182, Application US/09072967
; Patent No. 6592877
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  355
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,967
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  182:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  985 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-967-182

  Query Match             3.2%;  Score 41.6;  DB 3;  Length 985;
  Best Local Similarity   51.0%;  Pred. No. 0.37;
  Matches   98;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        108 GGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGG 167
              ||| | || |||| || || | | ||      || | |||| | | | |||| |||||
Db        345 GGGCCTCTCCGGCTTTGACGGCGGCCAAGGCGGCCAAGGCGGGGCCGGCGGCAGCGCCGG 404

Qy        168 CAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGC 227
              |   | ||| |  ||| ||| | |   |  |||  || ||    || |  || |
Db        405 CGCCGGCGGCATCAACGGGGCCGGCGGGGCCGGCGGCAACGGCGGCGACGGCGGGGACGG 464

Qy        228 GGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCA 287
              ||  |   |  | || ||  |||     || ||| |  |     ||| |||  ||
Db        465 CGCAACCGGTGCCGCAGGTCTCGGCGACAACGGCGGGGTCGGCGGTGACGGTGGGGCCGG 524

Qy        288 GCGTCCTGCGGG 299
                |  | || ||
Db        525 TGGCGCCGCCGG 536


RESULT 72
US-10-193-002-177
; Sequence 177, Application US/10193002
; Patent No. 6949246
;  GENERAL INFORMATION:
;       APPLICANT: Reed, Steven G.
;                  Skeiky, Yasir A.W.
;                  Dillon, Davin C.
;                  Campos-Neto, Antonia
;                  Houghton, Raymond
;                  Vedvick, Thomas S.
;                  Twardzik, Daniel R.
;                  Lodes, Michael J.
;                  Hendrickson, Ronald C.
;       TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                           TUBERCULOSIS
;       NUMBER OF SEQUENCES: 350
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SEED and BERRY LLP
;            STREET: 6300 Columbia Center, 701 Fifth Avenue
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/193,002
;            FILING DATE: 10-Jul-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/072,596
;            FILING DATE: 05-MAY-1998
;
```

```
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Maki, David J.
;            REGISTRATION NUMBER: 31,392
;            REFERENCE/DOCKET NUMBER: 210121.417C9
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;  INFORMATION FOR SEQ ID NO: 177:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 985 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 177:
US-10-193-002-177

  Query Match             3.2%;  Score 41.6;  DB 3;  Length 985;
  Best Local Similarity   51.0%;  Pred. No. 0.37;
  Matches   98;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        108 GGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGG 167
              ||| | || |||| || || | | ||       || | |||| | | | |||| |||||
Db        345 GGGCCTCTCCGGCTTTGACGGCGGCCAAGGCGGCCAAGGCGGGGCCGGCGGCAGCGCCGG 404

Qy        168 CAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGC 227
              |   | ||| | ||| ||| | |   |   ||| || ||    || |  || |
Db        405 CGCCGGCGGCATCAACGGGGCCGGCGGGGCCGGCGGCAACGGCGGCGACGGCGGGGACGG 464

Qy        228 GGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCA 287
              ||  |   | || ||  |||     || ||| |  |     ||| |||  ||
Db        465 CGCAACCGGTGCCGCAGGTCTCGGCGACAACGGCGGGGTCGGCGGTGACGGTGGGGCCGG 524

Qy        288 GCGTCCTGCGGG 299
              |  | || ||
Db        525 TGGCGCCGCCGG 536


RESULT 73
US-10-084-843-182
; Sequence 182, Application US/10084843
; Patent No. 6962710
;   GENERAL INFORMATION:
;       APPLICANT: Reed, Steven G.
;                  Skeiky, Yasir A.W.
;                  Dillon, Davin C.
;                  Campos-Neto, Antonio
;                  Houghton, Raymond
;                  Vedvick, Thomas S.
;                  Twardzik, Daniel R.
;                  Lodes, Michael J.
;                  Hendrickson, Ronald C.
;       TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;                           AND DIAGNOSIS OF TUBERCULOSIS
;       NUMBER OF SEQUENCES: 355
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SEED and BERRY LLP
;            STREET: 6300 Columbia Center, 701 Fifth Avenue
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/084,843
;            FILING DATE: 25-Feb-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/072,967
;            FILING DATE: 05-MAY-1998
```

```
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Maki, David J.
;            REGISTRATION NUMBER: 31,392
;            REFERENCE/DOCKET NUMBER: 210121.411C9
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;  INFORMATION FOR SEQ ID NO: 182:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 985 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 182:
US-10-084-843-182

  Query Match             3.2%;  Score 41.6;  DB 3;  Length 985;
  Best Local Similarity   51.0%;  Pred. No. 0.37;
  Matches   98;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        108 GGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGG 167
              ||| | || |||| || || | | ||       || | |||| | | | |||| |||||
Db        345 GGGCCTCTCCGGCTTTGACGGCGGCCAAGGCGGCCAAGGCGGGGCCGGCGGCAGCGCCGG 404

Qy        168 CAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGC 227
              |   | ||| | ||| ||| | |   |   ||| || ||    || |  || |
Db        405 CGCCGGCGGCATCAACGGGGCCGGCGGGGCCGGCGGCAACGGCGGCGACGGCGGGGACGG 464

Qy        228 GGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCA 287
              ||  |   | || ||  |||     || ||| |  |     ||| |||  ||
Db        465 CGCAACCGGTGCCGCAGGTCTCGGCGACAACGGCGGGGTCGGCGGTGACGGTGGGGCCGG 524

Qy        288 GCGTCCTGCGGG 299
              |  | || ||
Db        525 TGGCGCCGCCGG 536


RESULT 74
US-09-902-540-977
; Sequence 977, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 977
;   LENGTH: 10178
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-977

  Query Match             3.2%;  Score 41.6;  DB 3;  Length 10178;
  Best Local Similarity   49.1%;  Pred. No. 1.2;
  Matches  110;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

Qy        149 GCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAG 208
              ||   |||||| | ||   |  | | |||| |  |  ||| |  |  ||   || | |
Db       7869 GCCACGCCCGCCGCGCGCACGCCGGGCGCGGTGGGGCTGCCGGTGGGCGCGGCGGTGCTG 7928

Qy        209 GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGC 268
              | ||||| ||   || | | |  ||  |     |  ||| | |||||  | | |    ||
Db       7929 GTGGACCTGGAAGAGGGCCCGGTGCCGCTCACCGCGGGTTATGCGTGCGCGTTGACGCGC 7988

Qy        269 GACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTG 328
              ||  |  || || |      | | |||  || |    | | || |  |  | |   ||
```

```
Db       7989 GAGGCGGCGTTGCTGAAGGCCTTGCTGGAGGCGGCGCAGTCGCGGCTGACGGACATCCAC 8048

Qy        329 AGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCG 372
              ||||   |   ||  || | |     ||||  | |  | ||||
Db       8049 GGCGCGCGCGAGGACGTGGCCGCCGCGGACCGGGAGGCGGCGCG 8092


RESULT 75
US-09-902-540-5035/c
; Sequence 5035, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5035
;   LENGTH: 2755
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-5035

  Query Match             3.2%;  Score 41.4;  DB 3;  Length 2755;
  Best Local Similarity   57.3%;  Pred. No. 0.71;
  Matches   75;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy        152 GAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAG 211
              | |  ||||||  | || || || |||| | ||||    |  || || |  |||  || |
Db        275 GCGGGCGCAGGCGCTGCGGGACGCGCGGTGGCGGCCTGGGCGGCGGCCCCCGCGTCGGCG 216

Qy        212 GACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGAC 271
              | |||| ||| | |  || || |   || ||   ||      |  || |  ||   ||||
Db        215 GCCCGGCTGAGGCGCAGGTGGATGAGGTCCCCTCGCCGCATGGCGGCGATGGACAGCGAG 156

Qy        272 CCCACGATGGT 282
              |||| |  |||
Db        155 CCCAGGCCGGT 145


Search completed: March 11, 2006, 14:55:24
Job time : 277 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 00:16:23 ; Search time 229.027 Seconds
(without alignments)
1588.688 Million cell updates/sec

Title: US-10-757-745-6_COPY_2_45
Perfect score: 229
Sequence: 1 GDPGSEIIESVPPAGPEASE.....ENEDDIQFVSEGPSRPVLEY 44

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170337_4279/app_query.fasta_1
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss08
-USER=US10757745_@CGN_1_1_2715_@runat_09032006_170337_4279 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	229	100.0	1536	9	US-10-757-745-5	Sequence 5, Appli
2	223	97.4	3042	7	US-10-362-463-5	Sequence 5, Appli
3	223	97.4	3042	7	US-10-362-463-6	Sequence 6, Appli
4	223	97.4	3183	7	US-10-362-463-9	Sequence 9, Appli
5	223	97.4	3183	7	US-10-362-463-10	Sequence 10, Appl
6	223	97.4	4961	8	US-10-775-169-139	Sequence 139, App
7	223	97.4	4999	7	US-10-362-463-1	Sequence 1, Appli
8	223	97.4	4999	7	US-10-362-463-2	Sequence 2, Appli
9	223	97.4	5144	7	US-10-362-463-3	Sequence 3, Appli
10	223	97.4	5144	7	US-10-362-463-4	Sequence 4, Appli
11	217	94.8	3363	9	US-10-450-763-17121	Sequence 17121, A
12	175	76.4	394	9	US-10-450-763-15209	Sequence 15209, A
13	175	76.4	394	9	US-10-450-763-17120	Sequence 17120, A
14	175	76.4	448	3	US-09-918-995-28032	Sequence 28032, A
15	175	76.4	2039	9	US-10-450-763-17122	Sequence 17122, A
c 16	75.5	33.0	2688	6	US-10-369-493-33508	Sequence 33508, A
17	68	29.7	430	3	US-09-918-995-5409	Sequence 5409, Ap
18	68	29.7	493	3	US-09-918-995-24799	Sequence 24799, A
19	68	29.7	493	3	US-09-918-995-32051	Sequence 32051, A
20	68	29.7	500	3	US-09-918-995-30429	Sequence 30429, A
21	68	29.7	2438	6	US-10-264-049-796	Sequence 796, App
22	68	29.7	3713	9	US-10-491-467-70	Sequence 70, Appl
c 23	68	29.7	3818	7	US-10-112-944-663	Sequence 663, App
24	67	29.3	8100	3	US-09-190-246-4	Sequence 4, Appli
25	67	29.3	10610	3	US-09-994-412-3	Sequence 3, Appli
26	67	29.3	11517	3	US-09-901-106-1	Sequence 1, Appli
27	67	29.3	15538	3	US-09-190-246-1	Sequence 1, Appli
c 28	67	29.3	77536	3	US-09-940-316B-1	Sequence 1, Appli
29	65	28.4	2436	7	US-10-320-797-2316	Sequence 2316, Ap
30	65	28.4	3346	7	US-10-320-797-1316	Sequence 1316, Ap
31	65	28.4	5346	7	US-10-320-797-316	Sequence 316, App
32	65	28.4	77536	3	US-09-940-316B-1	Sequence 1, Appli
33	64	27.9	7390	9	US-10-765-700-95	Sequence 95, Appl
34	64	27.9	7666	7	US-10-363-616-164	Sequence 164, App
35	64	27.9	7758	5	US-10-171-581-295	Sequence 295, App
36	64	27.9	7758	8	US-10-723-860-1239	Sequence 1239, Ap
37	64	27.9	7792	8	US-10-723-860-5743	Sequence 5743, Ap
38	64	27.9	171096	8	US-10-612-869-8	Sequence 8, Appli
39	63.5	27.7	428	7	US-10-424-599-48990	Sequence 48990, A
40	63	27.5	270	7	US-10-437-963-59361	Sequence 59361, A
c 41	63	27.5	2736	7	US-10-437-963-82114	Sequence 82114, A
42	63	27.5	3120	7	US-10-112-944-191	Sequence 191, App
c 43	63	27.5	9795	9	US-10-937-379-44	Sequence 44, Appl
c 44	63	27.5	73599	9	US-10-937-379-24	Sequence 24, Appl
45	63	27.5	9025608	6	US-10-156-761-1	Sequence 1, Appli
c 46	62	27.1	557	3	US-09-867-701-5034	Sequence 5034, Ap
47	62	27.1	651	7	US-10-767-701-2245	Sequence 2245, Ap
c 48	62	27.1	712	9	US-10-956-157-5003	Sequence 5003, Ap
c 49	62	27.1	712	9	US-10-956-157-10238	Sequence 10238, A
50	62	27.1	762	9	US-10-927-644-4	Sequence 4, Appli
51	62	27.1	1056	3	US-09-778-844-110	Sequence 110, App
52	62	27.1	1135	3	US-09-778-844-108	Sequence 108, App
53	62	27.1	1137	3	US-09-778-844-109	Sequence 109, App
54	62	27.1	3393	9	US-10-927-644-2	Sequence 2, Appli
55	62	27.1	3393	10	US-11-067-321A-3	Sequence 3, Appli
c 56	62	27.1	4984	6	US-10-116-788A-32	Sequence 32, Appl
c 57	62	27.1	7518	9	US-10-494-484A-1	Sequence 1, Appli
58	61.5	26.9	494	7	US-10-767-701-3659	Sequence 3659, Ap
59	61.5	26.9	1031	6	US-10-419-341-4	Sequence 4, Appli
60	61	26.6	380	8	US-10-425-115-35966	Sequence 35966, A
61	61	26.6	393	3	US-09-796-692-2534	Sequence 2534, Ap
62	61	26.6	393	5	US-10-040-862-2534	Sequence 2534, Ap
63	61	26.6	393	6	US-10-057-475B-2534	Sequence 2534, Ap
64	61	26.6	393	6	US-10-154-884B-2534	Sequence 2534, Ap
65	61	26.6	393	8	US-10-764-324-2534	Sequence 2534, Ap
66	61	26.6	445	3	US-09-918-995-29169	Sequence 29169, A
67	61	26.6	2502	6	US-10-172-118-1080	Sequence 1080, Ap
68	61	26.6	2502	7	US-10-342-887-1080	Sequence 1080, Ap
69	61	26.6	3727	5	US-10-171-581-49	Sequence 49, Appl
70	61	26.6	4184	8	US-10-483-506-24	Sequence 24, Appl
c 71	61	26.6	4466	3	US-09-940-316B-20	Sequence 20, Appl
c 72	61	26.6	4478	3	US-09-940-316B-16	Sequence 16, Appl
c 73	61	26.6	4547	3	US-09-940-316B-22	Sequence 22, Appl
c 74	61	26.6	4571	3	US-09-940-316B-18	Sequence 18, Appl
75	61	26.6	4587	10	US-11-097-143-2012	Sequence 2012, Ap

ALIGNMENTS

RESULT 1

```
US-10-757-745-5
; Sequence 5, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 1536
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1530)..(1531)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (209)..(1534)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1392)..(1392)
;   OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-5

Alignment Scores:
Pred. No.:                1.79e-23        Length:          1536
Score:                    229.00          Matches:         44
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-757-745-5 (1-1536)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 271

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 331

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        332 GTTCTTGAATAC 343

RESULT 2
US-10-362-463-5
; Sequence 5, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
;  APPLICANT: Akzo No. US20040059089A1el N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 3042
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-5

Alignment Scores:
Pred. No.:                2.9e-22         Length:          3042
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-5 (1-3042)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 3
US-10-362-463-6
; Sequence 6, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
;  APPLICANT: Akzo No. US20040059089A1el N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 3042
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-6

Alignment Scores:
Pred. No.:                2.9e-22         Length:          3042
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-6 (1-3042)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 4
```

```
US-10-362-463-9
; Sequence 9, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
;  APPLICANT: Akzo No. US20040059089A1el N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 3183
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-9

Alignment Scores:
Pred. No.:                3.05e-22        Length:         3183
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-9 (1-3183)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 5
US-10-362-463-10
; Sequence 10, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
;  APPLICANT: Akzo No. US20040059089A1el N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 3183
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-10

Alignment Scores:
Pred. No.:                3.05e-22        Length:         3183
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-10 (1-3183)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 6
US-10-775-169-139
; Sequence 139, Application US/10775169
; Publication No. US20040175743A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Burczynski, Michael
;  APPLICANT:  Twine, Natalie
;  APPLICANT:  Dorner, Andrew
;  APPLICANT:  Trepicchio, William
;  TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
;  FILE REFERENCE: AM101080 (031896-013000)
;  CURRENT APPLICATION NUMBER: US/10/775,169
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 5278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 139
;   LENGTH: 4961
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-775-169-139

Alignment Scores:
Pred. No.:                4.99e-22        Length:         4961
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-775-169-139 (1-4961)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 106

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        107 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 166

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        167 GTTCTTGAATAC 178

RESULT 7
US-10-362-463-1
; Sequence 1, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
;  APPLICANT: Akzo No. US20040059089A1el N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
```

```
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4999
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-1

Alignment Scores:
Pred. No.:                5.03e-22        Length:          4999
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-1 (1-4999)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         349 GTTCTTGAATAC 360

RESULT 8
US-10-362-463-2
; Sequence 2, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
;  APPLICANT: Akzo No. US20040059089A1el N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4999
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-2

Alignment Scores:
Pred. No.:                5.03e-22        Length:          4999
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-2 (1-4999)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         349 GTTCTTGAATAC 360

RESULT 9
US-10-362-463-3
; Sequence 3, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
;  APPLICANT: Akzo No. US20040059089A1el N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 5144
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-3

Alignment Scores:
Pred. No.:                5.19e-22        Length:          5144
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-3 (1-5144)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         349 GTTCTTGAATAC 360

RESULT 10
US-10-362-463-4
; Sequence 4, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
;  APPLICANT: Akzo No. US20040059089A1el N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 5144
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-4

Alignment Scores:
Pred. No.:                5.19e-22        Length:          5144
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       7               Gaps:            0
```

```
US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-4 (1-5144)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 11
US-10-450-763-17121
; Sequence 17121, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 17121
;   LENGTH: 3363
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (771)..(211)
;   OTHER INFORMATION: 98% homologous to Escherichia coli seryl-tRNA
;   OTHER INFORMATION: synthetase,accession number X05017,Smith-Waterman Score=934.
US-10-450-763-17121

Alignment Scores:
Pred. No.:              2.47e-21       Length:        3363
Score:                  217.00         Matches:       42
Percent Similarity:     95.5%          Conservative:  0
Best Local Similarity:  95.5%          Mismatches:    2
Query Match:            94.8%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-450-763-17121 (1-3363)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        181 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATTTGAG 240

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        241 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 300

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        301 GTTCTTGAATAC 312

RESULT 12
US-10-450-763-15209
; Sequence 15209, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 15209
;   LENGTH: 394
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (79)..(183)
;   OTHER INFORMATION: 100% homologous to Homo sapiens Human secreted protein, SEQ
;   OTHER INFORMATION: ID NO: 5728,accession number G01647,Smith-Waterman Score=180.
US-10-450-763-15209

Alignment Scores:
Pred. No.:              3.42e-16       Length:        394
Score:                  175.00         Matches:       34
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            76.4%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-450-763-15209 (1-394)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 141

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              ||||||||||||||||||||||||||||||||||||||||||
Db        142 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGT 183

RESULT 13
US-10-450-763-17120
; Sequence 17120, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 17120
;   LENGTH: 394
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (79)..(183)
;   OTHER INFORMATION: 100% homologous to Homo sapiens Human secreted protein, SEQ
;   OTHER INFORMATION: ID NO: 5728,accession number G01647,Smith-Waterman Score=180.
US-10-450-763-17120

Alignment Scores:
Pred. No.:              3.42e-16       Length:        394
Score:                  175.00         Matches:       34
Percent Similarity:     100.0%         Conservative:  0
```

```
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            76.4%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-450-763-17120 (1-394)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 141

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              ||||||||||||||||||||||||||||||||||||
Db        142 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGT 183

RESULT 14
US-09-918-995-28032
; Sequence 28032, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 28032
;   LENGTH: 448
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(448)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-28032

Alignment Scores:
Pred. No.:              3.94e-16        Length:         448
Score:                  175.00          Matches:        34
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            76.4%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-918-995-28032 (1-448)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 178

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              ||||||||||||||||||||||||||||||||||||
Db        179 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGT 220

RESULT 15
US-10-450-763-17122
; Sequence 17122, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 17122
;   LENGTH: 2039
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (121)..(561)
;   OTHER INFORMATION: 98% homologous to Escherichia coli similar to,accession
;   OTHER INFORMATION: number D90875,Smith-Waterman Score=762.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2039)
;   OTHER INFORMATION: n = a,t,c or g
US-10-450-763-17122

Alignment Scores:
Pred. No.:              2.12e-15        Length:         2039
Score:                  175.00          Matches:        34
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            76.4%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-450-763-17122 (1-2039)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 130

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              ||||||||||||||||||||||||||||||||||||
Db        131 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGT 172

RESULT 16
US-10-369-493-33508/c
; Sequence 33508, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33508
;   LENGTH: 2688
;   TYPE: DNA
;   ORGANISM: Desulfitobacterium hafniense
US-10-369-493-33508

Alignment Scores:
Pred. No.:              1.21            Length:         2688
Score:                  75.50           Matches:        15
Percent Similarity:     61.1%           Conservative:   7
Best Local Similarity:  41.7%           Mismatches:     13
Query Match:            33.0%           Indels:         1
DB:                     6               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-369-493-33508 (1-2688)

Qy          1 GlyAspProGlySerGlu---IleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||||||   |||   :::|||||:::   ||||||      |||
Db        819 GGAGACCCCGGCCAGGAACATCTTATAGAAGCCTTCCCCCCGCAGAACCCCGGTCAGGAC 760
```

```
Qy         20 GluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
              :::|||   |||:::|||   |||:::   |||      ::::::
Db        759 TACAATACCGCAGACGATAATACTGATAATGACCAAATATTTGCCCAG 712

RESULT 17
US-09-918-995-5409
; Sequence 5409, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5409
;   LENGTH: 430
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(430)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-5409

Alignment Scores:
Pred. No.:              2.01          Length:         430
Score:                  68.00         Matches:        14
Percent Similarity:     48.8%         Conservative:   7
Best Local Similarity:  32.6%         Mismatches:     22
Query Match:            29.7%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-918-995-5409 (1-430)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||          |||
Db        197 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 256

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||::::::      |||   :::|||          |||   |||:::
Db        257 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 316

Qy         41 ValLeuGlu 43
              :::|||
Db        317 GAGATGGAG 325

RESULT 18
US-09-918-995-24799
; Sequence 24799, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24799
;   LENGTH: 493
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(493)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-24799

Alignment Scores:
Pred. No.:              2.34          Length:         493
Score:                  68.00         Matches:        14
Percent Similarity:     48.8%         Conservative:   7
Best Local Similarity:  32.6%         Mismatches:     22
Query Match:            29.7%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-918-995-24799 (1-493)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||          |||
Db        178 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 237

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||::::::      |||   :::|||          |||   |||:::
Db        238 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 297

Qy         41 ValLeuGlu 43
              :::|||
Db        298 GAGATGGAG 306

RESULT 19
US-09-918-995-32051
; Sequence 32051, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32051
;   LENGTH: 493
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(493)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-32051

Alignment Scores:
Pred. No.:              2.34          Length:         493
Score:                  68.00         Matches:        14
Percent Similarity:     48.8%         Conservative:   7
Best Local Similarity:  32.6%         Mismatches:     22
Query Match:            29.7%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-918-995-32051 (1-493)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||          |||
Db        336 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 395

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||::::::      |||   :::|||          |||   |||:::
Db        396 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 455

Qy         41 ValLeuGlu 43
              :::|||
Db        456 GAGATGGAG 464

RESULT 20
```

```
US-09-918-995-30429
; Sequence 30429, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 30429
;   LENGTH: 500
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(500)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30429

Alignment Scores:
Pred. No.:                2.38          Length:          500
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-918-995-30429 (1-500)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||         |||
Db        329 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 388

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::      |||   :::|||          |||   |||:::
Db        389 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 448

Qy         41 ValLeuGlu 43
                :::|||
Db        449 GAGATGGAG 457

RESULT 21
US-10-264-049-796
; Sequence 796, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA133P1
;  CURRENT APPLICATION NUMBER: US/10/264,049
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/18569
;  PRIOR FILING DATE: 2001-06-07
;  PRIOR APPLICATION NUMBER: US 60/209,467
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 4360
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 796
;   LENGTH: 2438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-264-049-796

Alignment Scores:
Pred. No.:                13.8          Length:          2438
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-264-049-796 (1-2438)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||         |||
Db       2081 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 2140

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::      |||   :::|||          |||   |||:::
Db       2141 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 2200

Qy         41 ValLeuGlu 43
                :::|||
Db       2201 GAGATGGAG 2209

RESULT 22
US-10-491-467-70
; Sequence 70, Application US/10491467
; Publication No. US20050186568A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE CORPORATION; BANDMAN, Olga;
;  APPLICANT:  BAUGHN, Mariah R.; BECHA, Shanya D.;
;  APPLICANT:  BOROWSKY, Mark L.; DUGGAN, Brendan M.;
;  APPLICANT:  EMERLING, Brooke M.; FORSYTHE, Ian J.;
;  APPLICANT:  GANDHI, Ameena R.; GORVAD, Ann E.;
;  APPLICANT:  GRIFFIN, Jennifer A.; GURURAJAN, Rajagopal;
;  APPLICANT:  HAFALIA, April J.A.; KHAN, Farrah A.;
;  APPLICANT:  LAL, Preeti G.; LEE, Ernestine A.;
;  APPLICANT:  LEE, Soo Yeun; LINDQUIST, Erika A.;
;  APPLICANT:  LU, Dyung Aina M.; LU, Yan;
;  APPLICANT:  MARQUIS, Joseph P.; NGUYEN, Danniel B.;
;  APPLICANT:  ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi;
;  APPLICANT:  RECIPON, Shirley A.; RICHARDSON, Thomas W.;
;  APPLICANT:  SWARNAKAR, Anita; TANG, Y. Tom;
;  APPLICANT:  THORNTON, Michael B.; TRAN, Uyen K.;
;  APPLICANT:  CHAWLA, Narinder K.; WARREN, Bridget A.;
;  APPLICANT:  YANG, Junming; YAO, Monique G.;
;  APPLICANT:  YUE, Henry; ZEBARJADIAN, Yeganeh
;  TITLE OF INVENTION: KINASES AND PHOSPHATASES
;  FILE REFERENCE: PF-1244 USN
;  CURRENT APPLICATION NUMBER: US/10/491,467
;  CURRENT FILING DATE:  2004-03-31
;  PRIOR APPLICATION NUMBER: PCT/US02/33723
;  PRIOR FILING DATE: 2002-10-17
;  PRIOR APPLICATION NUMBER: US 60/345,474
;  PRIOR FILING DATE: 2001-10-19
;  PRIOR APPLICATION NUMBER: US 60/343,910
;  PRIOR FILING DATE: 2001-11-02
;  PRIOR APPLICATION NUMBER: US 60/333,098
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/332,424
;  PRIOR FILING DATE: 2001-11-16
;  PRIOR APPLICATION NUMBER: US 60/334,288
;  PRIOR FILING DATE: 2001-11-30
;  NUMBER OF SEQ ID NOS: 104
;  SOFTWARE: PERL Program
; SEQ ID NO 70
;   LENGTH: 3713
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 2969494CB1
US-10-491-467-70

Alignment Scores:
Pred. No.:                22            Length:          3713
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
```

```
DB:                 9             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-491-467-70 (1-3713)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||     |||:::::::|||       |||   |||   |||          |||
Db       3373 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 3432

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::       |||   :::|||          |||   |||:::
Db       3433 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3492

Qy         41 ValLeuGlu 43
                :::|||
Db       3493 GAGATGGAG 3501

RESULT 23
US-10-112-944-663/c
; Sequence 663, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wang, Zhiwei
;  TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
;  TITLE OF INVENTION:  Secreted Polypeptides
;  FILE REFERENCE: 805A
;  CURRENT APPLICATION NUMBER: US/10/112,944
;  CURRENT FILING DATE:  2002-03-28
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: US 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  PRIOR APPLICATION NUMBER: US 09/515,126
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: US 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: US 09/552,929
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: US 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 924
;  SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 663
;   LENGTH: 3818
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-112-944-663

Alignment Scores:
Pred. No.:              22.7          Length:        3818
Score:                  68.00         Matches:       14
Percent Similarity:     48.8%         Conservative:  7
Best Local Similarity:  32.6%         Mismatches:    22
Query Match:            29.7%         Indels:        0
DB:                     7             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-112-944-663 (1-3818)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||     |||:::::::|||       |||   |||   |||          |||
Db        316 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 257

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::       |||   :::|||          |||   |||:::
Db        256 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 197

Qy         41 ValLeuGlu 43
                :::|||
Db        196 GAGATGGAG 188

RESULT 24
US-09-190-246-4
; Sequence 4, Application US/09190246
; Publication No. US20030180257A1
; GENERAL INFORMATION:
;  APPLICANT: Parrington, Mark
;  APPLICANT: Li, Xiaomao
;  APPLICANT: Klein, Michel
;  TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
;  FILE REFERENCE: Parrington et al.
;  CURRENT APPLICATION NUMBER: US/09/190,246
;  CURRENT FILING DATE: 1998-11-13
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 8100
;   TYPE: DNA
;   ORGANISM: Semliki Forest virus
US-09-190-246-4

Alignment Scores:
Pred. No.:              73.2          Length:        8100
Score:                  67.00         Matches:       14
Percent Similarity:     51.2%         Conservative:  7
Best Local Similarity:  34.1%         Mismatches:    20
Query Match:            29.3%         Indels:        0
DB:                     3             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-190-246-4 (1-8100)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||       :::::::::   ::::::|||           |||:::
Db       2701 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 2760

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                        |||       |||   |||       |||   ||||||||||:::   |||
Db       2761 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 2820

Qy         41 Val 41
              |||
Db       2821 GTG 2823

RESULT 25
US-09-994-412-3
; Sequence 3, Application US/09994412
; Publication No. US20030059943A1
; GENERAL INFORMATION:
;  APPLICANT: Certa, Ulrich
;  APPLICANT:  Lundstrom, Kenneth
;  TITLE OF INVENTION: Inhibition of Expression of a Target Gene
;  FILE REFERENCE: 20787
;  CURRENT APPLICATION NUMBER: US/09/994,412
;  CURRENT FILING DATE:  2002-03-12
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 10610
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: pSFV2gen(PD) vector
US-09-994-412-3
```

```
Alignment Scores:
Pred. No.:                98.8          Length:          10610
Score:                    67.00         Matches:         14
Percent Similarity:       51.2%         Conservative:    7
Best Local Similarity:    34.1%         Mismatches:      20
Query Match:              29.3%         Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-994-412-3 (1-10610)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::::|||                 |||:::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||  |||      |||  |||||||||:::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101

RESULT 26
US-09-901-106-1
; Sequence 1, Application US/09901106
; Patent No. US20020151067A1
;   GENERAL INFORMATION:
;        APPLICANT: Garoff, Henrik
;                   Liljestrom, Peter
;        TITLE OF INVENTION: DNA Expression Systems Based on
;                            Alphaviruses
;        NUMBER OF SEQUENCES: 27
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Birch, Stewart, Kolasch & Birch
;             STREET: P.O. Box 747
;             CITY: Falls Church
;             STATE: Virginia
;             COUNTRY: USA
;             ZIP: 22040-0747
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/901,106
;             FILING DATE: 10-Jul-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/07/920,281C
;             FILING DATE: 13-AUG-1992
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Murphy Jr., Gerald M.
;             REGISTRATION NUMBER: 28,977
;             REFERENCE/DOCKET NUMBER: 828-103P
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-241-1300
;             TELEFAX: 703-241-2848
;             TELEX: 248345
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 11517 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: RNA (genomic)
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        ORIGINAL SOURCE:
;             ORGANISM: Semliki Forest Virus
;        FEATURE:
;             NAME/KEY:  -
;             LOCATION:  1..11517
;             OTHER INFORMATION:  /label= genome
;             /note= "Semliki Forest Virus complete nucleotide
;             sequence, presented as a cloned DNA sequence; see
;             Figure 5."
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  87..7379
;             OTHER INFORMATION:  /product= "SFV polyprotein"
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  7421..11179
;             OTHER INFORMATION:  /product= "SFV polyprotein"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-901-106-1

Alignment Scores:
Pred. No.:                108           Length:          11517
Score:                    67.00         Matches:         14
Percent Similarity:       51.2%         Conservative:    7
Best Local Similarity:    34.1%         Mismatches:      20
Query Match:              29.3%         Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-901-106-1 (1-11517)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::|||                 |||:::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||  |||      |||  |||||||||:::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101

RESULT 27
US-09-190-246-1
; Sequence 1, Application US/09190246
; Publication No. US20030180257A1
; GENERAL INFORMATION:
;  APPLICANT: Parrington, Mark
;  APPLICANT: Li, Xiaomao
;  APPLICANT: Klein, Michel
;  TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
;  FILE REFERENCE: Parrington et al.
;  CURRENT APPLICATION NUMBER: US/09/190,246
;  CURRENT FILING DATE: 1998-11-13
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 15538
;   TYPE: DNA
;   ORGANISM: Semliki Forest virus
US-09-190-246-1

Alignment Scores:
Pred. No.:                151           Length:          15538
Score:                    67.00         Matches:         14
Percent Similarity:       51.2%         Conservative:    7
Best Local Similarity:    34.1%         Mismatches:      20
Query Match:              29.3%         Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-190-246-1 (1-15538)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::|||                 |||:::
Db       5525 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 5584

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
```

```
                    |||      |||  |||      |||  |||||||||:::   |||
Db       5585 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 5644

Qy         41 Val 41
              |||
Db       5645 GTG 5647
```

RESULT 28
US-09-940-316B-1/c
; Sequence 1, Application US/09940316B
; Publication No. US20030175901A1
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
; TITLE OF INVENTION: GENE CLUSTER
; FILE REFERENCE: 30062-20026.11
; CURRENT APPLICATION NUMBER: US/09/940,316B
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/410,551
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 77536
; TYPE: DNA
; ORGANISM: Streptomyces hygroscopicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52275)...(71465)
US-09-940-316B-1

Alignment Scores:

Pred. No.:	896	Length:	77536
Score:	67.00	Matches:	17
Percent Similarity:	46.9%	Conservative:	6
Best Local Similarity:	34.7%	Mismatches:	16
Query Match:	29.3%	Indels:	10
DB:	3	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-1 (1-77536)

```
Qy          2 AspProGlySerGluIleIleGluSerValProProAla----------GlyPro 16
              :::|||      |||  |||         ||||||                  |||
Db      52933 AACCCCGCCGGCGAGGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCG 52874

Qy         17 GluAlaSerGlu----------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||  |||:::              |||  |||  ::::::|||      :::
Db      52873 GAGCGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCG 52814

Qy         32 PheValSerGluGlyProSerArgPro 40
                 |||:::  |||||||||||||||
Db      52813 ACCGTGACCGCCGGACCCTCCAGACCG 52787
```

RESULT 29
US-10-320-797-2316
; Sequence 2316, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2316
; LENGTH: 2436
; TYPE: DNA
; ORGANISM: Cryptococcus neoformans
US-10-320-797-2316

Alignment Scores:

Pred. No.:	38	Length:	2436
Score:	65.00	Matches:	16
Percent Similarity:	57.1%	Conservative:	8
Best Local Similarity:	38.1%	Mismatches:	12
Query Match:	28.4%	Indels:	6
DB:	7	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-320-797-2316 (1-2436)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   |||:::   ::::::              :::|||   |||   |||
Db        799 GGTGATGAAGGTACTTCCGTAGTT--------------AGCGGCTCAGAAGAAAGC--- 840

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||   ||||||||||||||||::::::   :::          |||   |||   :::|||
Db        841 AGCAACACGGATGAAAATGAAGAGGAGGAGGAGGAGGAAGAAGAACCTGTCAAGCCT 900

Qy         41 ValLeu 42
                 |||
Db        901 ACCCTC 906
```

RESULT 30
US-10-320-797-1316
; Sequence 1316, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1316
; LENGTH: 3346
; TYPE: DNA
; ORGANISM: Cryptococcus neoformans
US-10-320-797-1316

Alignment Scores:

Pred. No.:	54.1	Length:	3346
Score:	65.00	Matches:	16
Percent Similarity:	57.1%	Conservative:	8
Best Local Similarity:	38.1%	Mismatches:	12
Query Match:	28.4%	Indels:	6
DB:	7	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-320-797-1316 (1-3346)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   |||:::   ::::::              :::|||   |||   |||
Db       1357 GGTGATGAAGGTACTTCCGTAGTT--------------AGCGGCTCAGAAGAAAGC--- 1398

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||   ||||||||||||||||::::::   :::          |||   |||   :::|||
```

```
Db       1399 AGCAACACGGATGAAAATGAAGAGGAGGAGGAGGAGGAGGAAGAAGAACCTGTCAAGCCT 1458

Qy         41 ValLeu 42
                |||
Db       1459 ACCCTC 1464
```

RESULT 31
US-10-320-797-316
; Sequence 316, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 316
; LENGTH: 5346
; TYPE: DNA
; ORGANISM: Cryptococcus neoformans
US-10-320-797-316

Alignment Scores:

Pred. No.:	90.9	Length:	5346
Score:	65.00	Matches:	16
Percent Similarity:	57.1%	Conservative:	8
Best Local Similarity:	38.1%	Mismatches:	12
Query Match:	28.4%	Indels:	6
DB:	7	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-320-797-316 (1-5346)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   ||:::   ::::::           :::|||   |||   |||
Db       2357 GGTGATGAAGGTACTTCCGTAGTT--------------AGCGGCTCAGAAGAAAGC--- 2398

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||   ||||||||||||||||:::::   :::         |||   |||   :::|||
Db       2399 AGCAACACGGATGAAAATGAAGAGGAGGAGGAGGAGGAGGAAGAAGAACCTGTCAAGCCT 2458

Qy         41 ValLeu 42
                |||
Db       2459 ACCCTC 2464
```

RESULT 32
US-09-940-316B-1
; Sequence 1, Application US/09940316B
; Publication No. US20030175901A1
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHU, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI
; TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
; TITLE OF INVENTION: GENE CLUSTER
; FILE REFERENCE: 30062-20026.11
; CURRENT APPLICATION NUMBER: US/09/940,316B
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/410,551
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/139,650
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: US 60/123,810
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/102,748
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 77536
; TYPE: DNA
; ORGANISM: Streptomyces hygroscopicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52275)...(71465)
US-09-940-316B-1

Alignment Scores:

Pred. No.:	1.76e+03	Length:	77536
Score:	65.00	Matches:	17
Percent Similarity:	45.8%	Conservative:	5
Best Local Similarity:	35.4%	Mismatches:	16
Query Match:	28.4%	Indels:	10
DB:	3	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-1 (1-77536)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------GlyProGlu 17
              |||   |||   |||   |||||                              ||||||
Db      28077 CCGGACGGCGAGGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 28136

Qy         18 AlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                 |||:::                |||   |||   ::::::|||       :::
Db      28137 CGCAGTGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 28196

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   ||||||||||||||||
Db      28197 GTGACCGCCGGACCCTCCAGACCG 28220
```

RESULT 33
US-10-765-700-95
; Sequence 95, Application US/10765700
; Publication No. US20050130171A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/10/765,700
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: US/09/566,921
; PRIOR FILING DATE: PRIORT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 95
; LENGTH: 7390
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 337896.11
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 5097-5416
; OTHER INFORMATION: a, t, c, g, or other
US-10-765-700-95

Alignment Scores:

Pred. No.:	183	Length:	7390
Score:	64.00	Matches:	17
Percent Similarity:	45.7%	Conservative:	4
Best Local Similarity:	37.0%	Mismatches:	17
Query Match:	27.9%	Indels:	8
DB:	9	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-765-700-95 (1-7390)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        332 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 391

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||      ||||||||||||   |||   |||:::   |||           :::
Db        392 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 451

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        452 GAGTCGCCCTCGCACCCG 469
```

RESULT 34
US-10-363-616-164
; Sequence 164, Application US/10363616
; Publication No. US20040044181A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-113 (793)
; CURRENT APPLICATION NUMBER: US/10/363,616
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 09/654,935
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 490
; SEQ ID NO 164
; LENGTH: 7666
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (632)..(1552)
US-10-363-616-164

Alignment Scores:

Pred. No.:	190	Length:	7666
Score:	64.00	Matches:	17
Percent Similarity:	45.7%	Conservative:	4
Best Local Similarity:	37.0%	Mismatches:	17
Query Match:	27.9%	Indels:	8
DB:	7	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-363-616-164 (1-7666)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||      ||||||||||||   |||   |||:::   |||           :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        806 GAGTCGCCCTCGCACCCG 823
```

RESULT 35
US-10-171-581-295
; Sequence 295, Application US/10171581
; Publication No. US20030104426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 295
; LENGTH: 7758
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB011085
; DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-295

Alignment Scores:

Pred. No.:	193	Length:	7758
Score:	64.00	Matches:	17
Percent Similarity:	45.7%	Conservative:	4
Best Local Similarity:	37.0%	Mismatches:	17
Query Match:	27.9%	Indels:	8
DB:	5	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-171-581-295 (1-7758)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||      ||||||||||||   |||   |||:::   |||           :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        806 GAGTCGCCCTCGCACCCG 823
```

RESULT 36
US-10-723-860-1239
; Sequence 1239, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1239
; LENGTH: 7758
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-1239

Alignment Scores:

Pred. No.:	193	Length:	7758
Score:	64.00	Matches:	17
Percent Similarity:	45.7%	Conservative:	4
Best Local Similarity:	37.0%	Mismatches:	17
Query Match:	27.9%	Indels:	8
DB:	8	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-723-860-1239 (1-7758)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||      ||||||||||||   |||   |||:::   |||           :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        806 GAGTCGCCCTCGCACCCG 823
```

```
RESULT 37
US-10-723-860-5743
; Sequence 5743, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5743
;   LENGTH: 7792
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (511)..(536)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-5743

Alignment Scores:
Pred. No.:                194           Length:        7792
Score:                    64.00         Matches:       17
Percent Similarity:       45.7%         Conservative:  4
Best Local Similarity:    37.0%         Mismatches:    17
Query Match:              27.9%         Indels:        8
DB:                       8             Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-723-860-5743 (1-7792)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||     ||||||||||||   |||   |||:::  |||            :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||  |||
Db        806 GAGTCGCCCTCGCACCCG 823

RESULT 38
US-10-612-869-8
; Sequence 8, Application US/10612869
; Publication No. US20040214161A1
; GENERAL INFORMATION:
;  APPLICANT: HEALTH PROTECTION AGENCY
;  TITLE OF INVENTION: DETECTION OF EPSTEI N BARR VIRUS
;  FILE REFERENCE: MRM/PJG/25431
;  CURRENT APPLICATION NUMBER: US/10/612,869
;  CURRENT FILING DATE:  2003-07-02
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
;   LENGTH: 171096
;   TYPE: DNA
;   ORGANISM: Cercopithecine herpesvirus 15 strain LCL8664
US-10-612-869-8

Alignment Scores:
Pred. No.:                5.95e+03      Length:        171096
Score:                    64.00         Matches:       19
Percent Similarity:       53.1%         Conservative:  7
Best Local Similarity:    38.8%         Mismatches:    14
Query Match:              27.9%         Indels:        9
DB:                       8             Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-612-869-8 (1-171096)

Qy          1 GlyAspProGlySerGluIleIle------GluSerValProProAlaGlyProGluAla 18
              ||||||||||||||||              |||:::|||   |||            |||
Db      76730 GGTGATCCCGGGTCTAATCCCAGTCCACGTGAAACTGTGTCCCCGGGACCGTCTACTGCC 76789

Qy         19 SerGluSer-----------------ThrThrAspGluAsnGluAspAspIleGlnPhe 32
              ||||||                  :::   ::::::||||||::::::
Db      76790 AGCGAGCAAATGTTGCAACACGAGCCCATGGAAGAGAATTCAGATGATGACGTGGAAGAA 76849

Qy         33 ValSerGlu-GlyProSerArgPro 40
              |||||||:: ||||||   ||||||
Db      76850 GTGAGCCAACGGGCCAGAACGCCCC 76874

RESULT 39
US-10-424-599-48990
; Sequence 48990, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 48990
;   LENGTH: 428
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_15244C.1
US-10-424-599-48990

Alignment Scores:
Pred. No.:                9.19          Length:        428
Score:                    63.50         Matches:       14
Percent Similarity:       48.8%         Conservative:  7
Best Local Similarity:    32.6%         Mismatches:    17
Query Match:              27.7%         Indels:        5
DB:                       7             Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-424-599-48990 (1-428)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||::::::              |||   |||   |||   |||:::
Db         25 GGGGACCCGGGGGCGGAC--------------CCGGGGGCGGCCCCCGGGGCCACCCCG 69

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||   |||   :::     ||||||    :::|||    :::
Db         70 AAAAACCCGGACCGGAACCGGGAGGCAAACCAATTTCCCGAAAAAGGGGGGGAAAAAAAG 129

Qy         41 ValLeuGlu 43
              :::   |||
Db        130 ATTGCCGAA 138

RESULT 40
US-10-437-963-59361
; Sequence 59361, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
```

```
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 59361
;   LENGTH: 270
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(270)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_60993C.1
US-10-437-963-59361

Alignment Scores:
Pred. No.:              6.53          Length:         270
Score:                  63.00         Matches:        17
Percent Similarity:     53.8%         Conservative:   4
Best Local Similarity:  43.6%         Mismatches:     17
Query Match:            27.5%         Indels:         2
DB:                     7             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-437-963-59361 (1-270)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||   |||   :::         :::||||||      |||   |||   |||:::|||
Db         13 CCCCCATCCTCCCTATTCCCTCCCCTTCCTCCGCCACCCCCTCCCGCTCCTGAGGCGACG 72

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProVal 41
              |||           ||||||   |||        |||||||||   ||||||:::
Db         73 ACCACCTTCGAGAGCGACGACGA-CAA---CGACGAGAAGGTCCCCGAAGACCCATC 125

RESULT 41
US-10-437-963-82114/c
; Sequence 82114, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 82114
;   LENGTH: 2736
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_81576C.1
US-10-437-963-82114

Alignment Scores:
Pred. No.:              85.1          Length:         2736
Score:                  63.00         Matches:        11
Percent Similarity:     66.7%         Conservative:   5
Best Local Similarity:  45.8%         Mismatches:     8
Query Match:            27.5%         Indels:         0
DB:                     7             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-437-963-82114 (1-2736)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||   |||       :::   |||:::||||||:::   |||   :::||||||
Db       2720 GGGTCGCCGGCCTCGTCGTCGCTGTCGTCCATACCACCGTCGCCGCCCTCCTCGTCGGAG 2661

Qy         21 SerThrThrAsp 24
              |||   |||:::
Db       2660 TCGAGGACGAAC 2649

RESULT 42
US-10-112-944-191
; Sequence 191, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wang, Zhiwei
;  TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
;  TITLE OF INVENTION:  Secreted Polypeptides
;  FILE REFERENCE: 805A
;  CURRENT APPLICATION NUMBER: US/10/112,944
;  CURRENT FILING DATE:  2002-03-28
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: US 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  PRIOR APPLICATION NUMBER: US 09/515,126
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: US 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: US 09/552,929
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: US 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 924
;  SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 191
;   LENGTH: 3120
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (115)..(3063)
US-10-112-944-191

Alignment Scores:
Pred. No.:              98.5          Length:         3120
Score:                  63.00         Matches:        13
Percent Similarity:     46.5%         Conservative:   7
Best Local Similarity:  30.2%         Mismatches:     23
Query Match:            27.5%         Indels:         0
DB:                     7             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-112-944-191 (1-3120)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||:::::::|||       |||          |||          |||
Db       2812 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTCCTCCCCCAACCCTGTTGGAA 2871
```

```
Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               |||::::::  |||  ::|||      |||   |||:::
Db        2872 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 2931

Qy          41 ValLeuGlu 43
                :::|||
Db        2932 GAGATGGAG 2940
```

RESULT 43
US-10-937-379-44/c
; Sequence 44, Application US/10937379
; Publication No. US20050176653A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: McAlpine, James B.
; APPLICANT: Farnet, Chris M.
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Dan Sorensen
; TITLE OF INVENTION: POLYENE POLYKETIDES AND METHODS OF PRODUCTION
; FILE REFERENCE: 3010-4US
; CURRENT APPLICATION NUMBER: US/10/937,379
; CURRENT FILING DATE: 2004-09-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 9795
; TYPE: DNA
; ORGANISM: Amycolatopsis orientalis
US-10-937-379-44

Alignment Scores:

Pred. No.:	350	Length:	9795
Score:	63.00	Matches:	17
Percent Similarity:	46.7%	Conservative:	4
Best Local Similarity:	37.8%	Mismatches:	14
Query Match:	27.5%	Indels:	10
DB:	9	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-937-379-44 (1-9795)

```
Qy           6 GluIleIleGluSerValProProAla--------------GlyProGluAlaSerGlu 20
               |||  |||       |||||||||                 |||      |||:::
Db         722 GAGGCCATCACGGTCACCCCGCCCGCCAGCGCGAGCGTGGACTCCCCGCGCCGGAGCGAC 663

Qy          21 --------------SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                           |||  |||  |||:::|||     :::      |||:::
Db         662 TCGGCAGCGAGGTGTAGCGCGACCAGCGAAGACGAACACGCGGTGTCGATGGTGACGGCC 603

Qy          36 GlyProSerArgPro 40
               |||||||||||||||
Db         602 GGGCCTTCGAGGCCG 588
```

RESULT 44
US-10-937-379-24/c
; Sequence 24, Application US/10937379
; Publication No. US20050176653A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: McAlpine, James B.
; APPLICANT: Farnet, Chris M.
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Dan Sorensen
; TITLE OF INVENTION: POLYENE POLYKETIDES AND METHODS OF PRODUCTION
; FILE REFERENCE: 3010-4US
; CURRENT APPLICATION NUMBER: US/10/937,379
; CURRENT FILING DATE: 2004-09-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 73599
; TYPE: DNA
; ORGANISM: Amycolatopsis orientalis
US-10-937-379-24

Alignment Scores:

Pred. No.:	3.28e+03	Length:	73599
Score:	63.00	Matches:	17
Percent Similarity:	46.7%	Conservative:	4
Best Local Similarity:	37.8%	Mismatches:	14
Query Match:	27.5%	Indels:	10
DB:	9	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-937-379-24 (1-73599)

```
Qy           6 GluIleIleGluSerValProProAla--------------GlyProGluAlaSerGlu 20
               |||  |||       |||||||||                 |||      |||:::
Db       43338 GAGGCCATCACGGTCACCCCGCCCGCCAGCGCGAGCGTGGACTCCCCGCGCCGGAGCGAC 43279

Qy          21 --------------SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                           |||  |||  |||:::|||     :::      |||:::
Db       43278 TCGGCAGCGAGGTGTAGCGCGACCAGCGAAGACGAACACGCGGTGTCGATGGTGACGGCC 43219

Qy          36 GlyProSerArgPro 40
               |||||||||||||||
Db       43218 GGGCCTTCGAGGCCG 43204
```

RESULT 45
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Alignment Scores:

Pred. No.:	6.66e+05	Length:	9025608
Score:	63.00	Matches:	13
Percent Similarity:	60.7%	Conservative:	4
Best Local Similarity:	46.4%	Mismatches:	11
Query Match:	27.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-156-761-1 (1-9025608)

```
Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||   |||   |||   :::|||:::     |||    ::::||||||:::   |||
Db      966353 CCACCAGAGGGGGCTGAGGAAGCCGAGGCCGAACAAGACCACGATGAGGATGAGAAGCAG 966412

Qy          32 PheValSerGluGlyProSerArg 39
               |||||||||||||
Db      966413 AACGAACACGGCGGGCCTTCCCGG 966436
```

RESULT 46

```
US-09-867-701-5034/c
; Sequence 5034, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5034
;   LENGTH: 557
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(557)
;   OTHER INFORMATION: n = A,T,C or G
US-09-867-701-5034

Alignment Scores:
Pred. No.:              20.4          Length:        557
Score:                  62.00         Matches:       17
Percent Similarity:     50.0%         Conservative:  6
Best Local Similarity:  37.0%         Mismatches:    11
Query Match:            27.1%         Indels:        12
DB:                     3             Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-867-701-5034 (1-557)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn------ 26
              ::::::::::||||||||||||||||||||||||||                ||||||
Db        180 AAAGCCATTCCACCGGCCGGCCCCGAAGCCTCGCTTTTTCATGCCCAGGAGAACTCGCTG 121

Qy         27 ------GluAspAspIleGlnPheValSerGluGly---------------------- 36
                  :::|||       |||   :::       |||
Db        120 AGGCTGCAGGACCAGCACCAGGGTATAGAGCCTGGAGGCGCCCATGGTCCCTGGGTTTCT 61

Qy         37 ProSerArgProValLeu 42
              |||||||||   |||:::
Db         60 CCAAGCCGCCGTGTGATC 43

RESULT 47
US-10-767-701-2245
; Sequence 2245, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof For Plant Improvement
;  FILE REFERENCE: 38-21(53535)B
;  CURRENT APPLICATION NUMBER: US/10/767,701
;  CURRENT FILING DATE:  2004-01-29
;  NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 2245
;   LENGTH: 651
;   TYPE: DNA
;   ORGANISM: Sorghum bicolor
;   FEATURE:
;   OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS37640_1
US-10-767-701-2245

Alignment Scores:
Pred. No.:              24.3          Length:        651
Score:                  62.00         Matches:       17
Percent Similarity:     44.7%         Conservative:  4
Best Local Similarity:  36.2%         Mismatches:    16
Query Match:            27.1%         Indels:        10
DB:                     7             Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-767-701-2245 (1-651)

Qy          6 GluIleIleGluSerValProProAlaGlyProGlu---------------------- 17
              :::   |||       |||||||||:::       |||
Db        176 CAGTTCATCCTCCGCGTGCCCCCATCCGTGGCGGAGCGGATCGAGCGCCTCATGAATGAA 235

Qy         18 ------AlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                    |||:::   |||       ||||||          |||:::   |||       |||
Db        236 TCCGCCGCTGCCTCCTCCAACCCCGACGAGGCCTCCCTCGACCTCTCCTTCTCAGGGGAG 295

Qy         36 GlyProSerArgProValLeu 42
              ||||||   ||||||   |||
Db        296 GGTCCCCATCGCCCCCGTCTC 316

RESULT 48
US-10-956-157-5003/c
; Sequence 5003, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WIT
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5003
;   LENGTH: 712
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-5003

Alignment Scores:
Pred. No.:              26.8          Length:        712
Score:                  62.00         Matches:       17
Percent Similarity:     50.0%         Conservative:  6
Best Local Similarity:  37.0%         Mismatches:    11
Query Match:            27.1%         Indels:        12
DB:                     9             Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-956-157-5003 (1-712)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn------ 26
              ::::::::::||||||||||||||||||||||||||                ||||||
Db        202 AAAGCCATTCCACCGGCCGGCCCCGAAGCCTCGCTTTTTCATGCCCAGGAGAACTCGCTG 143

Qy         27 ------GluAspAspIleGlnPheValSerGluGly---------------------- 36
                  :::|||       |||   :::       |||
Db        142 AGGCTGCAGGACCAGCACCAGGGTATAGAGCCTGGAGGCGCCCATGGTCCCTGGGTTTCT 83

Qy         37 ProSerArgProValLeu 42
              |||||||||   |||:::
Db         82 CCAAGCCGCCGTGTGATC 65

RESULT 49
US-10-956-157-10238/c
; Sequence 10238, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 10238
;   LENGTH: 712
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-10238

Alignment Scores:
Pred. No.:              26.8          Length:         712
Score:                  62.00         Matches:        17
Percent Similarity:     50.0%         Conservative:   6
Best Local Similarity:  37.0%         Mismatches:     11
Query Match:            27.1%         Indels:         12
DB:                     9             Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-956-157-10238 (1-712)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn------ 26
              :::::::::|||||||||||||||||||||||||||||              ||||||
Db        202 AAAGCCATTCCACCGGCCGGCCCCGAAGCCTCGCTTTTTCATGCCCAGGAGAACTCGCTG 143

Qy         27 ------GluAspAspIleGlnPheValSerGluGly------------------------ 36
                  :::|||      |||   :::     |||
Db        142 AGGCTGCAGGACCAGCACCAGGGTATAGAGCCTGGAGGCGCCCATGGTCCCTGGGTTTCT 83

Qy         37 ProSerArgProValLeu 42
              |||||||||   |||:::
Db         82 CCAAGCCGCCGTGTGATC 65

RESULT 50
US-10-927-644-4
; Sequence 4, Application US/10927644
; Publication No. US20050095241A1
; GENERAL INFORMATION:
;  APPLICANT: Vadlamudi, Ratna K.
;  APPLICANT:  Kumar, Rakesh
;  APPLICANT:  Balasenthil, Seetharaman
;  TITLE OF INVENTION: ESTROGEN RECEPTOR MODULATORS AND USES THEREOF
;  FILE REFERENCE: AND541/4-011US
;  CURRENT APPLICATION NUMBER: US/10/927,644
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/498,118
;  PRIOR FILING DATE: 2003-08-26
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
;   LENGTH: 762
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-927-644-4

Alignment Scores:
Pred. No.:              28.9          Length:         762
Score:                  62.00         Matches:        13
Percent Similarity:     46.5%         Conservative:   7
Best Local Similarity:  30.2%         Mismatches:     23
Query Match:            27.1%         Indels:         0
DB:                     9             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-927-644-4 (1-762)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||::::::|||     |||   |||   |||     |||
Db        508 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 567

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  ::::::      |||   :::|||          |||   |||:::
Db        568 GAGGAGCCTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 627

Qy         41 ValLeuGlu 43
                 :::|||
Db        628 GAGATGGAG 636

RESULT 51
```

```
US-09-778-844-110
; Sequence 110, Application US/09778844
; Patent No. US20020150971A1
; GENERAL INFORMATION:
;  APPLICANT: JOHANSEN, JEANETTE ELISABETH
;  APPLICANT:  SCHALLING, MARTIN
;  TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES FOR CONTROLLING FOOD
;  TITLE OF INVENTION:  INTAKE AND/OR BODY WEIGHT
;  FILE REFERENCE: 030307/0195
;  CURRENT APPLICATION NUMBER: US/09/778,844
;  CURRENT FILING DATE:  2001-02-08
;  NUMBER OF SEQ ID NOS: 206
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 110
;   LENGTH: 1056
;   TYPE: DNA
;   ORGANISM: Rattus sp.
US-09-778-844-110

Alignment Scores:
Pred. No.:              41.6          Length:         1056
Score:                  62.00         Matches:        13
Percent Similarity:     56.2%         Conservative:   5
Best Local Similarity:  40.6%         Mismatches:     14
Query Match:            27.1%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-778-844-110 (1-1056)

Qy         11 ValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIle 30
              |||||||||:::||||||      |||      :::      |||   ||||||:::
Db        191 GTACCCCCTTCTGGACCAGCTCTTTCCTCAGACAGTGACAAGGAGGGAGAAGATGAGGGT 250

Qy         31 GlnPheValSerGluGlyProSerArgProValLeu 42
                      |||   |||:::   |||:::|||
Db        251 ACTGAAGAAGAAGAACTGCCAGCTCTGCCCATCCTT 286

RESULT 52
US-09-778-844-108
; Sequence 108, Application US/09778844
; Patent No. US20020150971A1
; GENERAL INFORMATION:
;  APPLICANT: JOHANSEN, JEANETTE ELISABETH
;  APPLICANT:  SCHALLING, MARTIN
;  TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES FOR CONTROLLING FOOD
;  TITLE OF INVENTION:  INTAKE AND/OR BODY WEIGHT
;  FILE REFERENCE: 030307/0195
;  CURRENT APPLICATION NUMBER: US/09/778,844
;  CURRENT FILING DATE:  2001-02-08
;  NUMBER OF SEQ ID NOS: 206
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 108
;   LENGTH: 1135
;   TYPE: DNA
;   ORGANISM: Rattus sp.
US-09-778-844-108

Alignment Scores:
Pred. No.:              45            Length:         1135
Score:                  62.00         Matches:        13
Percent Similarity:     56.2%         Conservative:   5
Best Local Similarity:  40.6%         Mismatches:     14
Query Match:            27.1%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-778-844-108 (1-1135)

Qy         11 ValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIle 30
              |||||||||:::||||||      |||      :::      |||   ||||||:::
Db        843 GTACCCCCTTCTGGACCAGCTCTTTCCTCAGACAGTGACAAGGAGGGAGAAGATGAGGGT 902

Qy         31 GlnPheValSerGluGlyProSerArgProValLeu 42
                      |||   |||:::   |||:::|||
```

```
Db        903 ACTGAAGAAGAAGAACTGCCAGCTCTGCCCATCCTT 938

RESULT 53
US-09-778-844-109
; Sequence 109, Application US/09778844
; Patent No. US20020150971A1
; GENERAL INFORMATION:
;  APPLICANT: JOHANSEN, JEANETTE ELISABETH
;  APPLICANT:  SCHALLING, MARTIN
;  TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES FOR CONTROLLING FOOD
;  TITLE OF INVENTION:  INTAKE AND/OR BODY WEIGHT
;  FILE REFERENCE: 030307/0195
;  CURRENT APPLICATION NUMBER: US/09/778,844
;  CURRENT FILING DATE:  2001-02-08
;  NUMBER OF SEQ ID NOS: 206
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
;   LENGTH: 1137
;   TYPE: DNA
;   ORGANISM: Rattus sp.
US-09-778-844-109

Alignment Scores:
Pred. No.:              45.1          Length:         1137
Score:                  62.00         Matches:        13
Percent Similarity:     56.2%         Conservative:   5
Best Local Similarity:  40.6%         Mismatches:     14
Query Match:            27.1%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-778-844-109 (1-1137)

Qy         11 ValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIle 30
              ||||||||||:::||||||      |||      :::      |||   ||||||:::
Db        517 GTACCCCCTTCTGGACCAGCTCTTTCCTCAGACAGTGACAAGGAGGGAGAAGATGAGGGT 576

Qy         31 GlnPheValSerGluGlyProSerArgProValLeu 42
                        |||   |||:::   |||:::|||
Db        577 ACTGAAGAAGAAGAACTGCCAGCTCTGCCCATCCTT 612

RESULT 54
US-10-927-644-2
; Sequence 2, Application US/10927644
; Publication No. US20050095241A1
; GENERAL INFORMATION:
;  APPLICANT: Vadlamudi, Ratna K.
;  APPLICANT:  Kumar, Rakesh
;  APPLICANT:  Balasenthil, Seetharaman
;  TITLE OF INVENTION: ESTROGEN RECEPTOR MODULATORS AND USES THEREOF
;  FILE REFERENCE: AND541/4-011US
;  CURRENT APPLICATION NUMBER: US/10/927,644
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/498,118
;  PRIOR FILING DATE: 2003-08-26
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
;   LENGTH: 3393
;   TYPE: DNA
;   ORGANISM: Homosapiens
US-10-927-644-2

Alignment Scores:
Pred. No.:              152           Length:         3393
Score:                  62.00         Matches:        13
Percent Similarity:     46.5%         Conservative:   7
Best Local Similarity:  30.2%         Mismatches:     23
Query Match:            27.1%         Indels:         0
DB:                     9             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-927-644-2 (1-3393)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
                 |||   |||      |||:::::::|||      |||   |||   |||         |||
Db       3139 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 3198

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                       ::::::      |||   :::|||         |||   |||:::
Db       3199 GAGGAGCCTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3258

Qy         41 ValLeuGlu 43
                 :::|||
Db       3259 GAGATGGAG 3267

RESULT 55
US-11-067-321A-3
; Sequence 3, Application US/11067321A
; Publication No. US20050202515A1
; GENERAL INFORMATION:
;  APPLICANT: Barletta, Frank
;  APPLICANT:  O'Donnell, Greg
;  APPLICANT:  Cheskis, Boris
;  TITLE OF INVENTION: FUNCTIONAL SCREENING METHOD FOR IDENTIFYING COMPOUNDS FOR THE
;  TITLE OF INVENTION:  TREATMENT OF OSTEOPOROSIS USING THE VITAMIN D RECEPTOR
;  FILE REFERENCE: 00630/1200916-US1
;  CURRENT APPLICATION NUMBER: US/11/067,321A
;  CURRENT FILING DATE:  2005-02-25
;  PRIOR APPLICATION NUMBER: 60/548,352
;  PRIOR FILING DATE: 2004-02-27
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
;   LENGTH: 3393
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-067-321A-3

Alignment Scores:
Pred. No.:              152           Length:         3393
Score:                  62.00         Matches:        13
Percent Similarity:     46.5%         Conservative:   7
Best Local Similarity:  30.2%         Mismatches:     23
Query Match:            27.1%         Indels:         0
DB:                     10            Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-067-321A-3 (1-3393)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
                 |||   |||      |||:::::::|||      |||   |||   |||         |||
Db       3139 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 3198

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                       ::::::      |||   :::|||         |||   |||:::
Db       3199 GAGGAGCCTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3258

Qy         41 ValLeuGlu 43
                 :::|||
Db       3259 GAGATGGAG 3267

RESULT 56
US-10-116-788A-32/c
; Sequence 32, Application US/10116788A
; Publication No. US20030170840A1
; GENERAL INFORMATION:
;  APPLICANT: McDonald, Thomas L.
;  APPLICANT:  Larson, Marilyn A.
;  APPLICANT:  Weber, Annika
;  TITLE OF INVENTION: Genomic Mammary Amyloid A Sequence
;  FILE REFERENCE: P04557US1
;  CURRENT APPLICATION NUMBER: US/10/116,788A
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: 09/425,679
;  PRIOR FILING DATE: 1999-10-22
;  PRIOR APPLICATION NUMBER: 60/218,482
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/218,611
```

```
;  PRIOR FILING DATE: 2000-07-17
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
;   LENGTH: 4984
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3191, 3222, 3229, 3323, 3402
;   OTHER INFORMATION: n can be a, t, c, or g
US-10-116-788A-32

Alignment Scores:
Pred. No.:              232          Length:          4984
Score:                  62.00        Matches:         11
Percent Similarity:     50.0%        Conservative:    3
Best Local Similarity:  39.3%        Mismatches:      14
Query Match:            27.1%        Indels:          0
DB:                     6            Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-116-788A-32 (1-4984)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||           :::||||||:::   |||       |||
Db       4169 GGAGATCCAGGTGGCAGTTTCTGGAGCAGGATTCCACCATCCAGTCCTCCTGTGTCCATG 4110

Qy         21 SerThrThrAspGluAsnGluAsp 28
                    |||:::      ||||||
Db       4109 CCCCACACAGAGAGCACAGAGGAC 4086

RESULT 57
US-10-494-484A-1/c
; Sequence 1, Application US/10494484A
; Publication No. US20050107315A1
; GENERAL INFORMATION:
;  APPLICANT: Board of Regents, University of Nebraska
;  APPLICANT:  McDonald, Thomas
;  APPLICANT:  Larson, Marilynn
;  APPLICANT:  Weber, Annika
;  TITLE OF INVENTION: Mammary-Associated Serum Amyloid A3 Promoter Sequences and Uses f
;  FILE REFERENCE: U051-167
;  CURRENT APPLICATION NUMBER: US/10/494,484A
;  CURRENT FILING DATE:  2004-04-29
;  PRIOR APPLICATION NUMBER: 60/333,261
;  PRIOR FILING DATE: 2001-11-21
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 7518
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(7518)
;   OTHER INFORMATION: "n" can be any nucleotide
US-10-494-484A-1

Alignment Scores:
Pred. No.:              366          Length:          7518
Score:                  62.00        Matches:         11
Percent Similarity:     50.0%        Conservative:    3
Best Local Similarity:  39.3%        Mismatches:      14
Query Match:            27.1%        Indels:          0
DB:                     9            Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-494-484A-1 (1-7518)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||           :::||||||:::   |||       |||
Db       6703 GGAGATCCAGGTGGCAGTTTCTGGAGCAGGATTCCACCATCCAGTCCTCCTGTGTCCATG 6644

Qy         21 SerThrThrAspGluAsnGluAsp 28
                    |||:::      ||||||
Db       6643 CCCCACACAGAGAGCACAGAGGAC 6620

RESULT 58
US-10-767-701-3659
; Sequence 3659, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof For Plant Improvement
;  FILE REFERENCE: 38-21(53535)B
;  CURRENT APPLICATION NUMBER: US/10/767,701
;  CURRENT FILING DATE:  2004-01-29
;  NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 3659
;   LENGTH: 494
;   TYPE: DNA
;   ORGANISM: Sorghum bicolor
;   FEATURE:
;   OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS62000_1
US-10-767-701-3659

Alignment Scores:
Pred. No.:              21.2         Length:          494
Score:                  61.50        Matches:         17
Percent Similarity:     39.1%        Conservative:    8
Best Local Similarity:  26.6%        Mismatches:      10
Query Match:            26.9%        Indels:          29
DB:                     7            Gaps:            3

US-10-757-745-6_COPY_2_45 (1-44) x US-10-767-701-3659 (1-494)

Qy          1 GlyAspProGlySerGluIleIleGluSer---------------------------- 10
              ||||||||||||           ::::::
Db        115 GGAGATCCGGGGCTCCATGTCGTCGCCGAGCAGGATGTTGGTCGGCACCAGACGGCCGTG 174

Qy         11 ------------ValProProAlaGlyPro--------------------------Glu 17
                          ||||||||||||||||                           :::
Db        175 CACGATGGGCCGCCGCCGCCGGCCGGACCCGGCGGCCGCCCCCGTCCACCCTTGGTGCAG 234

Qy         18 AlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyPro 37
              |||:::::::::|||     :::|||  :::                    |||
Db        235 GAACGCGAGACCTCGAGCGACGCCGAGGATGATGCGGTT-----------------CCG 276

Qy         38 SerArgProVal 41
              ||||||||||||
Db        277 AGTCGGCCAGTC 288

RESULT 59
US-10-419-341-4
; Sequence 4, Application US/10419341
; Publication No. US20030180785A1
; GENERAL INFORMATION:
;  APPLICANT: Krishnan, Rajendra
;  APPLICANT:  Coleman, Rebecca A.
;  APPLICANT:  Yoder, Christine C.
;  APPLICANT:  Durtschi, Becky A.
;  APPLICANT:  Brake, David
;  TITLE OF INVENTION: POLYNUCLEOTIDE MOLECULES ENCODING NEOSPORA PROTEINS
;  FILE REFERENCE: PC9943A
;  CURRENT APPLICATION NUMBER: US/10/419,341
;  CURRENT FILING DATE:  2003-04-21
;  PRIOR APPLICATION NUMBER: US/09/276,438
;  PRIOR FILING DATE: 1999-03-25
;  PRIOR APPLICATION NUMBER: 60/079,389
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: 60/112,282
;  PRIOR FILING DATE: 1998-12-15
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 2.0 - beta
```

```
; SEQ ID NO 4
;   LENGTH: 1031
;   TYPE: DNA
;   ORGANISM: Neospora caninum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (25)..(660)
US-10-419-341-4

Alignment Scores:
Pred. No.:                47.9          Length:         1031
Score:                    61.50         Matches:        13
Percent Similarity:       61.1%         Conservative:   9
Best Local Similarity:    36.1%         Mismatches:     13
Query Match:              26.9%         Indels:         1
DB:                       6             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-419-341-4 (1-1031)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGlu---SerThrThrAspGluAsnGlu 27
              |||   |||||||||:::||||||   :::   |||         |||   :::   :::
Db        493 GAAAAAGTGCCCCCCAGTGGCCCGCGATCAGGTGAAGTTCAGCGTACTCGTAAAGAGCAA 552

Qy         28 AspAspIleGlnPheValSerGluGlyProSerArgProValLeuGlu 43
              :::|||:::|||      :::|||      :::      :::||||||
Db        553 AATGACGTGCAGCAAACCGCAGAGATGTTGGCTGAGGAAATTCTTGAG 600

RESULT 60
US-10-425-115-35966
; Sequence 35966, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 35966
;   LENGTH: 380
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(380)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_132800C.1
US-10-425-115-35966

Alignment Scores:
Pred. No.:                18.8          Length:         380
Score:                    61.00         Matches:        15
Percent Similarity:       47.8%         Conservative:   7
Best Local Similarity:    32.6%         Mismatches:     18
Query Match:              26.6%         Indels:         6
DB:                       8             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-425-115-35966 (1-380)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAla---------- 18
              |||   :::|||   ::::::      ||||||   |||      |||
Db        183 CCGGAGGCTGAGACGGTCGATCCGCCGCCCCCGGGGGGGGGGCCAGCGCCCAAGACGGTC 242

Qy         19 ------SerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGly 36
                 ||||||      |||   |||:::         :::      |||:::   |||
Db        243 ACCTGGAGCGAGAAGCTCACGTCCGAGTCGTCGACGCACGTGCACGCCGTCGCTGCGGGG 302

Qy         37 ProSerArgProValLeu 42
                 |||   ||||||:::
Db        303 GTGTCCAGCCCAGTCGTG 320

RESULT 61
US-09-796-692-2534
; Sequence 2534, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAP
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2534
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (53)
;   OTHER INFORMATION: n=A,T,C or G
US-09-796-692-2534

Alignment Scores:
Pred. No.:                19.5          Length:         393
Score:                    61.00         Matches:        11
Percent Similarity:       62.1%         Conservative:   7
Best Local Similarity:    37.9%         Mismatches:     11
Query Match:              26.6%         Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-796-692-2534 (1-393)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
              ||||||         :::|||:::||||||::::::|||   ::::::   |||
Db        306 CCACCTCCAGCAAAGAAAGCAACAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 365

Qy         32 PheValSerGluGlyProSerArgPro 40
                 |||   ||||||:::|||
Db        366 GAGGATGATGAAGCTCCTTCTAAGCCA 392

RESULT 62
US-10-040-862-2534
; Sequence 2534, Application US/10040862
```

```
; Publication No. US20030078396A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013520US
;  CURRENT APPLICATION NUMBER: US/10/040,862
;  CURRENT FILING DATE:  2001-11-06
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: US 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: US 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: US 09/796,692
;  PRIOR FILING DATE: 2001-03-01
;  NUMBER OF SEQ ID NOS: 10467
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2534
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (53)
;   OTHER INFORMATION: n=A,T,C or G
US-10-040-862-2534

Alignment Scores:
Pred. No.:                19.5          Length:          393
Score:                    61.00         Matches:         11
Percent Similarity:       62.1%         Conservative:    7
Best Local Similarity:    37.9%         Mismatches:      11
Query Match:              26.6%         Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-040-862-2534 (1-393)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||         :::|||:::|||||||:::::::|||   ::::::   |||
Db         306 CCACCTCCAGCAAAGAAAGCAACAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 365

Qy          32 PheValSerGluGlyProSerArgPro 40
                  |||   |||||||:::|||
Db         366 GAGGATGATGAAGCTCCTTCTAAGCCA 392

RESULT 63
US-10-057-475B-2534
; Sequence 2534, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Clapper, Jonathan David
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Ordonez, Nadia
;  APPLICANT:  Carter, Lauren
;  APPLICANT:  McNeill, Patricia Dianne
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-014402US
;  CURRENT APPLICATION NUMBER: US/10/057,475B
;  CURRENT FILING DATE:  2002-01-22
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10979
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2534
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(393)
;   OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-2534

Alignment Scores:
Pred. No.:                19.5          Length:          393
Score:                    61.00         Matches:         11
Percent Similarity:       62.1%         Conservative:    7
Best Local Similarity:    37.9%         Mismatches:      11
Query Match:              26.6%         Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-057-475B-2534 (1-393)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||         :::|||:::|||||||:::::::|||   ::::::   |||
Db         306 CCACCTCCAGCAAAGAAAGCAACAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 365

Qy          32 PheValSerGluGlyProSerArgPro 40
                  |||   |||||||:::|||
Db         366 GAGGATGATGAAGCTCCTTCTAAGCCA 392

RESULT 64
US-10-154-884B-2534
; Sequence 2534, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
```

```
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013521US
;  CURRENT APPLICATION NUMBER: US/10/154,884B
;  CURRENT FILING DATE:  2002-05-23
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 11290
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2534
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(393)
;   OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-2534

Alignment Scores:
Pred. No.:                19.5          Length:         393
Score:                    61.00         Matches:        11
Percent Similarity:       62.1%         Conservative:   7
Best Local Similarity:    37.9%         Mismatches:     11
Query Match:              26.6%         Indels:         0
DB:                       6             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-154-884B-2534 (1-393)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||          :::|||:::||||||:::::|||   ::::::   |||
Db         306 CCACCTCCAGCAAAGAAAGCAACAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 365

Qy          32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db         366 GAGGATGATGAAGCTCCTTCTAAGCCA 392

RESULT 65
US-10-764-324-2534
; Sequence 2534, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013520US
;  CURRENT APPLICATION NUMBER: US/10/764,324
;  CURRENT FILING DATE:  2004-01-23
;  PRIOR APPLICATION NUMBER: US/10/040,862
;  PRIOR FILING DATE: 2001-11-06
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10467
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2534
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (53)
;   OTHER INFORMATION: n=A,T,C or G
US-10-764-324-2534

Alignment Scores:
Pred. No.:                19.5          Length:         393
Score:                    61.00         Matches:        11
Percent Similarity:       62.1%         Conservative:   7
Best Local Similarity:    37.9%         Mismatches:     11
Query Match:              26.6%         Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-764-324-2534 (1-393)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||          :::|||:::||||||:::::|||   ::::::   |||
Db         306 CCACCTCCAGCAAAGAAAGCAACAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 365

Qy          32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db         366 GAGGATGATGAAGCTCCTTCTAAGCCA 392

RESULT 66
US-09-918-995-29169
; Sequence 29169, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 29169
;   LENGTH: 445
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(445)
```

```
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-29169

Alignment Scores:
Pred. No.:                22.4          Length:         445
Score:                    61.00         Matches:        11
Percent Similarity:       62.1%         Conservative:   7
Best Local Similarity:    37.9%         Mismatches:     11
Query Match:              26.6%         Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-918-995-29169 (1-445)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||        :::|||:::||||||::::::|||   ::::::   |||
Db         166 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 225

Qy          32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db         226 GAGGATGATGAAGCTCCTTCTAAGCCA 252

RESULT 67
US-10-172-118-1080
; Sequence 1080, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Chris
;  APPLICANT:  Van 't Veer, Laura
;  APPLICANT:  Van de Vijver, Marc
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-175-999
;  CURRENT APPLICATION NUMBER: US/10/172,118
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: 60/380,770
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1080
;   LENGTH: 2502
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: NM_004741
;   DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1080

Alignment Scores:
Pred. No.:                152           Length:         2502
Score:                    61.00         Matches:        11
Percent Similarity:       62.1%         Conservative:   7
Best Local Similarity:    37.9%         Mismatches:     11
Query Match:              26.6%         Indels:         0
DB:                       6             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-172-118-1080 (1-2502)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||        :::|||:::||||||::::::|||   ::::::   |||
Db        1108 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 1167

Qy          32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db        1168 GAGGATGATGAAGCTCCTTCTAAGCCA 1194

RESULT 68
US-10-342-887-1080
; Sequence 1080, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter S.
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Christopher J.
;  APPLICANT:  Van 't Veer, Laura Johanna
;  APPLICANT:  Van de Vijver, Marc J.
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-188-999
;  CURRENT APPLICATION NUMBER: US/10/342,887
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: 60/298,918
;  PRIOR FILING DATE: 2001-06-18
;  PRIOR APPLICATION NUMBER: 60/380,710
;  PRIOR FILING DATE: 2002-05-14
;  PRIOR APPLICATION NUMBER: 10/172,118
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1080
;   LENGTH: 2502
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-1080

Alignment Scores:
Pred. No.:                152           Length:         2502
Score:                    61.00         Matches:        11
Percent Similarity:       62.1%         Conservative:   7
Best Local Similarity:    37.9%         Mismatches:     11
Query Match:              26.6%         Indels:         0
DB:                       7             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-342-887-1080 (1-2502)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||        :::|||:::||||||::::::|||   ::::::   |||
Db        1108 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 1167

Qy          32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db        1168 GAGGATGATGAAGCTCCTTCTAAGCCA 1194

RESULT 69
US-10-171-581-49
; Sequence 49, Application US/10171581
; Publication No. US20030104426A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  Linsley, Peter
;  APPLICANT:  Mao Mao
;  TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
;  FILE REFERENCE: 9301-157-999
;  CURRENT APPLICATION NUMBER: US/10/171,581
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: 60/298,914
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 49
;   LENGTH: 3727
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: D21262
;   DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-49

Alignment Scores:
Pred. No.:                236           Length:         3727
Score:                    61.00         Matches:        11
Percent Similarity:       62.1%         Conservative:   7
Best Local Similarity:    37.9%         Mismatches:     11
Query Match:              26.6%         Indels:         0
```

```
DB:                    5              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-171-581-49 (1-3727)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||          :::|||:::||||||:::::::|||   ::::::   |||
Db        1080 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 1139

Qy          32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db        1140 GAGGATGATGAAGCTCCTTCTAAGCCA 1166

RESULT 70
US-10-483-506-24
; Sequence 24, Application US/10483506
; Publication No. US20040249127A1
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT:  Baughn, Mariah R.
;  APPLICANT:  Becha, Shanya D.
;  APPLICANT:  Chinn, Anna M.
;  APPLICANT:  Cocks, Benjamin G.
;  APPLICANT:  Ding, Li
;  APPLICANT:  Forsythe, Ian J.
;  APPLICANT:  Gietzen, Kimberly J.
;  APPLICANT:  Griffin, Jennifer A.
;  APPLICANT:  Hafalia, April J. A.
;  APPLICANT:  Khan, Farrah A.
;  APPLICANT:  Lal, Preeti G.
;  APPLICANT:  Lee, Ernestine A.
;  APPLICANT:  Li, Joana X.
;  APPLICANT:  Lu, Dyung Aina M.
;  APPLICANT:  Ramkumar, Jayalaxmi
;  APPLICANT:  Richardson, Thomas W.
;  APPLICANT:  Swarnakar, Anita
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Walia, Narinder K.
;  APPLICANT:  Warren, Bridget A.
;  APPLICANT:  Yang, Junming
;  APPLICANT:  Yao, Monique G.
;  APPLICANT:  Yue, Henry
;  APPLICANT:  Zebarjadian, Yeganeh
;  TITLE OF INVENTION: Intracellular Signaling Molecules
;  FILE REFERENCE: PF-1066 USN
;  CURRENT APPLICATION NUMBER: US/10/483,506
;  CURRENT FILING DATE:  2004-01-12
;  PRIOR APPLICATION NUMBER: PCT/US02/22379
;  PRIOR FILING DATE: 2002-07-11
;  PRIOR APPLICATION NUMBER: US 60/305,113
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: US 60/305,367
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/306,966
;  PRIOR FILING DATE: 2001-07-19
;  PRIOR APPLICATION NUMBER: US 60/308,175
;  PRIOR FILING DATE: 2001-07-27
;  PRIOR APPLICATION NUMBER: US 60/308,327
;  PRIOR FILING DATE: 2001-07-27
;  PRIOR APPLICATION NUMBER: US 60/309,902
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: US 60/310,752
;  PRIOR FILING DATE: 2001-08-07
;  PRIOR APPLICATION NUMBER: US 60/311,636
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PERL Program
; SEQ ID NO 24
;   LENGTH: 4184
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7461789CB1

US-10-483-506-24

Alignment Scores:
Pred. No.:                268             Length:          4184
Score:                    61.00           Matches:         12
Percent Similarity:       55.6%           Conservative:    3
Best Local Similarity:    44.4%           Mismatches:      12
Query Match:              26.6%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-483-506-24 (1-4184)

Qy           3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
               |||  |||      :::         ||||||      ||||||  ||||||   :::
Db         304 CCCCAGAGCCATGGGCTCGCTCACGCCCCTCCGCACTCCCCGGAGGGCTCCGAGGGCAGC 363

Qy          23 ThrAspGluAsnGluAspAsp 29
                  :::|||   |||||||||
Db         364 GGAGAGGAGGAGGAAGACGAC 384

RESULT 71
US-09-940-316B-20/c
; Sequence 20, Application US/09940316B
; Publication No. US20030175901A1
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTI
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;   LENGTH: 4466
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (9)..(4454)
US-09-940-316B-20

Alignment Scores:
Pred. No.:                288             Length:          4466
Score:                    61.00           Matches:         16
Percent Similarity:       45.8%           Conservative:    6
Best Local Similarity:    33.3%           Mismatches:      16
Query Match:              26.6%           Indels:          10
DB:                       3               Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-20 (1-4466)

Qy           3 ProGlySerGluIleIleGluSerValProProAla----------GlyProGlu 17
               |||      :::   |||         ||||||                ||||||
Db         910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy          18 AlaSerGlu----------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
```

```
                ||||:::                  |||   |||   ::::::|||       :::
Db        850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   |||||||||||||||
Db        790 GTGACCGCCGGACCCTCCAGACCG 767
```

```
RESULT 72
US-09-940-316B-16/c
; Sequence 16, Application US/09940316B
; Publication No. US20030175901A1
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 4478
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4466)
US-09-940-316B-16

Alignment Scores:
Pred. No.:                289             Length:         4478
Score:                    61.00           Matches:        16
Percent Similarity:       45.8%           Conservative:   6
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              26.6%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-16 (1-4478)

Qy           3 ProGlySerGluIleIleGluSerValProProAla---------------GlyProGlu 17
               |||      :::   |||          ||||||                  ||||||
Db         910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy          18 AlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                |||:::                 |||   |||   ::::::|||       :::
Db         850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy          33 ValSerGluGlyProSerArgPro 40
               |||:::   |||||||||||||||
Db         790 GTGACCGCCGGACCCTCCAGACCG 767
```

```
RESULT 73
US-09-940-316B-22/c
; Sequence 22, Application US/09940316B
; Publication No. US20030175901A1
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTHA
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
;   LENGTH: 4547
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (9)..(4535)
US-09-940-316B-22

Alignment Scores:
Pred. No.:                294             Length:         4547
Score:                    61.00           Matches:        16
Percent Similarity:       45.8%           Conservative:   6
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              26.6%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-22 (1-4547)

Qy           3 ProGlySerGluIleIleGluSerValProProAla---------------GlyProGlu 17
               |||      :::   |||          ||||||                  ||||||
Db         910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy          18 AlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                |||:::                 |||   |||   ::::::|||       :::
Db         850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy          33 ValSerGluGlyProSerArgPro 40
               |||:::   |||||||||||||||
Db         790 GTGACCGCCGGACCCTCCAGACCG 767
```

```
RESULT 74
US-09-940-316B-18/c
; Sequence 18, Application US/09940316B
; Publication No. US20030175901A1
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNT
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
```

```
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
;   LENGTH: 4571
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (9)..(4559)
US-09-940-316B-18

Alignment Scores:
Pred. No.:              296          Length:        4571
Score:                  61.00        Matches:       16
Percent Similarity:     45.8%        Conservative:  6
Best Local Similarity:  33.3%        Mismatches:    16
Query Match:            26.6%        Indels:        10
DB:                     3            Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-18 (1-4571)

Qy          3 ProGlySerGluIleIleGluSerValProProAla--------------GlyProGlu 17
              |||      :::   |||         ||||||                      ||||||
Db        910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy         18 AlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                ||:::                |||   |||   ::::::|||      :::
Db        850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   ||||||||||||||
Db        790 GTGACCGCCGGACCCTCCAGACCG 767

RESULT 75
US-11-097-143-2012
; Sequence 2012, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2012
;   LENGTH: 4587
;   TYPE: DNA
;   ORGANISM: DROSOPHILA
US-11-097-143-2012

Alignment Scores:
Pred. No.:              297          Length:        4587
Score:                  61.00        Matches:       11
Percent Similarity:     54.8%        Conservative:  6
Best Local Similarity:  35.5%        Mismatches:    14
Query Match:            26.6%        Indels:        0
DB:                     10           Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-097-143-2012 (1-4587)

Qy         13 ProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPhe 32
              |||   |||||||||           |||   |||::::::   :::          :::
Db       3669 CCAATTGGCCCGGAGTTTGGTTCTCCAACTGCAGATGATCATCGAGAACAATGTAGCTGG 3728

Qy         33 ValSerGluGlyProSerArgProValLeuGlu 43
                :::||||||   :::   |||   ||||||
Db       3729 TGCAACGAGGGAACTACGCCGCCTTATCTCGAG 3761


Search completed: March 11, 2006, 05:53:46
Job time : 364.027 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 20:06:44 ; Search time 50.038 Seconds
(without alignments)
1563.065 Million cell updates/sec

Title: US-10-757-745-6_COPY_2_45
Perfect score: 229
Sequence: 1 GDPGSEIIESVPPAGPEASE.....ENEDDIQFVSEGPSRPVLEY 44

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170331_4111/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_417_@runat_09032006_170331_4111 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	229	100.0	274	3	US-09-513-999C-1651	Sequence 1651, Ap
	2	229	100.0	1536	3	US-09-697-863A-5	Sequence 5, Appli
	3	67	29.3	8100	3	US-09-554-337-4	Sequence 4, Appli
	4	67	29.3	11517	2	US-07-920-281C-1	Sequence 1, Appli
	5	67	29.3	11517	3	US-08-466-277-1	Sequence 1, Appli
	6	67	29.3	11517	3	US-09-688-842-1	Sequence 1, Appli
	7	67	29.3	15538	3	US-09-554-337-1	Sequence 1, Appli
c	8	67	29.3	77536	3	US-09-410-551B-1	Sequence 1, Appli
c	9	67	29.3	77536	3	US-09-940-316B-1	Sequence 1, Appli
	10	65	28.4	77536	3	US-09-410-551B-1	Sequence 1, Appli
	11	65	28.4	77536	3	US-09-940-316B-1	Sequence 1, Appli
	12	64	27.9	3211	2	US-08-574-959A-8	Sequence 8, Appli
	13	64	27.9	3211	3	US-09-357-014-8	Sequence 8, Appli
	14	64	27.9	3901	2	US-08-574-959A-6	Sequence 6, Appli
	15	64	27.9	3901	3	US-09-357-014-6	Sequence 6, Appli
	16	64	27.9	7390	3	US-09-566-921-95	Sequence 95, Appl
c	17	64	27.9	58857	3	US-09-477-962-1	Sequence 1, Appli
	18	61.5	26.9	1031	3	US-09-276-438-4	Sequence 4, Appli
c	19	61.5	26.9	1758	3	US-09-252-991A-1644	Sequence 1644, Ap
c	20	61	26.6	4466	3	US-09-410-551B-20	Sequence 20, Appl
c	21	61	26.6	4466	3	US-09-940-316B-20	Sequence 20, Appl
c	22	61	26.6	4478	3	US-09-410-551B-16	Sequence 16, Appl
c	23	61	26.6	4478	3	US-09-940-316B-16	Sequence 16, Appl
c	24	61	26.6	4547	3	US-09-410-551B-22	Sequence 22, Appl
c	25	61	26.6	4547	3	US-09-940-316B-22	Sequence 22, Appl
c	26	61	26.6	4571	3	US-09-410-551B-18	Sequence 18, Appl
c	27	61	26.6	4571	3	US-09-940-316B-18	Sequence 18, Appl
	28	61	26.6	13290	3	US-09-949-016-13937	Sequence 13937, A
	29	61	26.6	13290	3	US-09-949-016-13938	Sequence 13938, A
c	30	61	26.6	24707	3	US-09-740-027-3	Sequence 3, Appli
c	31	61	26.6	24707	3	US-10-274-968-3	Sequence 3, Appli
c	32	61	26.6	24720	3	US-09-949-016-12341	Sequence 12341, A
c	33	61	26.6	24721	3	US-09-949-016-15610	Sequence 15610, A
	34	60	26.2	9164	3	US-09-814-915A-80	Sequence 80, Appl
	35	59.5	26.0	2383	3	US-09-523-487-9	Sequence 9, Appli
	36	59.5	26.0	2409	3	US-09-320-095-9	Sequence 9, Appli
c	37	59.5	26.0	18836	3	US-09-949-016-12654	Sequence 12654, A
c	38	59.5	26.0	18838	3	US-09-949-016-15633	Sequence 15633, A
	39	59.5	26.0	36731	3	US-09-949-016-13770	Sequence 13770, A
c	40	59	25.8	335	3	US-09-702-705-372	Sequence 372, App
c	41	59	25.8	335	3	US-09-736-457-372	Sequence 372, App
c	42	59	25.8	335	3	US-09-614-124B-372	Sequence 372, App
c	43	59	25.8	335	3	US-09-671-325-372	Sequence 372, App
c	44	59	25.8	335	3	US-09-589-184-372	Sequence 372, App
c	45	59	25.8	335	3	US-09-658-824-372	Sequence 372, App
c	46	59	25.8	335	3	US-10-017-754-372	Sequence 372, App
c	47	59	25.8	335	3	US-09-651-563-372	Sequence 372, App
c	48	59	25.8	335	3	US-09-519-642-372	Sequence 372, App
	49	59	25.8	427	3	US-09-513-999C-14793	Sequence 14793, A
	50	59	25.8	543	3	US-09-949-016-2252	Sequence 2252, Ap
	51	59	25.8	543	3	US-09-949-016-5158	Sequence 5158, Ap
c	52	59	25.8	601	3	US-09-949-016-77086	Sequence 77086, A
	53	59	25.8	645	3	US-09-854-133-627	Sequence 627, App
	54	59	25.8	944	2	US-08-786-606-4	Sequence 4, Appli
	55	59	25.8	1112	2	US-08-933-750C-97	Sequence 97, Appl
	56	59	25.8	1112	3	US-09-234-613-97	Sequence 97, Appl
	57	59	25.8	1114	3	US-09-690-454-39	Sequence 39, Appl
	58	59	25.8	1329	3	US-09-902-540-64	Sequence 64, Appl
	59	59	25.8	3172	3	US-09-976-594-400	Sequence 400, App
	60	59	25.8	3817	3	US-09-949-016-5552	Sequence 5552, Ap
	61	59	25.8	3835	3	US-09-949-016-1067	Sequence 1067, Ap
	62	59	25.8	4543	3	US-09-949-016-13994	Sequence 13994, A
c	63	59	25.8	4674	3	US-09-410-551B-26	Sequence 26, Appl
c	64	59	25.8	4674	3	US-09-940-316B-26	Sequence 26, Appl
c	65	59	25.8	4725	3	US-09-410-551B-24	Sequence 24, Appl
c	66	59	25.8	4725	3	US-09-940-316B-24	Sequence 24, Appl
c	67	59	25.8	4737	3	US-09-410-551B-30	Sequence 30, Appl
c	68	59	25.8	4737	3	US-09-940-316B-30	Sequence 30, Appl
c	69	59	25.8	4767	3	US-09-410-551B-28	Sequence 28, Appl
c	70	59	25.8	4767	3	US-09-940-316B-28	Sequence 28, Appl
c	71	59	25.8	4818	3	US-09-410-551B-32	Sequence 32, Appl
c	72	59	25.8	4818	3	US-09-940-316B-32	Sequence 32, Appl
	73	59	25.8	7137	3	US-09-949-002-28	Sequence 28, Appl
	74	59	25.8	7434	3	US-09-949-002-112	Sequence 112, App
	75	59	25.8	27236	3	US-09-949-016-12809	Sequence 12809, A

ALIGNMENTS

RESULT 1
US-09-513-999C-1651
; Sequence 1651, Application US/09513999C

; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 1651
; LENGTH: 274
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 108..272
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 18
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 43
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 45
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 52
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 54
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 60
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 62
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128
; OTHER INFORMATION: r=a or g
US-09-513-999C-1651

Alignment Scores:

Pred. No.:	1.25e-23	Length:	274
Score:	229.00	Matches:	44
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-513-999C-1651 (1-274)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||
Db        111 GGAGACCCGGGGTCGGARATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 170

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||
Db        171 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 230

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        231 GTTCTTGAATAC 242
```

RESULT 2
US-09-697-863A-5
; Sequence 5, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/09/697,863A
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1530)..(1531)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (209)..(1534)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1392)..(1392)
; OTHER INFORMATION: N stands for any nucleotide.
US-09-697-863A-5

Alignment Scores:

Pred. No.:	1.09e-22	Length:	1536
Score:	229.00	Matches:	44
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-697-863A-5 (1-1536)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||
Db        212 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 271

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||
Db        272 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 331

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        332 GTTCTTGAATAC 343
```

RESULT 3
US-09-554-337-4
; Sequence 4, Application US/09554337
; Patent No. 6475780
; GENERAL INFORMATION:
; APPLICANT: Parrington, Mark
; APPLICANT: Li, Xiaomao
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
; FILE REFERENCE: 1038-1042 MIS
; CURRENT APPLICATION NUMBER: US/09/554,337
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/065,791

```
;  PRIOR FILING DATE: 1997-11-14
;  PRIOR APPLICATION NUMBER: PCT/CA98/01064
;  PRIOR FILING DATE: 1998-11-13
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 8100
;   TYPE: DNA
;   ORGANISM: respiratory syncytial virus
US-09-554-337-4

Alignment Scores:
Pred. No.:                55.4          Length:         8100
Score:                    67.00         Matches:        14
Percent Similarity:       51.2%         Conservative:   7
Best Local Similarity:    34.1%         Mismatches:     20
Query Match:              29.3%         Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-554-337-4 (1-8100)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||              |||:::
Db       2701 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 2760

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||   |||      |||   |||||||||:::   |||
Db       2761 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 2820

Qy         41 Val 41
              |||
Db       2821 GTG 2823

RESULT 4
US-07-920-281C-1
; Sequence 1, Application US/07920281C
; Patent No. 5739026
;  GENERAL INFORMATION:
;    APPLICANT:  Garoff, Henrik
;    APPLICANT:  Liljestrom, Peter
;    TITLE OF INVENTION:  DNA Expression Systems Based on
;    TITLE OF INVENTION:  Alphaviruses
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch
;      STREET:  P.O. Box 747
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/920,281C
;      FILING DATE:  13-AUG-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  828-103P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11517 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Semliki Forest Virus
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..11517
;      OTHER INFORMATION:  /label= genome
;      OTHER INFORMATION:  /note= "Semliki Forest Virus complete nucleotide
;      OTHER INFORMATION:  sequence, presented as a cloned DNA sequence; see
;      OTHER INFORMATION:  Figure 5."
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  87..7379
;      OTHER INFORMATION:  /product= "SFV polyprotein"
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  7421..11179
;      OTHER INFORMATION:  /product= "SFV polyprotein"
US-07-920-281C-1

Alignment Scores:
Pred. No.:                86.3          Length:         11517
Score:                    67.00         Matches:        14
Percent Similarity:       51.2%         Conservative:   7
Best Local Similarity:    34.1%         Mismatches:     20
Query Match:              29.3%         Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-07-920-281C-1 (1-11517)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||              |||:::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||   |||      |||   |||||||||:::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101

RESULT 5
US-08-466-277-1
; Sequence 1, Application US/08466277
; Patent No. 6190666
;  GENERAL INFORMATION:
;       APPLICANT: Garoff, Henrik
;                  Liljestrom, Peter
;       TITLE OF INVENTION: DNA Expression Systems Based on
;                           Alphaviruses
;       NUMBER OF SEQUENCES: 27
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Birch, Stewart, Kolasch & Birch
;            STREET: P.O. Box 747
;            CITY: Falls Church
;            STATE: Virginia
;            COUNTRY: USA
;            ZIP: 22040-0747
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/466,277
;            FILING DATE: 06-Jun-1995
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 07/920,281
;            FILING DATE: <Unknown>
```

```
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Murphy Jr., Gerald M.
;            REGISTRATION NUMBER: 28,977
;            REFERENCE/DOCKET NUMBER: 828-103P
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-241-1300
;            TELEFAX: 703-241-2848
;            TELEX: 248345
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 11517 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: RNA (genomic)
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       ORIGINAL SOURCE:
;            ORGANISM: Semliki Forest Virus
;       FEATURE:
;            NAME/KEY:  -
;            LOCATION:  1..11517
;            OTHER INFORMATION:  /label= genome
;            /note= "Semliki Forest Virus complete nucleotide
;            sequence, presented as a cloned DNA sequence; see
;            Figure 5."
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  87..7379
;            OTHER INFORMATION:  /product= "SFV polyprotein"
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  7421..11179
;            OTHER INFORMATION:  /product= "SFV polyprotein"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-466-277-1

Alignment Scores:
Pred. No.:                86.3          Length:        11517
Score:                    67.00         Matches:       14
Percent Similarity:       51.2%         Conservative:  7
Best Local Similarity:    34.1%         Mismatches:    20
Query Match:              29.3%         Indels:        0
DB:                       3             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-08-466-277-1 (1-11517)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::|||              |||:::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||    |||   |||     |||   ||||||||::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101

RESULT 6
US-09-688-842-1
; Sequence 1, Application US/09688842
; Patent No. 6770283
;   GENERAL INFORMATION:
;       APPLICANT: Garoff, Henrik
;                  Liljestrom, Peter
;       TITLE OF INVENTION: DNA Expression Systems Based on
;                           Alphaviruses
;       NUMBER OF SEQUENCES: 27
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Birch, Stewart, Kolasch & Birch
;            STREET: P.O. Box 747
;            CITY: Falls Church
;            STATE: Virginia
;            COUNTRY: USA
;            ZIP: 22040-0747
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/688,842
;            FILING DATE: 17-Oct-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/466,277
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Murphy Jr., Gerald M.
;            REGISTRATION NUMBER: 28,977
;            REFERENCE/DOCKET NUMBER: 828-103P
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-241-1300
;            TELEFAX: 703-241-2848
;            TELEX: 248345
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 11517 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: RNA (genomic)
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       ORIGINAL SOURCE:
;            ORGANISM: Semliki Forest Virus
;       FEATURE:
;            NAME/KEY:  -
;            LOCATION:  1..11517
;            OTHER INFORMATION:  /label= genome
;            /note= "Semliki Forest Virus complete nucleotide
;            sequence, presented as a cloned DNA sequence; see
;            Figure 5."
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  87..7379
;            OTHER INFORMATION:  /product= "SFV polyprotein"
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  7421..11179
;            OTHER INFORMATION:  /product= "SFV polyprotein"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-688-842-1

Alignment Scores:
Pred. No.:                86.3          Length:        11517
Score:                    67.00         Matches:       14
Percent Similarity:       51.2%         Conservative:  7
Best Local Similarity:    34.1%         Mismatches:    20
Query Match:              29.3%         Indels:        0
DB:                       3             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-688-842-1 (1-11517)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::|||              |||:::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||    |||   |||     |||   ||||||||::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101
```

```
RESULT 7
US-09-554-337-1
; Sequence 1, Application US/09554337
; Patent No. 6475780
; GENERAL INFORMATION:
;  APPLICANT: Parrington, Mark
;  APPLICANT:  Li, Xiaomao
;  APPLICANT:  Klein, Michel H.
;  TITLE OF INVENTION: ALPHAVIRUS VECTORS FOR PARAMYXOVIRUS VACCINES
;  FILE REFERENCE: 1038-1042 MIS
;  CURRENT APPLICATION NUMBER: US/09/554,337
;  CURRENT FILING DATE:  2000-10-20
;  PRIOR APPLICATION NUMBER: 60/065,791
;  PRIOR FILING DATE: 1997-11-14
;  PRIOR APPLICATION NUMBER: PCT/CA98/01064
;  PRIOR FILING DATE: 1998-11-13
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 15538
;   TYPE: DNA
;   ORGANISM: respiratory syncytial virus
US-09-554-337-1

Alignment Scores:
Pred. No.:                126             Length:         15538
Score:                    67.00           Matches:        14
Percent Similarity:       51.2%           Conservative:   7
Best Local Similarity:    34.1%           Mismatches:     20
Query Match:              29.3%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-554-337-1 (1-15538)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||               |||:::
Db       5525 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 5584

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                        |||      |||   |||      |||   |||||||||:::   |||
Db       5585 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 5644

Qy         41 Val 41
              |||
Db       5645 GTG 5647

RESULT 8
US-09-410-551B-1/c
; Sequence 1, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 77536
;   TYPE: DNA
;   ORGANISM: Streptomyces hygroscopicus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (52275)...(71465)
US-09-410-551B-1

Alignment Scores:
Pred. No.:                952             Length:         77536
Score:                    67.00           Matches:        17
Percent Similarity:       46.9%           Conservative:   6
Best Local Similarity:    34.7%           Mismatches:     16
Query Match:              29.3%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-1 (1-77536)

Qy          2 AspProGlySerGluIleIleGluSerValProProAla----------GlyPro 16
              :::|||      |||   |||        ||||||                |||
Db       52933 AACCCCGCCGGCGAGGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCG 52874

Qy         17 GluAlaSerGlu----------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||   |||:::          |||   |||   ::::::|||      :::
Db       52873 GAGCGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCG 52814

Qy         32 PheValSerGluGlyProSerArgPro 40
                 |||:::   ||||||||||||||
Db       52813 ACCGTGACCGCCGGACCCTCCAGACCG 52787

RESULT 9
US-09-940-316B-1/c
; Sequence 1, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTI
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 77536
;   TYPE: DNA
;   ORGANISM: Streptomyces hygroscopicus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (52275)...(71465)
US-09-940-316B-1

Alignment Scores:
Pred. No.:                952             Length:         77536
Score:                    67.00           Matches:        17
Percent Similarity:       46.9%           Conservative:   6
Best Local Similarity:    34.7%           Mismatches:     16
Query Match:              29.3%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-1 (1-77536)
```

```
Qy          2 AspProGlySerGluIleIleGluSerValProProAla---------------GlyPro 16
              :::|||     |||   |||        ||||||                    |||
Db      52933 AACCCCGCCGGCGAGGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCG 52874

Qy         17 GluAlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||   |||:::               |||   |||   :::::::|||      :::
Db      52873 GAGCGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCG 52814

Qy         32 PheValSerGluGlyProSerArgPro 40
                |||:::   ||||||||||||||||
Db      52813 ACCGTGACCGCCGGACCCTCCAGACCG 52787

RESULT 10
US-09-410-551B-1
; Sequence 1, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 77536
;   TYPE: DNA
;   ORGANISM: Streptomyces hygroscopicus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (52275)...(71465)
US-09-410-551B-1

Alignment Scores:
Pred. No.:                1.82e+03        Length:          77536
Score:                    65.00           Matches:         17
Percent Similarity:       45.8%           Conservative:    5
Best Local Similarity:    35.4%           Mismatches:      16
Query Match:              28.4%           Indels:          10
DB:                       3               Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-1 (1-77536)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------GlyProGlu 17
              |||       |||   |||        ||||||                    ||||||
Db      28077 CCGGACGGCGAGGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 28136

Qy         18 AlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
              |||:::               |||   |||   :::::::|||      :::
Db      28137 CGCAGTGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 28196

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   ||||||||||||||||
Db      28197 GTGACCGCCGGACCCTCCAGACCG 28220

RESULT 11
US-09-940-316B-1
; Sequence 1, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH.
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 77536
;   TYPE: DNA
;   ORGANISM: Streptomyces hygroscopicus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (52275)...(71465)
US-09-940-316B-1

Alignment Scores:
Pred. No.:                1.82e+03        Length:          77536
Score:                    65.00           Matches:         17
Percent Similarity:       45.8%           Conservative:    5
Best Local Similarity:    35.4%           Mismatches:      16
Query Match:              28.4%           Indels:          10
DB:                       3               Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-1 (1-77536)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------GlyProGlu 17
              |||       |||   |||        ||||||                    ||||||
Db      28077 CCGGACGGCGAGGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 28136

Qy         18 AlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
              |||:::               |||   |||   :::::::|||      :::
Db      28137 CGCAGTGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 28196

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   ||||||||||||||||
Db      28197 GTGACCGCCGGACCCTCCAGACCG 28220

RESULT 12
US-08-574-959A-8
; Sequence 8, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
;   APPLICANT:  Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;   APPLICANT:  and Jack L. Strominger
;   TITLE OF INVENTION:  p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;   TITLE OF INVENTION:  AND USES THEREFOR
;   NUMBER OF SEQUENCES:  22
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  LAHIVE & COCKFIELD
;     STREET:  60 State Street, Suite 510
;     CITY:  Boston
;     STATE:  Massachusetts
;     COUNTRY:  USA
;     ZIP:  02109-1875
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/574,959A
;       FILING DATE:  19-DEC-95
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mandragouras, Amy E.
;       REGISTRATION NUMBER:  36,207
;       REFERENCE/DOCKET NUMBER:  DFN-008
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617)227-7400
;       TELEFAX:  (617)227-5941
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3211 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  439..3157
US-08-574-959A-8

Alignment Scores:
Pred. No.:                45.7          Length:         3211
Score:                    64.00         Matches:        13
Percent Similarity:       46.5%         Conservative:   7
Best Local Similarity:    30.2%         Mismatches:     23
Query Match:              27.9%         Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-08-574-959A-8 (1-3211)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||          |||         |||
Db       2902 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTNCTCCCCCAACCCTGTTGGAA 2961

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                |||::::::     |||   :::|||         |||   |||:::
Db       2962 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3021

Qy         41 ValLeuGlu 43
                :::|||
Db       3022 GAGATGGAG 3030

RESULT 13
US-09-357-014-8
; Sequence 8, Application US/09357014
; Patent No. 6291645
;   GENERAL INFORMATION:
;       APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;                  and Jack L. Strominger
;       TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;                           AND USES THEREFOR
;       NUMBER OF SEQUENCES: 22
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: LAHIVE & COCKFIELD
;            STREET: 60 State Street, Suite 510
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02109-1875
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/357,014
;            FILING DATE: 19-Jul-1999
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/574,959
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mandragouras, Amy E.
;            REGISTRATION NUMBER: 36,207
;            REFERENCE/DOCKET NUMBER: DFN-008
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)227-7400
;            TELEFAX: (617)227-5941
;   INFORMATION FOR SEQ ID NO: 8:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 3211 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  439..3157
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-357-014-8

Alignment Scores:
Pred. No.:                45.7          Length:         3211
Score:                    64.00         Matches:        13
Percent Similarity:       46.5%         Conservative:   7
Best Local Similarity:    30.2%         Mismatches:     23
Query Match:              27.9%         Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-357-014-8 (1-3211)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||          |||         |||
Db       2902 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTNCTCCCCCAACCCTGTTGGAA 2961

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                |||::::::     |||   :::|||         |||   |||:::
Db       2962 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3021

Qy         41 ValLeuGlu 43
                :::|||
Db       3022 GAGATGGAG 3030

RESULT 14
US-08-574-959A-6
; Sequence 6, Application US/08574959A
; Patent No. 5962224
;  GENERAL INFORMATION:
;    APPLICANT:  Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;    APPLICANT:  and Jack L. Strominger
;    TITLE OF INVENTION:  p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/574,959A
;      FILING DATE:  19-DEC-95
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  DFN-008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
```

```
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3901 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  439..3847
US-08-574-959A-6

Alignment Scores:
Pred. No.:                58.4          Length:         3901
Score:                    64.00         Matches:        13
Percent Similarity:       46.5%         Conservative:   7
Best Local Similarity:    30.2%         Mismatches:     23
Query Match:              27.9%         Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-08-574-959A-6 (1-3901)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||  |||      |||:::::::|||      |||          |||          |||
Db       3592 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTNCTCCCCCAACCCTGTTGGAA 3651

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                |||::::::     |||   :::|||         |||   |||:::
Db       3652 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3711

Qy         41 ValLeuGlu 43
                :::|||
Db       3712 GAGATGGAG 3720

RESULT 15
US-09-357-014-6
; Sequence 6, Application US/09357014
; Patent No. 6291645
;   GENERAL INFORMATION:
;        APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
;                   and Jack L. Strominger
;        TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
;                            AND USES THEREFOR
;        NUMBER OF SEQUENCES: 22
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: LAHIVE & COCKFIELD
;             STREET: 60 State Street, Suite 510
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02109-1875
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/357,014
;             FILING DATE: 19-Jul-1999
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/574,959
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mandragouras, Amy E.
;             REGISTRATION NUMBER: 36,207
;             REFERENCE/DOCKET NUMBER: DFN-008
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)227-7400
;             TELEFAX: (617)227-5941
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 3901 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  439..3847
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-357-014-6

Alignment Scores:
Pred. No.:                58.4          Length:         3901
Score:                    64.00         Matches:        13
Percent Similarity:       46.5%         Conservative:   7
Best Local Similarity:    30.2%         Mismatches:     23
Query Match:              27.9%         Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-357-014-6 (1-3901)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||  |||      |||:::::::|||      |||          |||          |||
Db       3592 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTNCTCCCCCAACCCTGTTGGAA 3651

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                |||::::::     |||   :::|||         |||   |||:::
Db       3652 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3711

Qy         41 ValLeuGlu 43
                :::|||
Db       3712 GAGATGGAG 3720

RESULT 16
US-09-566-921-95
; Sequence 95, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
;  APPLICANT:  Loring, Jeanne F.
;  APPLICANT:  Tingley, Debora W.
;  APPLICANT:  Edwards, Carla M.
;  TITLE OF INVENTION:  GENES EXPRESSED IN ALZHEIMER'S DISEASE
;  FILE REFERENCE:  PA-0024 US
;  CURRENT APPLICATION NUMBER: US/09/566,921
;  CURRENT FILING DATE:  2000-05-05
;  NUMBER OF SEQ ID NOS:  138
;  SOFTWARE:  PERL Program
; SEQ ID NO 95
;   LENGTH: 7390
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6682888 337896.11
;   NAME/KEY: unsure
;   LOCATION: 5097-5416
;   OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-95

Alignment Scores:
Pred. No.:                130           Length:         7390
Score:                    64.00         Matches:        17
Percent Similarity:       45.7%         Conservative:   4
Best Local Similarity:    37.0%         Mismatches:     17
Query Match:              27.9%         Indels:         8
DB:                       3             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-566-921-95 (1-7390)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||  |||      :::|||:::   ||||||
Db        332 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 391

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
```

```
                ||||      ||||||||||||   |||   |||:::   |||            :::
Db        392 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 451

Qy         35 GluGlyProSerArgPro 40
                |||   ||||||   |||
Db        452 GAGTCGCCCTCGCACCCG 469
```

RESULT 17
US-09-477-962-1/c
; Sequence 1, Application US/09477962
; Patent No. 6927286
; GENERAL INFORMATION:
; APPLICANT: SHEN, BEN
; APPLICANT: DU, LIANGCHENG
; APPLICANT: SANCHEZ, CESAR
; APPLICANT: CHEN, MEI
; APPLICANT: EDWARDS, DANIEL J.
; TITLE OF INVENTION: BLEOMYCIN GENE CLUSTER COMPONENTS AND THEIR USES
; FILE REFERENCE: 407T-895820US
; CURRENT APPLICATION NUMBER: US/09/477,962
; CURRENT FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: 60/115,435
; PRIOR FILING DATE: 1999-01-06
; PRIOR APPLICATION NUMBER: 60/118,848
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1
; LENGTH: 58857
; TYPE: DNA
; ORGANISM: Streptomyces verticillus
US-09-477-962-1

Alignment Scores:

Pred. No.:	1.78e+03	Length:	58857
Score:	64.00	Matches:	14
Percent Similarity:	52.5%	Conservative:	7
Best Local Similarity:	35.0%	Mismatches:	15
Query Match:	27.9%	Indels:	4
DB:	3	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-477-962-1 (1-58857)

```
Qy          1 GlyAspProGlySerGluIleIleGluSer-----------ValProProAlaGlyPro 16
              ||||||   |||   |||:::::::::                  ||||||   |||
Db      34476 GGCGACGGCGGCGGGGAAGTTCTCGACGACCATGATGCTGTCGAACAGCCCGCCGCCCCG 34417

Qy         17 GluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGly 36
                ||||||:::      |||      |||  |||::::::   |||   :::|||
Db      34416 GCCCAGTCCGAGACCGCGCTCGACGCGGCTGAGCGGGACGTGCGCGTGGTCCCACAGGGC 34357
```

RESULT 18
US-09-276-438-4
; Sequence 4, Application US/09276438
; Patent No. 6600027
; GENERAL INFORMATION:
; APPLICANT: Krishnan, Rajendra
; APPLICANT: Coleman, Rebecca A.
; APPLICANT: Yoder, Christine C.
; APPLICANT: Durtschi, Becky A.
; APPLICANT: Brake, David
; TITLE OF INVENTION: POLYNUCLEOTIDE MOLECULES ENCODING NEOSPORA PROTEINS
; FILE REFERENCE: PC9943A
; CURRENT APPLICATION NUMBER: US/09/276,438
; CURRENT FILING DATE: 1999-03-25
; EARLIER APPLICATION NUMBER: 60/079,389
; EARLIER FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: 60/112,282
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 4
; LENGTH: 1031
; TYPE: DNA
; ORGANISM: Neospora caninum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(660)
US-09-276-438-4

Alignment Scores:

Pred. No.:	24.6	Length:	1031
Score:	61.50	Matches:	13
Percent Similarity:	61.1%	Conservative:	9
Best Local Similarity:	36.1%	Mismatches:	13
Query Match:	26.9%	Indels:	1
DB:	3	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-276-438-4 (1-1031)

```
Qy          9 GluSerValProProAlaGlyProGluAlaSerGlu---SerThrThrAspGluAsnGlu 27
              |||   |||||||||:::||||||   :::   |||        |||   :::   :::
Db        493 GAAAAAGTGCCCCCCAGTGGCCCGCGATCAGGTGAAGTTCAGCGTACTCGTAAAGAGCAA 552

Qy         28 AspAspIleGlnPheValSerGluGlyProSerArgProValLeuGlu 43
              :::|||:::|||      :::|||      :::      :::||||||
Db        553 AATGACGTGCAGCAAACCGCAGAGATGTTGGCTGAGGAAATTCTTGAG 600
```

RESULT 19
US-09-252-991A-1644/c
; Sequence 1644, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1644
; LENGTH: 1758
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1644

Alignment Scores:

Pred. No.:	48.1	Length:	1758
Score:	61.50	Matches:	19
Percent Similarity:	49.1%	Conservative:	9
Best Local Similarity:	33.3%	Mismatches:	14
Query Match:	26.9%	Indels:	15
DB:	3	Gaps:	4

US-10-757-745-6_COPY_2_45 (1-44) x US-09-252-991A-1644 (1-1758)

```
Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              |||||||||:::|||:::::::::   ||||||   |||   |||   |||   :::
Db        933 GACCCCGGCGCCGAGCTGCTCCAGCAGGTCCCGCGGGCCGATCCCGGAGCGTTGCAGCAG 874

Qy         22 Thr------ThrAspGlu------AsnGluAspAspIleGln----------------- 31
                  ||||||           :::||||||:::
Db        873 GCACGGCGAACCGATCGCGCCGGCGCTCAGGACGATCTCGCGGCGCGCCTCAAAGCGCTG 814

Qy         32 PheValSer---------------GluGlyProSerArgProValLeuGlu 43
              |||   :::                :::||||||         |||||||||
Db        813 TTCGCGGCCCTGCCAACGGACCGACAGGGCCCTGGCGCGGCCGTCCTCGAG 763
```

RESULT 20
US-09-410-551B-20/c
; Sequence 20, Application US/09410551B
; Patent No. 6503737

```
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;   LENGTH: 4466
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (9)..(4454)
US-09-410-551B-20

Alignment Scores:
Pred. No.:                183             Length:         4466
Score:                    61.00           Matches:        16
Percent Similarity:       45.8%           Conservative:   6
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              26.6%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-20 (1-4466)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------GlyProGlu 17
              |||       :::   |||       ||||||                      ||||||
Db        910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy         18 AlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                 |||:::               |||   |||   ::::::|||      :::
Db        850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   ||||||||||||||
Db        790 GTGACCGCCGGACCCTCCAGACCG 767

RESULT 21
US-09-940-316B-20/c
; Sequence 20, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;   LENGTH: 4466
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (9)..(4454)
US-09-940-316B-20

Alignment Scores:
Pred. No.:                183             Length:         4466
Score:                    61.00           Matches:        16
Percent Similarity:       45.8%           Conservative:   6
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              26.6%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-20 (1-4466)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------GlyProGlu 17
              |||       :::   |||       ||||||                      ||||||
Db        910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy         18 AlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                 |||:::               |||   |||   ::::::|||      :::
Db        850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   ||||||||||||||
Db        790 GTGACCGCCGGACCCTCCAGACCG 767

RESULT 22
US-09-410-551B-16/c
; Sequence 16, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 4478
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (3)..(4466)
US-09-410-551B-16

Alignment Scores:
Pred. No.:                184             Length:         4478
Score:                    61.00           Matches:        16
Percent Similarity:       45.8%           Conservative:   6
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              26.6%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-16 (1-4478)

Qy          3 ProGlySerGluIleIleGluSerValProProAla--------------GlyProGlu 17
              |||      :::   |||         ||||||                    ||||||
Db        910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy         18 AlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                 |||:::               |||   |||   ::::::|||      :::
Db        850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   |||||||||||||||
Db        790 GTGACCGCCGGACCCTCCAGACCG 767

RESULT 23
US-09-940-316B-16/c
; Sequence 16, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 4478
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4466)
US-09-940-316B-16

Alignment Scores:
Pred. No.:                184             Length:         4478
Score:                    61.00           Matches:        16
Percent Similarity:       45.8%           Conservative:   6
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              26.6%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-16 (1-4478)

Qy          3 ProGlySerGluIleIleGluSerValProProAla--------------GlyProGlu 17
              |||      :::   |||         ||||||                    ||||||
Db        910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy         18 AlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                 |||:::               |||   |||   ::::::|||      :::
Db        850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   |||||||||||||||
Db        790 GTGACCGCCGGACCCTCCAGACCG 767

RESULT 24
US-09-410-551B-22/c
; Sequence 22, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
;   LENGTH: 4547
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (9)..(4535)
US-09-410-551B-22

Alignment Scores:
Pred. No.:                187             Length:         4547
Score:                    61.00           Matches:        16
Percent Similarity:       45.8%           Conservative:   6
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              26.6%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-22 (1-4547)

Qy          3 ProGlySerGluIleIleGluSerValProProAla--------------GlyProGlu 17
              |||      :::   |||         ||||||                    ||||||
Db        910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy         18 AlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                 |||:::               |||   |||   ::::::|||      :::
Db        850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   |||||||||||||||
Db        790 GTGACCGCCGGACCCTCCAGACCG 767

RESULT 25
US-09-940-316B-22/c
```

```
; Sequence 22, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
;   LENGTH: 4547
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (9)..(4535)
US-09-940-316B-22

Alignment Scores:
Pred. No.:              187           Length:         4547
Score:                  61.00         Matches:        16
Percent Similarity:     45.8%         Conservative:   6
Best Local Similarity:  33.3%         Mismatches:     16
Query Match:            26.6%         Indels:         10
DB:                     3             Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-22 (1-4547)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------GlyProGlu 17
              |||       :::   |||         ||||||                  ||||||
Db        910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy         18 AlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                 |||:::               |||   |||   ::::::|||      :::
Db        850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   ||||||||||||||
Db        790 GTGACCGCCGGACCCTCCAGACCG 767

RESULT 26
US-09-410-551B-18/c
; Sequence 18, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
;   LENGTH: 4571
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (9)..(4559)
US-09-410-551B-18

Alignment Scores:
Pred. No.:              188           Length:         4571
Score:                  61.00         Matches:        16
Percent Similarity:     45.8%         Conservative:   6
Best Local Similarity:  33.3%         Mismatches:     16
Query Match:            26.6%         Indels:         10
DB:                     3             Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-18 (1-4571)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------GlyProGlu 17
              |||       :::   |||         ||||||                  ||||||
Db        910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy         18 AlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                 |||:::               |||   |||   ::::::|||      :::
Db        850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   ||||||||||||||
Db        790 GTGACCGCCGGACCCTCCAGACCG 767

RESULT 27
US-09-940-316B-18/c
; Sequence 18, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
;   LENGTH: 4571
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
```

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (9)..(4559)
US-09-940-316B-18

Alignment Scores:
Pred. No.:                188             Length:         4571
Score:                    61.00           Matches:        16
Percent Similarity:       45.8%           Conservative:   6
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              26.6%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-18 (1-4571)

Qy          3 ProGlySerGluIleIleGluSerValProProAla----------------GlyProGlu 17
              |||       :::   |||          ||||||                    ||||||
Db        910 CCGCCGGGAGACGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 851

Qy         18 AlaSerGlu----------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                |||:::                |||   |||   ::::::|||      :::
Db        850 CGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 791

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   ||||||||||||||||
Db        790 GTGACCGCCGGACCCTCCAGACCG 767

RESULT 28
US-09-949-016-13937
; Sequence 13937, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13937
;   LENGTH: 13290
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13937

Alignment Scores:
Pred. No.:                722             Length:         13290
Score:                    61.00           Matches:        16
Percent Similarity:       60.5%           Conservative:   7
Best Local Similarity:    42.1%           Mismatches:     13
Query Match:              26.6%           Indels:         2
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-13937 (1-13290)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||       :::   |||   |||||||||      |||::::::::::|||
Db      12458 CCTGGTTCCCACTGCCTAGTTTCTGGGCCCCCGGCATCC---GAGTCGGCCAAAAGCCGG 12514

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                :::|||:::      |||   ||||||         |||   |||:::|||
Db      12515 GCAGAAGAGAGCGCCAAGGACTGGCAGTTC---CGGAACGGGCAATCCCAGCCG 12565

RESULT 29
US-09-949-016-13938
; Sequence 13938, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13938
;   LENGTH: 13290
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13938

Alignment Scores:
Pred. No.:                722             Length:         13290
Score:                    61.00           Matches:        16
Percent Similarity:       60.5%           Conservative:   7
Best Local Similarity:    42.1%           Mismatches:     13
Query Match:              26.6%           Indels:         2
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-13938 (1-13290)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||       :::   |||   |||||||||      |||::::::::::|||
Db      12458 CCTGGTTCCCACTGCCTAGTTTCTGGGCCCCCGGCATCC---GAGTCGGCCAAAAGCCGG 12514

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                :::|||:::      |||   ||||||         |||   |||:::|||
Db      12515 GCAGAAGAGAGCGCCAAGGACTGGCAGTTC---CGGAACGGGCAATCCCAGCCG 12565

RESULT 30
US-09-740-027-3/c
; Sequence 3, Application US/09740027
; Patent No. 6485939
; GENERAL INFORMATION:
;  APPLICANT: YE, Jane et al
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER COFACTOR
;  TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER
;  TITLE OF INVENTION:  COFACTOR PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL001052
;  CURRENT APPLICATION NUMBER: US/09/740,027
;  CURRENT FILING DATE:  2000-12-20
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 24707
;   TYPE: DNA
;   ORGANISM: Human
US-09-740-027-3

Alignment Scores:
Pred. No.:                1.58e+03        Length:         24707
Score:                    61.00           Matches:        18
Percent Similarity:       56.1%           Conservative:   5
Best Local Similarity:    43.9%           Mismatches:     18
Query Match:              26.6%           Indels:         1
DB:                       3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-740-027-3 (1-24707)
```

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||:::      :::   :::|||   ||||||||||||      |||   ||||||
Db      14331 CCTGGTGCAGGGCCACTGGCTGCTGTCCAGCCTGCTGGCCCTGTGTTATCACTCAGCACA 14272

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
              :::|||      :::   |||   |||           |||   |||   ||||||
Db      14271 TCAGACAATGAGCAGTG-GATAGTCAAGGCTCCCAGCCAGGTTCCAGCTCCCCTGTTTAG 14213

Qy         43 Glu 43
              |||
Db      14212 GAG 14210
```

RESULT 31
US-10-274-968-3/c
; Sequence 3, Application US/10274968
; Patent No. 6869928
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER COFACTOR
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER
; TITLE OF INVENTION: COFACTOR PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001052 DIV
; CURRENT APPLICATION NUMBER: US/10/274,968
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/740,027
; PRIOR FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 24707
; TYPE: DNA
; ORGANISM: Human
US-10-274-968-3

Alignment Scores:

Pred. No.:	1.58e+03	Length:	24707
Score:	61.00	Matches:	18
Percent Similarity:	56.1%	Conservative:	5
Best Local Similarity:	43.9%	Mismatches:	18
Query Match:	26.6%	Indels:	1
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-274-968-3 (1-24707)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||:::      :::   :::|||   ||||||||||||      |||   ||||||
Db      14331 CCTGGTGCAGGGCCACTGGCTGCTGTCCAGCCTGCTGGCCCTGTGTTATCACTCAGCACA 14272

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
              :::|||      :::   |||   |||           |||   |||   ||||||
Db      14271 TCAGACAATGAGCAGTG-GATAGTCAAGGCTCCCAGCCAGGTTCCAGCTCCCCTGTTTAG 14213

Qy         43 Glu 43
              |||
Db      14212 GAG 14210
```

RESULT 32
US-09-949-016-12341/c
; Sequence 12341, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12341
; LENGTH: 24720
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12341

Alignment Scores:

Pred. No.:	1.58e+03	Length:	24720
Score:	61.00	Matches:	18
Percent Similarity:	56.1%	Conservative:	5
Best Local Similarity:	43.9%	Mismatches:	18
Query Match:	26.6%	Indels:	1
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-12341 (1-24720)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||:::      :::   :::|||   ||||||||||||      |||   ||||||
Db      14343 CCTGGTGCAGGGCCACTGGCTGCTGTCCAGCCTGCTGGCCCTGTGTTATCACTCAGCACA 14284

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
              :::|||      :::   |||   |||           |||   |||   ||||||
Db      14283 TCAGACAATGAGCAGTG-GATAGTCAAGGCTCCCAGCCAGGTTCCAGCTCCCCTGTTTAG 14225

Qy         43 Glu 43
              |||
Db      14224 GAG 14222
```

RESULT 33
US-09-949-016-15610/c
; Sequence 15610, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15610
; LENGTH: 24721
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15610

Alignment Scores:

Pred. No.:	1.58e+03	Length:	24721
Score:	61.00	Matches:	18
Percent Similarity:	56.1%	Conservative:	5
Best Local Similarity:	43.9%	Mismatches:	18
Query Match:	26.6%	Indels:	1
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-15610 (1-24721)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||:::      :::   :::|||   ||||||||||||      |||   ||||||
Db      14343 CCTGGTGCAGGGCCACTGGCTGCTGTCCAGCCTGCTGGCCCTGTGTTATCACTCAGCACA 14284

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
              :::|||      :::   |||   |||           |||   |||   ||||||
Db      14283 TCAGACAATGAGCAGTG-GATAGTCAAGGCTCCCAGCCAGGTTCCAGCTCCCCTGTTTAG 14225
```

```
Qy          43 Glu 43
               |||
Db       14224 GAG 14222

RESULT 34
US-09-814-915A-80
; Sequence 80, Application US/09814915A
; Patent No. 6750015
; GENERAL INFORMATION:
;  APPLICANT: Horwitz, Kathryn
;  APPLICANT:  Richer, Jennifer
;  TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Relat
;  TITLE OF INVENTION:  Thereto
;  FILE REFERENCE: 2848-39
;  CURRENT APPLICATION NUMBER: US/09/814,915A
;  CURRENT FILING DATE:  2002-03-21
;  PRIOR APPLICATION NUMBER: 60/214,870
;  PRIOR FILING DATE: 2000-06-28
;  NUMBER OF SEQ ID NOS: 108
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
;   LENGTH: 9164
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-814-915A-80

Alignment Scores:
Pred. No.:                625             Length:         9164
Score:                    60.00           Matches:        20
Percent Similarity:       34.7%           Conservative:   6
Best Local Similarity:    26.7%           Mismatches:     17
Query Match:              26.2%           Indels:         32
DB:                       3               Gaps:           4

US-10-757-745-6_COPY_2_45 (1-44) x US-09-814-915A-80 (1-9164)

Qy           1 GlyAspProGlySerGluIleIle------GluSerValPro------------------ 12
               |||   ||||||      ::::::            ||||||
Db        8232 GGTCCCCCCGGAGAGAGAGTGGTGGGGGCTCCTGGGGTCCCTGGAGCTCCTGGCGAGAGA 8291

Qy          13 ---------------------ProAlaGlyProGluAlaSerGluSerThrThrAspGlu 25
                                    ||||||||||||        :::
Db        8292 GGGGAGCAGGGGCGGCCAGGGCCTGCCGGTCCTCGAGGCGAGAAGGGAGAAGCTGCACTG 8351

Qy          26 AsnGluAspAspIle---------------------------------------------- 30
                 ||||||||||||
Db        8352 ACGGAGGATGACATCCGGGGCTTTGTGCGCCAAGAGATGAGTCAGCACTGTGCCTGCCAG 8411

Qy          31 ---GlnPheValSerGluGlyProSerArgProValLeuGluTyr 44
                  ||||||::::::   |||   |||||||||:::      |||
Db        8412 GGCCAGTTCATCGCATCTGGA---TCACGACCCCTCCCTAGTTAT 8453

RESULT 35
US-09-523-487-9
; Sequence 9, Application US/09523487
; Patent No. 6232098
; GENERAL INFORMATION:
;  APPLICANT:  Conklin, Darrell C.
;  APPLICANT:  Foster, Donald C.
;  APPLICANT:  Gao, Zeren
;  TITLE OF INVENTION:  KUNITZ DOMAIN POLYPEPTIDE AND MATERIALS
;  TITLE OF INVENTION:  AND METHODS FOR MAKING IT
;  FILE REFERENCE:  98-22
;  CURRENT APPLICATION NUMBER: US/09/523,487
;  CURRENT FILING DATE:  2000-03-10
;  PRIOR APPLICATION NUMBER:  09/320,095
;  PRIOR FILING DATE:  1999-05-26
;  NUMBER OF SEQ ID NOS:  20
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 2383
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1674)
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2383)
;   OTHER INFORMATION: n = A,T,C or G
US-09-523-487-9

Alignment Scores:
Pred. No.:                135             Length:         2383
Score:                    59.50           Matches:        15
Percent Similarity:       56.2%           Conservative:   3
Best Local Similarity:    46.9%           Mismatches:     11
Query Match:              26.0%           Indels:         3
DB:                       3               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-523-487-9 (1-2383)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               |||   |||  |||:::  ||||||  |||        |||  ||||||
Db        1300 GGGATGTCAGGGGAAGAACTCAGTGAATCTACTCCAGAGCCTCAAAAAGAAATTTCTGAG 1359

Qy          21 Ser---------ThrThrAspGluAsnGluAspAsp 29
               |||         |||  |||::::::|||||||||
Db        1360 TCATTGAGTGTCACCAGAGACCAGGATGAAGATGAT 1395

RESULT 36
US-09-320-095-9
; Sequence 9, Application US/09320095
; Patent No. 6087473
; GENERAL INFORMATION:
;  APPLICANT: Conklin, Darrell C.
;  APPLICANT: Foster, Donald C.
;  APPLICANT: Gao, Zeren
;  TITLE OF INVENTION: KUNITZ DOMAIN POLYPEPTIDE AND MATERIALS
;  TITLE OF INVENTION: AND METHODS FOR MAKING IT
;  FILE REFERENCE: 98-22
;  CURRENT APPLICATION NUMBER: US/09/320,095
;  CURRENT FILING DATE: 1999-05-26
;  EARLIER APPLICATION NUMBER: US 60/087,032
;  EARLIER FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 2383
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1674)
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2383)
;   OTHER INFORMATION: n = A,T,C or G
US-09-320-095-9

Alignment Scores:
Pred. No.:                137             Length:         2409
Score:                    59.50           Matches:        15
Percent Similarity:       56.2%           Conservative:   3
Best Local Similarity:    46.9%           Mismatches:     11
Query Match:              26.0%           Indels:         3
DB:                       3               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-320-095-9 (1-2409)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               |||   |||  |||:::  ||||||  |||        |||  ||||||
Db        1300 GGGATGTCAGGGGAAGAACTCAGTGAATCTACTCCAGAGCCTCAAAAAGAAATTTCTGAG 1359

Qy          21 Ser---------ThrThrAspGluAsnGluAspAsp 29
               |||         |||  |||::::::|||||||||
Db        1360 TCATTGAGTGTCACCAGAGACCAGGATGAAGATGAT 1395
```

```
RESULT 37
US-09-949-016-12654/c
; Sequence 12654, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12654
;   LENGTH: 18836
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12654

Alignment Scores:
Pred. No.:              1.82e+03        Length:         18836
Score:                  59.50           Matches:        14
Percent Similarity:     50.0%           Conservative:   6
Best Local Similarity:  35.0%           Mismatches:     19
Query Match:            26.0%           Indels:         1
DB:                     3               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-12654 (1-18836)

Qy          2 AspPro---GlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   ||||||   :::|||   :::   ||||||      |||   |||   :::
Db       9717 GACCCCACAGGAAGCAGACTCATCCACACGAGCCCCCCACCCCAGCCCCCCGCAGGGGAT 9658

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              :::   ||||||   |||                 ::::::      |||      |||
Db       9657 GCGCACACGGACATGAACCCCCAGCGTACCTCCACCCTCACAGGGAGCCCACCCTACCCC 9598

RESULT 38
US-09-949-016-15633/c
; Sequence 15633, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15633
;   LENGTH: 18838
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15633

Alignment Scores:
Pred. No.:              1.82e+03        Length:         18838
Score:                  59.50           Matches:        14
Percent Similarity:     50.0%           Conservative:   6
Best Local Similarity:  35.0%           Mismatches:     19
Query Match:            26.0%           Indels:         1
DB:                     3               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-15633 (1-18838)

Qy          2 AspPro---GlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   ||||||   :::|||   :::   ||||||      |||   |||   :::
Db       9717 GACCCCACAGGAAGCAGACTCATCCACACGAGCCCCCCACCCCAGCCCCCCGCAGGGGAT 9658

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              :::   ||||||   |||                 ::::::      |||      |||
Db       9657 GCGCACACGGACATGAACCCCCAGCGTACCTCCACCCTCACAGGGAGCCCACCCTACCCC 9598

RESULT 39
US-09-949-016-13770
; Sequence 13770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13770
;   LENGTH: 36731
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13770

Alignment Scores:
Pred. No.:              4.22e+03        Length:         36731
Score:                  59.50           Matches:        16
Percent Similarity:     50.0%           Conservative:   5
Best Local Similarity:  38.1%           Mismatches:     12
Query Match:            26.0%           Indels:         9
DB:                     3               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-13770 (1-36731)

Qy         10 SerValProProAlaGlyProGluAlaSerGluSerThrThrAsp---------------- 24
              |||:::|||||||||||||   |||   |||   |||||||:::
Db      33513 TCGCTGCCTCCTGCAGGAAGGGAATTGTCAAGCTCAACATCTACTTCCAAGAATTTAACT 33572

Qy         25 ------------GluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                         :::|||:::      |||      |||   |||   |||:::   |||
Db      33573 ATCGCACCATTGAAGAATCAGCAGCCAATAACGTGAGTTTAGGAGTCTCCCAATACCCCA 33632

Qy         41 ValLeu 42
              |||
Db      33633 GCCCTG 33638

RESULT 40
US-09-702-705-372/c
; Sequence 372, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
```

; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 372
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-372

Alignment Scores:

Pred. No.:	13.4	Length:	335
Score:	59.00	Matches:	11
Percent Similarity:	70.6%	Conservative:	1
Best Local Similarity:	64.7%	Mismatches:	5
Query Match:	25.8%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-702-705-372 (1-335)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db        219 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 169
```

RESULT 41
US-09-736-457-372/c
; Sequence 372, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 372
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-372

Alignment Scores:

Pred. No.:	13.4	Length:	335
Score:	59.00	Matches:	11
Percent Similarity:	70.6%	Conservative:	1
Best Local Similarity:	64.7%	Mismatches:	5
Query Match:	25.8%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-736-457-372 (1-335)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db        219 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 169
```

RESULT 42
US-09-614-124B-372/c
; Sequence 372, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 372
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-372

Alignment Scores:

Pred. No.:	13.4	Length:	335
Score:	59.00	Matches:	11
Percent Similarity:	70.6%	Conservative:	1
Best Local Similarity:	64.7%	Mismatches:	5
Query Match:	25.8%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-614-124B-372 (1-335)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db        219 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 169
```

RESULT 43
US-09-671-325-372/c
; Sequence 372, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 372
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-372

Alignment Scores:

Pred. No.:	13.4	Length:	335
Score:	59.00	Matches:	11
Percent Similarity:	70.6%	Conservative:	1
Best Local Similarity:	64.7%	Mismatches:	5
Query Match:	25.8%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-671-325-372 (1-335)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              ||||||||||:::|||      |||   ||||||   |||||||||   |||
Db        219 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 169

RESULT 44
US-09-589-184-372/c
; Sequence 372, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C8
;  CURRENT APPLICATION NUMBER: US/09/589,184
;  CURRENT FILING DATE:  2000-06-05
;  NUMBER OF SEQ ID NOS:  827
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 372
;   LENGTH: 335
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-589-184-372

Alignment Scores:
Pred. No.:              13.4          Length:         335
Score:                  59.00         Matches:        11
Percent Similarity:     70.6%         Conservative:   1
Best Local Similarity:  64.7%         Mismatches:     5
Query Match:            25.8%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-589-184-372 (1-335)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              ||||||||||:::|||      |||   ||||||   |||||||||   |||
Db        219 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 169

RESULT 45
US-09-658-824-372/c
; Sequence 372, Application US/09658824
; Patent No. 6746846
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C11
;  CURRENT APPLICATION NUMBER: US/09/658,824
;  CURRENT FILING DATE:  2000-09-08
;  NUMBER OF SEQ ID NOS:  1788
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 372
;   LENGTH: 335
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-658-824-372

Alignment Scores:
Pred. No.:              13.4          Length:         335
Score:                  59.00         Matches:        11
Percent Similarity:     70.6%         Conservative:   1
Best Local Similarity:  64.7%         Mismatches:     5
Query Match:            25.8%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-658-824-372 (1-335)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              ||||||||||:::|||      |||   ||||||   |||||||||   |||
Db        219 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 169

RESULT 46
US-10-017-754-372/c
; Sequence 372, Application US/10017754
; Patent No. 6858204
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C18
;  CURRENT APPLICATION NUMBER: US/10/017,754
;  CURRENT FILING DATE:  2001-10-29
;  NUMBER OF SEQ ID NOS: 2004
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 372
;   LENGTH: 335
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-017-754-372

Alignment Scores:
Pred. No.:              13.4          Length:         335
Score:                  59.00         Matches:        11
Percent Similarity:     70.6%         Conservative:   1
Best Local Similarity:  64.7%         Mismatches:     5
Query Match:            25.8%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-017-754-372 (1-335)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              ||||||||||:::|||      |||   ||||||   |||||||||   |||
Db        219 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 169

RESULT 47
US-09-651-563-372/c
; Sequence 372, Application US/09651563
; Patent No. 6914132
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C10
```

```
;  CURRENT APPLICATION NUMBER: US/09/651,563
;  CURRENT FILING DATE:  2000-08-29
;  NUMBER OF SEQ ID NOS:  1679
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 372
;   LENGTH: 335
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-651-563-372

Alignment Scores:
Pred. No.:                13.4          Length:          335
Score:                    59.00         Matches:         11
Percent Similarity:       70.6%         Conservative:    1
Best Local Similarity:    64.7%         Mismatches:      5
Query Match:              25.8%         Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-651-563-372 (1-335)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db        219 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 169

RESULT 48
US-09-519-642-372/c
; Sequence 372, Application US/09519642
; Patent No. 6933363
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT: Bangur, Chaitanya S.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C4
;  CURRENT APPLICATION NUMBER: US/09/519,642
;  CURRENT FILING DATE: 2000-03-06
;  NUMBER OF SEQ ID NOS: 789
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 372
;   LENGTH: 335
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-519-642-372

Alignment Scores:
Pred. No.:                13.4          Length:          335
Score:                    59.00         Matches:         11
Percent Similarity:       70.6%         Conservative:    1
Best Local Similarity:    64.7%         Mismatches:      5
Query Match:              25.8%         Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-519-642-372 (1-335)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db        219 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 169

RESULT 49
US-09-513-999C-14793
; Sequence 14793, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 14793
;   LENGTH: 427
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-513-999C-14793

Alignment Scores:
Pred. No.:                18.2          Length:          427
Score:                    59.00         Matches:         11
Percent Similarity:       70.6%         Conservative:    1
Best Local Similarity:    64.7%         Mismatches:      5
Query Match:              25.8%         Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-513-999C-14793 (1-427)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db        186 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 236

RESULT 50
US-09-949-016-2252
; Sequence 2252, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2252
;   LENGTH: 543
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-2252

Alignment Scores:
Pred. No.:                24.6          Length:          543
Score:                    59.00         Matches:         11
Percent Similarity:       70.6%         Conservative:    1
Best Local Similarity:    64.7%         Mismatches:      5
Query Match:              25.8%         Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-2252 (1-543)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db        261 CCAGGATCCCAGATTGAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 311

RESULT 51
US-09-949-016-5158
; Sequence 5158, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
```

```
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5158
;   LENGTH: 543
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-5158

Alignment Scores:
Pred. No.:                24.6          Length:        543
Score:                    59.00         Matches:       11
Percent Similarity:       70.6%         Conservative:  1
Best Local Similarity:    64.7%         Mismatches:    5
Query Match:              25.8%         Indels:        0
DB:                       3             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-5158 (1-543)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              ||||||||||:::|||      |||   ||||||   |||||||||   |||
Db        261 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 311

RESULT 52
US-09-949-016-77086/c
; Sequence 77086, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 77086
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-77086

Alignment Scores:
Pred. No.:                28            Length:        601
Score:                    59.00         Matches:       11
Percent Similarity:       70.6%         Conservative:  1
Best Local Similarity:    64.7%         Mismatches:    5
Query Match:              25.8%         Indels:        0
DB:                       3             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-77086 (1-601)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db         63 CCAGGATCCCAGATTGAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 13

RESULT 53
US-09-854-133-627
; Sequence 627, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 627
;   LENGTH: 645
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-854-133-627

Alignment Scores:
Pred. No.:                30.6          Length:        645
Score:                    59.00         Matches:       15
Percent Similarity:       51.1%         Conservative:  8
Best Local Similarity:    33.3%         Mismatches:    16
Query Match:              25.8%         Indels:        6
DB:                       3             Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-854-133-627 (1-645)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              :::||||||:::|||      |||:::                  :::   ||||||:::
Db         65 GAACCAGGTGCTGAATCTTCTGAGAACGCTGATGATCCCAACAAAGATACAAGTGAAAAC 124

Qy         22 Thr---------ThrAspGluAsnGluAspAspIleGlnPheValSerGlu--------- 35
                         :::||||||||||:::||||||                |||
Db        125 GCAGATGGTCAAAGTGATGAGAACAAGGACGACTATACAATCCCAGATGAGTATAGAATT 184

Qy         36 GlyProSerArgPro 40
              ||||||   :::|||
Db        185 GGACCATATCAGCCC 199

RESULT 54
US-08-786-606-4
; Sequence 4, Application US/08786606
; Patent No. 5861495
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Coleman, Roger
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN ZINC-BINDING
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/786,606
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy RJ
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0173 US
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  944 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-786-606-4

Alignment Scores:
Pred. No.:              49.4          Length:         944
Score:                  59.00         Matches:        11
Percent Similarity:     70.6%         Conservative:   1
Best Local Similarity:  64.7%         Mismatches:     5
Query Match:            25.8%         Indels:         0
DB:                     2             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-08-786-606-4 (1-944)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db        358 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 408

RESULT 55
US-08-933-750C-97
; Sequence 97, Application US/08933750C
; Patent No. 5932442
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  HUMAN REGULATORY MOLECULES
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/933,750C
;      FILING DATE:  September 23, 1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0356 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  97:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1112 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  TESTNOT07
;      CLONE:  3217567
US-08-933-750C-97

Alignment Scores:
Pred. No.:              60.7          Length:         1112
Score:                  59.00         Matches:        11
Percent Similarity:     70.6%         Conservative:   1
Best Local Similarity:  64.7%         Mismatches:     5
Query Match:            25.8%         Indels:         0
DB:                     2             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-08-933-750C-97 (1-1112)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db        360 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 410

RESULT 56
US-09-234-613-97
; Sequence 97, Application US/09234613
; Patent No. 6132973
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  HUMAN REGULATORY MOLECULES
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/234,613
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/933,750
;      FILING DATE:  September 23, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0356 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  97:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1112 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  TESTNOT07
;      CLONE:  3217567
US-09-234-613-97
```

```
Alignment Scores:
Pred. No.:                60.7          Length:         1112
Score:                    59.00         Matches:        11
Percent Similarity:       70.6%         Conservative:   1
Best Local Similarity:    64.7%         Mismatches:     5
Query Match:              25.8%         Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-234-613-97 (1-1112)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||     |||   ||||||   |||||||||   |||
Db        360 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 410

RESULT 57
US-09-690-454-39
; Sequence 39, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
;  APPLICANT: Steven M. Ruben, et al.
;  TITLE OF INVENTION: 32 Human Secreted Proteins
;  FILE REFERENCE: PZ006P1
;  CURRENT APPLICATION NUMBER: US/09/690,454
;  CURRENT FILING DATE:  2000-10-18
;  PRIOR APPLICATION NUMBER: 09/189,144
;  PRIOR FILING DATE: 1998-11-10
;  PRIOR APPLICATION NUMBER: 60/044,039
;  PRIOR FILING DATE: May 30, 1997
;  PRIOR APPLICATION NUMBER: 60/048,093
;  PRIOR FILING DATE: May 30, 1997
;  PRIOR APPLICATION NUMBER: 60/048,190
;  PRIOR FILING DATE: May 30, 1997
;  PRIOR APPLICATION NUMBER: 60/050,935
;  PRIOR FILING DATE: May 30, 1997
;  PRIOR APPLICATION NUMBER: 60/048,101
;  PRIOR FILING DATE: May 30, 1997
;  PRIOR APPLICATION NUMBER: 60/048,356
;  PRIOR FILING DATE: May 30, 1997
;  PRIOR APPLICATION NUMBER: 60/056,250
;  PRIOR FILING DATE: August 29, 1997
;  PRIOR APPLICATION NUMBER: 60/056,296
;  PRIOR FILING DATE: August 29, 1997
;  PRIOR APPLICATION NUMBER: 60/056,293
;  PRIOR FILING DATE: August 29, 1997
;  NUMBER OF SEQ ID NOS: 229
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 1114
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-690-454-39

Alignment Scores:
Pred. No.:                60.9          Length:         1114
Score:                    59.00         Matches:        11
Percent Similarity:       70.6%         Conservative:   1
Best Local Similarity:    64.7%         Mismatches:     5
Query Match:              25.8%         Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-690-454-39 (1-1114)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||     |||   ||||||   |||||||||   |||
Db        309 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 359

RESULT 58
US-09-902-540-64
; Sequence 64, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 64
;   LENGTH: 1329
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-64

Alignment Scores:
Pred. No.:                76            Length:         1329
Score:                    59.00         Matches:        15
Percent Similarity:       48.2%         Conservative:   12
Best Local Similarity:    26.8%         Mismatches:     13
Query Match:              25.8%         Indels:         16
DB:                       3             Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-902-540-64 (1-1329)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyPro---------- 16
              |||      :::::::|||||||||   :::|||      ||||||
Db        651 GATCAGCAGGCCCAGATCATAGAAGGTCTGCCA------GGACCCATCGCCGTAGAAGTT 704

Qy         17 -------------------------GluAlaSerGluSerThrThrAspGluAsnGlu 27
                                       |||::::::::::      |||   :::|||
Db        705 CCGGAAGCCGCCGGCCACCGATAGATCGGGGCCGCGCAGGCTGCCGCGGATCACGACGAA 764

Qy         28 AspAspIleGlnPheValSerGluGlyProSerArgProValLeuGlu 43
              |||  :::     |||:::   ||||||:::   |||   ::::::
Db        765 GACCTCGTCCCCGTCGTAGCCCACGGTCCGGCTGCCCCCGAGGTCCAG 812

RESULT 59
US-09-976-594-400
; Sequence 400, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 400
;   LENGTH: 3172
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 3676741CB1
US-09-976-594-400

Alignment Scores:
Pred. No.:                227           Length:         3172
Score:                    59.00         Matches:        15
Percent Similarity:       51.1%         Conservative:   8
Best Local Similarity:    33.3%         Mismatches:     16
Query Match:              25.8%         Indels:         6
DB:                       3             Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-976-594-400 (1-3172)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              :::||||||:::|||     |||:::                  :::   |||||||:::
```

```
Db       2482 GAACCAGGTGCTGAATCTTCTGAGAACGCTGATGATCCCAACAAAGATACAAGTGAAAAC 2541

Qy         22 Thr---------ThrAspGluAsnGluAspAspIleGlnPheValSerGlu--------- 35
                          :::||||||||||:::||||||                 |||
Db       2542 GCAGATGGTCAAAGTGATGAGAACAAGGACGACTATACAATCCCAGATGAGTATAGAATT 2601

Qy         36 GlyProSerArgPro 40
              ||||||   :::|||
Db       2602 GGACCATATCAGCCC 2616

RESULT 60
US-09-949-016-5552
; Sequence 5552, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5552
;   LENGTH: 3817
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-5552

Alignment Scores:
Pred. No.:                287             Length:         3817
Score:                    59.00           Matches:        15
Percent Similarity:       51.1%           Conservative:   8
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              25.8%           Indels:         6
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-5552 (1-3817)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              :::||||||:::|||      |||:::                 :::   ||||||:::
Db       2478 GAACCAGGTGCTGAATCTTCTGAGAACGCTGATGATCCCAACAAAGATACAAGTGAAAAC 2537

Qy         22 Thr---------ThrAspGluAsnGluAspAspIleGlnPheValSerGlu--------- 35
                          :::||||||||||:::||||||                 |||
Db       2538 GCAGATGGTCAAAGTGATGAGAACAAGGACGACTATACAATCCCAGATGAGTATAGAATT 2597

Qy         36 GlyProSerArgPro 40
              ||||||   :::|||
Db       2598 GGACCATATCAGCCC 2612

RESULT 61
US-09-949-016-1067
; Sequence 1067, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1067
;   LENGTH: 3835
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-1067

Alignment Scores:
Pred. No.:                289             Length:         3835
Score:                    59.00           Matches:        15
Percent Similarity:       51.1%           Conservative:   8
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              25.8%           Indels:         6
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-1067 (1-3835)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              :::||||||:::|||      |||:::                 :::   ||||||:::
Db       2481 GAACCAGGTGCTGAATCTTCTGAGAACGCTGATGATCCCAACAAAGATACAAGTGAAAAC 2540

Qy         22 Thr---------ThrAspGluAsnGluAspAspIleGlnPheValSerGlu--------- 35
                          :::||||||||||:::||||||                 |||
Db       2541 GCAGATGGTCAAAGTGATGAGAACAAGGACGACTATACAATCCCAGATGAGTATAGAATT 2600

Qy         36 GlyProSerArgPro 40
              ||||||   :::|||
Db       2601 GGACCATATCAGCCC 2615

RESULT 62
US-09-949-016-13994
; Sequence 13994, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13994
;   LENGTH: 4543
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13994

Alignment Scores:
Pred. No.:                357             Length:         4543
Score:                    59.00           Matches:        11
Percent Similarity:       70.6%           Conservative:   1
Best Local Similarity:    64.7%           Mismatches:     5
Query Match:              25.8%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-13994 (1-4543)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db       2261 CCAGGATCCCAGATTGAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 2311

RESULT 63
US-09-410-551B-26/c
; Sequence 26, Application US/09410551B
```

```
; Patent No. 6503737
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
;   LENGTH: 4674
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4673)
US-09-410-551B-26

Alignment Scores:
Pred. No.:                370             Length:         4674
Score:                    59.00           Matches:        15
Percent Similarity:       44.9%           Conservative:   7
Best Local Similarity:    30.6%           Mismatches:     17
Query Match:              25.8%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-26 (1-4674)

Qy           2 AspProGlySerGluIleIleGluSerValProProAla--------------GlyPro 16
               :::|||      :::   |||         ||||||                   |||
Db        1078 AATCCGCCGGGCGACGCCATCACCGTGACACCGCCGACCAGGGCGAGCGAGCATTCGCCC 1019

Qy          17 GluAlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGln 31
               |||      :::                |||   ::::::|||      :::
Db        1018 GAGCGCAGGGACTGCCCTGCCTGGTGCAGGGCGACCAGTGACGACGAGCAGGCGGTGTCG 959

Qy          32 PheValSerGluGlyProSerArgPro 40
                 |||:::|||||||||||||||||||
Db         958 ACCGTGACCGAAGGGCCCTCCAGACCG 932

RESULT 64
US-09-940-316B-26/c
; Sequence 26, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
;   LENGTH: 4674
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4673)
US-09-940-316B-26

Alignment Scores:
Pred. No.:                370             Length:         4674
Score:                    59.00           Matches:        15
Percent Similarity:       44.9%           Conservative:   7
Best Local Similarity:    30.6%           Mismatches:     17
Query Match:              25.8%           Indels:         10
DB:                       3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-26 (1-4674)

Qy           2 AspProGlySerGluIleIleGluSerValProProAla--------------GlyPro 16
               :::|||      :::   |||         ||||||                   |||
Db        1078 AATCCGCCGGGCGACGCCATCACCGTGACACCGCCGACCAGGGCGAGCGAGCATTCGCCC 1019

Qy          17 GluAlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGln 31
               |||      :::                |||   ::::::|||      :::
Db        1018 GAGCGCAGGGACTGCCCTGCCTGGTGCAGGGCGACCAGTGACGACGAGCAGGCGGTGTCG 959

Qy          32 PheValSerGluGlyProSerArgPro 40
                 |||:::|||||||||||||||||||
Db         958 ACCGTGACCGAAGGGCCCTCCAGACCG 932

RESULT 65
US-09-410-551B-24/c
; Sequence 24, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
;   LENGTH: 4725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4724)
US-09-410-551B-24

Alignment Scores:
Pred. No.:                375           Length:         4725
Score:                    59.00         Matches:        15
Percent Similarity:       44.9%         Conservative:   7
Best Local Similarity:    30.6%         Mismatches:     17
Query Match:              25.8%         Indels:         10
DB:                       3             Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-24 (1-4725)

Qy          2 AspProGlySerGluIleIleGluSerValProProAla---------------GlyPro 16
              :::|||      :::   |||         ||||||                    |||
Db       1078 AATCCGCCGGGCGACGCCATCACCGTGACACCGCCGACCAGGGCGAGCGAGCATTCGCCC 1019

Qy         17 GluAlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||      :::                    |||   ::::::|||      :::
Db       1018 GAGCGCAGGGACTGCCCTGCCTGGTGCAGGGCGACCAGTGACGACGAGCAGGCGGTGTCG 959

Qy         32 PheValSerGluGlyProSerArgPro 40
              |||:::|||||||||||||||||||
Db        958 ACCGTGACCGAAGGGCCCTCCAGACCG 932

RESULT 66
US-09-940-316B-24/c
; Sequence 24, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
;   LENGTH: 4725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4724)
US-09-940-316B-24

Alignment Scores:
Pred. No.:                375           Length:         4725
Score:                    59.00         Matches:        15
Percent Similarity:       44.9%         Conservative:   7
Best Local Similarity:    30.6%         Mismatches:     17
Query Match:              25.8%         Indels:         10
DB:                       3             Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-24 (1-4725)

Qy          2 AspProGlySerGluIleIleGluSerValProProAla---------------GlyPro 16
              :::|||      :::   |||         ||||||                    |||
Db       1078 AATCCGCCGGGCGACGCCATCACCGTGACACCGCCGACCAGGGCGAGCGAGCATTCGCCC 1019

Qy         17 GluAlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||      :::                    |||   ::::::|||      :::
Db       1018 GAGCGCAGGGACTGCCCTGCCTGGTGCAGGGCGACCAGTGACGACGAGCAGGCGGTGTCG 959

Qy         32 PheValSerGluGlyProSerArgPro 40
              |||:::|||||||||||||||||||
Db        958 ACCGTGACCGAAGGGCCCTCCAGACCG 932

RESULT 67
US-09-410-551B-30/c
; Sequence 30, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
;   LENGTH: 4737
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4736)
US-09-410-551B-30

Alignment Scores:
Pred. No.:                377           Length:         4737
Score:                    59.00         Matches:        15
Percent Similarity:       44.9%         Conservative:   7
Best Local Similarity:    30.6%         Mismatches:     17
Query Match:              25.8%         Indels:         10
DB:                       3             Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-30 (1-4737)

Qy          2 AspProGlySerGluIleIleGluSerValProProAla---------------GlyPro 16
              :::|||      :::   |||         ||||||                    |||
Db       1078 AATCCGCCGGGCGACGCCATCACCGTGACACCGCCGACCAGGGCGAGCGAGCATTCGCCC 1019

Qy         17 GluAlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||      :::                    |||   ::::::|||      :::
Db       1018 GAGCGCAGGGACTGCCCTGCCTGGTGCAGGGCGACCAGTGACGACGAGCAGGCGGTGTCG 959

Qy         32 PheValSerGluGlyProSerArgPro 40
              |||:::|||||||||||||||||||
Db        958 ACCGTGACCGAAGGGCCCTCCAGACCG 932

RESULT 68
```

```
US-09-940-316B-30/c
; Sequence 30, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
;   LENGTH: 4737
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4736)
US-09-940-316B-30

Alignment Scores:
Pred. No.:                377             Length:          4737
Score:                    59.00           Matches:         15
Percent Similarity:       44.9%           Conservative:    7
Best Local Similarity:    30.6%           Mismatches:      17
Query Match:              25.8%           Indels:          10
DB:                       3               Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-30 (1-4737)

Qy          2 AspProGlySerGluIleIleGluSerValProProAla--------------GlyPro 16
              :::|||      :::   |||        ||||||                   |||
Db       1078 AATCCGCCGGGCGACGCCATCACCGTGACACCGCCGACCAGGGCGAGCGAGCATTCGCCC 1019

Qy         17 GluAlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||      :::                |||   ::::::|||     :::
Db       1018 GAGCGCAGGGACTGCCCTGCCTGGTGCAGGGCGACCAGTGACGACGAGCAGGCGGTGTCG 959

Qy         32 PheValSerGluGlyProSerArgPro 40
                |||:::|||||||||||||||||||
Db        958 ACCGTGACCGAAGGGCCCTCCAGACCG 932

RESULT 69
US-09-410-551B-28/c
; Sequence 28, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
;   LENGTH: 4767
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4766)
US-09-410-551B-28

Alignment Scores:
Pred. No.:                380             Length:          4767
Score:                    59.00           Matches:         15
Percent Similarity:       44.9%           Conservative:    7
Best Local Similarity:    30.6%           Mismatches:      17
Query Match:              25.8%           Indels:          10
DB:                       3               Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-28 (1-4767)

Qy          2 AspProGlySerGluIleIleGluSerValProProAla--------------GlyPro 16
              :::|||      :::   |||        ||||||                   |||
Db       1078 AATCCGCCGGGCGACGCCATCACCGTGACACCGCCGACCAGGGCGAGCGAGCATTCGCCC 1019

Qy         17 GluAlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||      :::                |||   ::::::|||     :::
Db       1018 GAGCGCAGGGACTGCCCTGCCTGGTGCAGGGCGACCAGTGACGACGAGCAGGCGGTGTCG 959

Qy         32 PheValSerGluGlyProSerArgPro 40
                |||:::|||||||||||||||||||
Db        958 ACCGTGACCGAAGGGCCCTCCAGACCG 932

RESULT 70
US-09-940-316B-28/c
; Sequence 28, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNTH
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
;   LENGTH: 4767
;   TYPE: DNA
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4766)
US-09-940-316B-28

Alignment Scores:
Pred. No.:                380          Length:         4767
Score:                    59.00        Matches:        15
Percent Similarity:       44.9%        Conservative:   7
Best Local Similarity:    30.6%        Mismatches:     17
Query Match:              25.8%        Indels:         10
DB:                       3            Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-28 (1-4767)

Qy          2 AspProGlySerGluIleIleGluSerValProProAla--------------GlyPro 16
              :::|||      :::   |||         ||||||                    |||
Db       1078 AATCCGCCGGGCGACGCCATCACCGTGACACCGCCGACCAGGGCGAGCGAGCATTCGCCC 1019

Qy         17 GluAlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||      :::                  |||   :::::::|||      :::
Db       1018 GAGCGCAGGGACTGCCCTGCCTGGTGCAGGGCGACCAGTGACGACGAGCAGGCGGTGTCG 959

Qy         32 PheValSerGluGlyProSerArgPro 40
                 |||:::|||||||||||||||||||
Db        958 ACCGTGACCGAAGGGCCCTCCAGACCG 932

RESULT 71
US-09-410-551B-32/c
; Sequence 32, Application US/09410551B
; Patent No. 6503737
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
;  TITLE OF INVENTION:  CONSTRUCTS THEREFOR
;  FILE REFERENCE: 30062-20026.00
;  CURRENT APPLICATION NUMBER: US/09/410,551B
;  CURRENT FILING DATE:  1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32
;   LENGTH: 4818
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4817)
US-09-410-551B-32

Alignment Scores:
Pred. No.:                385          Length:         4818
Score:                    59.00        Matches:        15
Percent Similarity:       44.9%        Conservative:   7
Best Local Similarity:    30.6%        Mismatches:     17
Query Match:              25.8%        Indels:         10
DB:                       3            Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-410-551B-32 (1-4818)

Qy          2 AspProGlySerGluIleIleGluSerValProProAla--------------GlyPro 16
              :::|||      :::   |||         ||||||                    |||
Db       1078 AATCCGCCGGGCGACGCCATCACCGTGACACCGCCGACCAGGGCGAGCGAGCATTCGCCC 1019

Qy         17 GluAlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||      :::                  |||   :::::::|||      :::
Db       1018 GAGCGCAGGGACTGCCCTGCCTGGTGCAGGGCGACCAGTGACGACGAGCAGGCGGTGTCG 959

Qy         32 PheValSerGluGlyProSerArgPro 40
                 |||:::|||||||||||||||||||
Db        958 ACCGTGACCGAAGGGCCCTCCAGACCG 932

RESULT 72
US-09-940-316B-32/c
; Sequence 32, Application US/09940316B
; Patent No. 6759536
; GENERAL INFORMATION:
;  APPLICANT: KOSAN BIOSCIENCES, Inc.
;  APPLICANT:  REEVES, CHRISTOPHER
;  APPLICANT:  CHU, DANIEL
;  APPLICANT:  KHOSLA, CHAITAN
;  APPLICANT:  SANTI, DANIEL
;  APPLICANT:  WU, KAI
;  TITLE OF INVENTION: POLYKETIDES ENCODING THE fkbA GENE OF THE FK-520 POLYKETIDE SYNT
;  TITLE OF INVENTION:  GENE CLUSTER
;  FILE REFERENCE: 30062-20026.11
;  CURRENT APPLICATION NUMBER: US/09/940,316B
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/410,551
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: US 60/139,650
;  PRIOR FILING DATE: 1999-06-17
;  PRIOR APPLICATION NUMBER: US 60/123,810
;  PRIOR FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: US 60/102,748
;  PRIOR FILING DATE: 1998-10-02
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32
;   LENGTH: 4818
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
;   OTHER INFORMATION: PKS synthase fragment
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(4817)
US-09-940-316B-32

Alignment Scores:
Pred. No.:                385          Length:         4818
Score:                    59.00        Matches:        15
Percent Similarity:       44.9%        Conservative:   7
Best Local Similarity:    30.6%        Mismatches:     17
Query Match:              25.8%        Indels:         10
DB:                       3            Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-940-316B-32 (1-4818)

Qy          2 AspProGlySerGluIleIleGluSerValProProAla--------------GlyPro 16
              :::|||      :::   |||         ||||||                    |||
Db       1078 AATCCGCCGGGCGACGCCATCACCGTGACACCGCCGACCAGGGCGAGCGAGCATTCGCCC 1019

Qy         17 GluAlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||      :::                  |||   :::::::|||      :::
Db       1018 GAGCGCAGGGACTGCCCTGCCTGGTGCAGGGCGACCAGTGACGACGAGCAGGCGGTGTCG 959

Qy         32 PheValSerGluGlyProSerArgPro 40
```

```
                  |||:::|||||||||||||||||||
Db        958 ACCGTGACCGAAGGGCCCTCCAGACCG 932

RESULT 73
US-09-949-002-28
; Sequence 28, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
;   LENGTH: 7137
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-28

Alignment Scores:
Pred. No.:              631          Length:         7137
Score:                  59.00        Matches:        14
Percent Similarity:     52.8%        Conservative:   5
Best Local Similarity:  38.9%        Mismatches:     15
Query Match:            25.8%        Indels:         2
DB:                     3            Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-002-28 (1-7137)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||      |||:::   :::  ||||||  |||  |||      |||||||||
Db       2320 CCAGGG------ATCCTTCCTACTTGGCCTCCTACTGGCGCCGAAACAGAGGAAAGTACA 2373

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSer 38
                :::     ::::::         ||||||  ||||||
Db       2374 GAAGGCCCTTCTGCAACTGAAGTGCCCTCTGCCTCAGAGGAACCATCC 2421

RESULT 74
US-09-949-002-112
; Sequence 112, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 112
;   LENGTH: 7434
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-112

Alignment Scores:
Pred. No.:              664          Length:         7434
Score:                  59.00        Matches:        14
Percent Similarity:     52.8%        Conservative:   5
Best Local Similarity:  38.9%        Mismatches:     15
Query Match:            25.8%        Indels:         2
DB:                     3            Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-002-112 (1-7434)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||      |||:::   :::  ||||||  |||  |||      |||||||||
Db       2320 CCAGGG------ATCCTTCCTACTTGGCCTCCTACTGGCGCCGAAACAGAGGAAAGTACA 2373

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSer 38
                :::     ::::::         ||||||  ||||||
Db       2374 GAAGGCCCTTCTGCAACTGAAGTGCCCTCTGCCTCAGAGGAACCATCC 2421

RESULT 75
US-09-949-016-12809
; Sequence 12809, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12809
;   LENGTH: 27236
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12809

Alignment Scores:
Pred. No.:              3.41e+03     Length:         27236
Score:                  59.00        Matches:        15
Percent Similarity:     51.1%        Conservative:   8
Best Local Similarity:  33.3%        Mismatches:     16
Query Match:            25.8%        Indels:         6
DB:                     3            Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-949-016-12809 (1-27236)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              :::||||||:::|||      |||:::                  :::   |||||:::
Db      20342 GAACCAGGTGCTGAATCTTCTGAGAACGCTGATGATCCCAACAAAGATACAAGTGAAAAC 20401

Qy         22 Thr---------ThrAspGluAsnGluAspAspIleGlnPheValSerGlu--------- 35
                  :::|||||||||:::||||||                  |||
Db      20402 GCAGATGGTCAAAGTGATGAGAACAAGGACGACTATACAATCCCAGATGAGTATAGAATT 20461

Qy         36 GlyProSerArgPro 40
              ||||||   :::|||
Db      20462 GGACCATATCAGCCC 20476


Search completed: March 10, 2006, 23:21:50
Job time : 69.038 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:38:30 ; Search time 170.261 Seconds
(without alignments)
1722.336 Million cell updates/sec

Title: US-10-757-745-6_COPY_2_45
Perfect score: 229
Sequence: 1 GDPGSEIIESVPPAGPEASE.....ENEDDIQFVSEGPSRPVLEY 44

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170325_3993/app_query.fasta_1
-DB=N_Geneseq -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1589_@runat_09032006_170325_3993 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	229	100.0	274	3	AAC01653	Aac01653 Human sec
2	229	100.0	632	4	AAI92516	Aai92516 Human pol
3	229	100.0	632	10	ADC32029	Adc32029 Human nov
4	229	100.0	1536	3	AAZ47120	Aaz47120 Human CD4
5	223	97.4	3042	14	ADV43036	Adv43036 Human psy
6	223	97.4	3183	6	ABK12397	Abk12397 cDNA #2 e
7	223	97.4	4961	13	ADR52788	Adr52788 Drug ther
8	223	97.4	4999	6	ABK12391	Abk12391 cDNA enco
9	223	97.4	4999	6	ABK12392	Abk12392 cDNA enco
10	223	97.4	4999	14	ADX07795	Adx07795 Cyclin-de
11	223	97.4	5143	6	ABK12407	Abk12407 cDNA #2 e
12	223	97.4	5144	6	ABK12394	Abk12394 cDNA #1 e
13	223	97.4	5144	6	ABK12393	Abk12393 cDNA #1 e
14	223	97.4	5245	10	ADC32216	Adc32216 Human nov
15	217	94.8	3363	5	AAS81317	Aas81317 DNA encod
16	216	94.3	2949	10	ADC30114	Adc30114 Human nov
17	215	93.9	5089	10	ADC30333	Adc30333 Human nov
18	175	76.4	394	5	AAS81316	Aas81316 DNA encod
19	175	76.4	394	5	AAS79405	Aas79405 DNA encod
20	175	76.4	448	9	ACH40820	Ach40820 Human foe
21	175	76.4	2039	5	AAS81318	Aas81318 DNA encod
c 22	89	38.9	121	4	AAI91954	Aai91954 Human pol
23	85.5	37.3	4629	13	ADU05801	Adu05801 Novel bro
c 24	75.5	33.0	2688	13	ADS57834	Ads57834 Bacterial
c 25	71	31.0	624	13	ADQ51222	Adq51222 Novel can
26	68	29.7	430	9	ACH18197	Ach18197 Human adu
27	68	29.7	493	9	ACH37587	Ach37587 Human end
28	68	29.7	493	9	ACH44839	Ach44839 Human foe
29	68	29.7	500	9	ACH43217	Ach43217 Human foe
30	68	29.7	2438	6	ABQ54916	Abq54916 Human ova
31	68	29.7	3713	10	ADC99117	Adc99117 Human KPP
c 32	68	29.7	3818	10	ADE09676	Ade09676 Novel DNA
c 33	68	29.7	3818	12	ADM87570	Adm87570 Human EST
34	67	29.3	7399	8	ACC45139	Acc45139 Alphaviru
35	67	29.3	8010	2	AAX78129	Aax78129 Semliki F
36	67	29.3	8100	2	AAX76588	Aax76588 SFV EcoRV
37	67	29.3	11060	2	AAX78130	Aax78130 pSFV link
38	67	29.3	11489	6	ABN86687	Abn86687 Nucleotid
39	67	29.3	11489	14	AEA40812	Aea40812 Anti-apop
40	67	29.3	11517	2	AAQ26021	Aaq26021 pSP6-SFV4
41	67	29.3	12347	14	AEA40832	Aea40832 Anti-apop
42	67	29.3	12464	3	AAC55473	Aac55473 Destinati
43	67	29.3	12612	14	AEA40833	Aea40833 Anti-apop
44	67	29.3	12837	10	ADL06634	Adl06634 SFV DNA c
45	67	29.3	12855	10	ADL06635	Adl06635 SFV DNA c
46	67	29.3	13599	6	ABN86690	Abn86690 Nucleotid
47	67	29.3	13599	14	AEA40871	Aea40871 Anti-apop
48	67	29.3	15538	2	AAX76582	Aax76582 Plasmid p
c 49	67	29.3	77536	3	AAA14651	Aaa14651 Nucleotid
50	65	28.4	2436	10	ADB69911	Adb69911 C. neofor
51	65	28.4	3346	10	ADB69550	Adb69550 C. neofor
52	65	28.4	5346	10	ADB69189	Adb69189 C. neofor
53	65	28.4	77536	3	AAA14651	Aaa14651 Nucleotid
54	64.5	28.2	2225	10	ADE55089	Ade55089 Rat gene
55	64	27.9	7390	12	ADI61727	Adi61727 Human cDN
56	64	27.9	7390	14	AEA43892	Aea43892 Human cDN
57	64	27.9	7666	6	ABQ93451	Abq93451 Human cDN
58	64	27.9	7758	6	ABK83454	Abk83454 Human cDN
59	64	27.9	7758	10	ADH29027	Adh29027 Human chr
60	64	27.9	7758	11	ADN95314	Adn95314 Human BEC
61	64	27.9	7758	12	ADQ18420	Adq18420 Human sof
62	64	27.9	7792	12	ADQ22923	Adq22923 Human sof
c 63	64	27.9	58857	3	AAA58471	Aaa58471 Nucleotid
c 64	64	27.9	104096	13	ADX56092	Adx56092 Streptomy
65	64	27.9	171096	13	ADT05874	Adt05874 Cercopith
66	63.5	27.7	132544	13	ADV99887	Adv99887 Nanchangm
67	63	27.5	3120	12	ADM87098	Adm87098 Human pro
68	63	27.5	3211	2	AAT89346	Aat89346 Human p16
69	63	27.5	3901	2	AAT89345	Aat89345 Human p16
c 70	63	27.5	9795	14	ADY72595	Ady72595 A. orient
c 71	63	27.5	73599	14	ADY72575	Ady72575 Polyene p
c 72	62	27.1	557	6	ABL82056	Abl82056 Human ova
c 73	62	27.1	574	14	ACL61290	Acl61290 Human col
74	62	27.1	762	14	ADX83075	Adx83075 DNA encod
75	62	27.1	1056	6	ABS65666	Abs65666 Mouse JNK

ALIGNMENTS

```
RESULT 1
AAC01653
ID   AAC01653 standard; cDNA; 274 BP.
XX
AC   AAC01653;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 1651.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   P-PSDB; AAG01647.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 1651; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. An ORF has been identified within the
CC   sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC   derived from 30 different tissues. EST sequences usually correspond
CC   mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC   often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC   well suited for isolating cDNA sequences derived from the 5' ends of
CC   mRNAs and even in those cases where longer cDNA sequences have been
CC   obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC   mRNAs with intact 5' ends and can therefore be used to obtain full length
CC   cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC   gene therapy and chromosome mapping procedures. They are used to obtain
CC   upstream regulatory sequences and to design expression and secretion
CC   vectors
XX
SQ   Sequence 274 BP; 62 A; 60 C; 96 G; 48 T; 0 U; 8 Other;

Alignment Scores:
Pred. No.:                5.21e-22        Length:         274
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AAC01653 (1-274)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GGAGACCCGGGGTCGGARATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 170

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 230

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        231 GTTCTTGAATAC 242

RESULT 2
AAI92516
ID   AAI92516 standard; cDNA; 632 BP.
XX
AC   AAI92516;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 12576.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US004927.
XX
PR   28-FEB-2000; 2000US-00515126.
PR   18-MAY-2000; 2000US-00577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   P-PSDB; AAO12585.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT   and treating e.g. leukemia, inflammation and immune disorders.
XX
PS   Claim 1; SEQ ID NO 12576; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 632 BP; 172 A; 132 C; 177 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.43e-21        Length:         632
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       4               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AAI92516 (1-632)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 278

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 338

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        339 GTTCTTGAATAC 350

RESULT 3
ADC32029
ID   ADC32029 standard; cDNA; 632 BP.
XX
AC   ADC32029;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human novel cDNA contig sequence, SEQ ID NO:2111.
XX
KW   Human; diagnostic; drug screening; forensics; gene mapping;
KW   biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW   neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW   ulcers; osteoporosis; autoimmune disease; cancer;
KW   molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW   neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW   antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW   gene therapy; chromosome 6; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003029271-A2.
XX
PD   10-APR-2003.
XX
PF   24-SEP-2002; 2002WO-US030474.
XX
PR   24-SEP-2001; 2001US-0324631P.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang TY,  Zhang J,  Ren F,  Xue AJ,  Zhao QA,  Wang J,  Wehrman T;
PI   Zhou P,  Ghosh M,  Wang D,  Ma Y,  Asundi V,  Wang Z,  Weng G;
PI   Haley-Vicente D,  Drmanac RT;
XX
DR   WPI; 2003-371981/35.
DR   P-PSDB; ADC32796.
XX
PT   New polynucleotide and polypeptide useful for diagnosing, preventing or
PT   treating conditions such as neurodegenerative diseases, anemias, platelet
PT   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT   cancer.
XX
PS   Example 2; SEQ ID NO 2111; 1185pp; English.
XX
CC   The invention relates to 971 novel human cDNA sequences (ADC29919-
CC   ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC   invention also relates to nucleic acid sequences over 99% identical with
CC   the novel human cDNAs. The invention additionally encompasses expression
CC   vectors and host cells comprising a nucleic acid of the invention; the
CC   recombinant production of a polypeptide of the invention; an antibody
CC   against a polypeptide of the invention; a method of detecting
CC   polynucleotides or polypeptides of the invention; and methods of
CC   identifying a compound which binds to a polypeptide of the invention. The
CC   invention further discloses methods of peventing, treating or
CC   ameliorating a medical condition; kits comprising polynucleotide probes
CC   and/or monoclonal antibodies for carrying out the methods of the
CC   invention; methods for the identification of compounds that modulate the
CC   expression or activity of the polynucleotide and/or polypeptide; and 767
CC   contig sequences corresponding to the cDNA sequences of the invention
CC   (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC   -ADC33394). The nucleic acids and polypeptides of the invention are
CC   useful in diagnostics, drug screening, forensics, gene mapping, in the
CC   identification of mutations responsible for genetic disorders or other
CC   traits, for assessing biodiversity, and in producing many other types of
CC   data and products dependent on DNA and amino acid sequences. They are
CC   also used for treating diseases such as Parkinson's disease, Alzheimer's
CC   disease and other neurodegenerative diseases, anaemia, platelet
CC   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC   cancer. The nucleic acids may also be used as hybridisation probes or
CC   primers, and in the recombinant production of a protein. The polypeptides
CC   are also useful in generating antibodies, as molecular weight markers,
CC   and as food supplements. The present sequence represents a human contig
CC   sequence used in an example of the invention. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 632 BP; 172 A; 132 C; 177 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.43e-21        Length:          632
Score:                    229.00          Matches:         44
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       10              Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ADC32029 (1-632)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 278

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 338

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        339 GTTCTTGAATAC 350

RESULT 4
AAZ47120
ID   AAZ47120 standard; cDNA; 1536 BP.
XX
AC   AAZ47120;
XX
DT   15-MAR-2000  (first entry)
XX
DE   Human CD40 receptor interacting protein 4C4 gene.
XX
KW   Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW   immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW   human; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW   tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW   TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW   diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW   arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW   graft versus host disease; autoimmune disease; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             209. .1534
FT                   /*tag= a
FT                   /product= "CD40 receptor interacting protein 4C4"
FT                   /note= "no stop codon is given at the 3' end of the
FT                   sequence"
FT                   /transl_except= (pos:1391. .1393, aa:Xaa)
FT                   /transl_except= (pos:1529. .1531, aa:Xaa)
FT                   /note= "Xaa = unknown"
XX
PN   WO9955859-A2.
XX
PD   04-NOV-1999.
XX
PF   28-APR-1999;   99WO-EP003025.
XX
```

```
PR   29-APR-1998;   98EP-00201392.
XX
PA   (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI   Pype SMC,  Remacle JEFJG,  Huylebroeck DFE;
XX
DR   WPI; 2000-062029/05.
DR   P-PSDB; AAY56021.
XX
PT   Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT   diseases and for improvement of anti-tumor treatments.
XX
PS   Claim 11; Page 45-46; 48pp; English.
XX
CC   This sequence represents the gene encoding human CD40 receptor
CC   interacting protein 4C4. The protein is a functional protein capable of
CC   interacting with the cytoplasmic domain of CD40 and/or other receptors of
CC   the tumour necrosis factor (TNF) receptor superfamily such as CD30 and
CC   TNF receptor I, where the protein has no homology to TNF receptor
CC   associated factor (TRAF)-proteins. The CD40 binding proteins can be used
CC   as modulators of the CD40 signalling pathway, especially to diagnose and
CC   treat TRAF-related, CD40-related, NF-kappaB related and/or Jun (kinase)-
CC   related diseases, and for the improvement of anti-tumour diseases.
CC   Diseases which may be treated include atherosclerosis, arthritis,
CC   multiple sclerosis, systemic lupus erythematosus, graft rejection, graft
CC   versus host disease, allergy, and autoimmune disease. The proteins can be
CC   used to sensitize tumour cells to anti-tumour treatments and to screen
CC   for compounds which interfere with the interaction of the proteins with
CC   other protein components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ   Sequence 1536 BP; 530 A; 248 C; 342 G; 413 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:                4.2e-21       Length:          1536
Score:                    229.00        Matches:         44
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AAZ47120 (1-1536)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 271

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 331

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        332 GTTCTTGAATAC 343

RESULT 5
ADV43036
ID   ADV43036 standard; cDNA; 3042 BP.
XX
AC   ADV43036;
XX
DT   10-MAR-2005  (first entry)
XX
DE   Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 664.
XX
KW   microarray; psychoneuroendocrinimmune; chronic fatigue;
KW   non-insulin dependent diabetes; allergy; immune disorder; inflammation;
KW   cancer; neoplasm; infection; expressed sequence tag; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004108899-A2.
XX
PD   16-DEC-2004.
XX
PF   04-JUN-2004; 2004WO-US017686.
XX
PR   04-JUN-2003; 2003US-0475915P.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Nicholson A,  Vernon SD;
XX
DR   WPI; 2005-031682/03.
XX
PT   New microarray comprising probes for genes involved in
PT   psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
PT   condition associated with PNI activity, e.g., inflammatory or infectious
PT   diseases.
XX
PS   Claim 1; SEQ ID NO 664; 254pp; English.
XX
CC   The invention relates to a new microarray which comprises probes for
CC   genes involved in psychoneuroendocrinimmune (PNI) activity. The
CC   microarray is useful in diagnosing a condition associated with PNI
CC   activity, such as CFS, type-2 diabetes, allergic condition, inflammation,
CC   cancer and infection. The present sequence represents a
CC   psychoneuroendocrinimmune gene expressed sequence tag. Note the
CC   specificatio mentions SEQ ID NO of up to 3314 but only sequences up to
CC   SEQ ID NO 1829 are provided.
XX
SQ   Sequence 3042 BP; 1009 A; 569 C; 648 G; 816 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.25e-20      Length:          3042
Score:                    223.00        Matches:         43
Percent Similarity:       97.7%         Conservative:    0
Best Local Similarity:    97.7%         Mismatches:      1
Query Match:              97.4%         Indels:          0
DB:                       14            Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ADV43036 (1-3042)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 6
ABK12397
ID   ABK12397 standard; cDNA; 3183 BP.
XX
AC   ABK12397;
XX
DT   05-JUN-2002  (first entry)
XX
DE   cDNA #2 encoding human COASTER protein sequence #3.
XX
KW   Human; COASTER; modulation of transcriptional activity; nuclear receptor;
KW   steroid receptor; coactivator; coactivation process; contraception;
KW   hormone-stimulated tumour growth; post-menopausal disorder; aging;
KW   heart disease; contraceptive; cytostatic; cardiant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .3183
FT                   /*tag=  a
FT                   /partial
FT                   /product= "COASTER protein #3"
```

```
FT                   /note= "This sequence lacks a stop codon"
XX
PN   WO200216426-A2.
XX
PD   28-FEB-2002.
XX
PF   16-AUG-2001; 2001WO-EP009499.
XX
PR   21-AUG-2000; 2000EP-00202905.
PR   14-MAY-2001; 2001EP-00201771.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Dechering KJ,  Mosselman S;
XX
DR   WPI; 2002-280916/32.
DR   P-PSDB; AAU77900.
XX
PT   Novel COASTER protein useful for producing anti-COASTER antibodies which
PT   are useful in diagnosis of disorders involving changes in the
PT   coactivation process.
XX
PS   Claim 11; Page 66-67; 78pp; English.
XX
CC   The present invention relates to the isolation of a human COASTER protein
CC   with several different amino acid sequences, and the polynucleotide
CC   sequences encoding them. The sequences of the invention can be used in a
CC   method of modulation of transcriptional activity promoted by a responsive
CC   nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a
CC   COASTER protein). The COASTER protein is useful in the diagnosis of
CC   disorders involving changes in the coactivation process. The method of
CC   the invention is useful for determining the action of hormones and
CC   mechanisms for control of transcription of genes in general. The method
CC   is also useful for developing new medicines to specifically influence
CC   physiological processes related to the functioning of a nuclear receptor
CC   for therapeutic, diagnostic, cosmetic and contraceptive purposes. The
CC   method can be used for the treatment of hormone-stimulated tumour growth,
CC   male or female contraception, menopausal and post-menopausal disorders in
CC   women, heart diseases and for the treatment of aging due to reduced
CC   hormonal activity. The present sequence encodes human COASTER protein
CC   sequence #3
XX
SQ   Sequence 3183 BP; 1048 A; 593 C; 678 G; 864 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.61e-20        Length:         3183
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x ABK12397 (1-3183)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 7
ADR52788
ID   ADR52788 standard; DNA; 4961 BP.
XX
AC   ADR52788;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Drug therapy altered expressed gene #139.
XX
KW   drug activity monitoring; expression profile; gene expression;
KW   peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
KW   CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
KW   mTOR; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004072265-A2.
XX
PD   26-AUG-2004.
XX
PF   11-FEB-2004; 2004WO-US004118.
XX
PR   11-FEB-2003; 2003US-0446133P.
PR   03-APR-2003; 2003US-0459782P.
PR   23-JAN-2004; 2004US-0538246P.
XX
PA   (AMHP ) WYETH.
PA   (BURC/) BURCZYNSKI M.
PA   (TWIN/) TWINE N.
PA   (DORN/) DORNER A J.
PA   (TREP/) TREPICCHIO W L.
XX
PI   Burczynski M,  Twine N,  Dorner AJ,  Trepicchio WL;
XX
DR   WPI; 2004-642301/62.
XX
PT   Monitoring drug activities in vivo comprises comparing an expression
PT   profile of a gene in a peripheral blood sample of a patient before and
PT   after drug therapy.
XX
PS   Disclosure; SEQ ID NO 139; 136pp; English.
XX
CC   The invention relates to a method of monitoring drug activities in vivo
CC   by comparing an expression profile of at least one gene in a peripheral
CC   blood sample of a patient to a reference expression profile of the at
CC   least one gene, where the at least one gene is differentially expressed
CC   in peripheral blood mononuclear cells (PBMCs) of patients who have a non-
CC   blood disease and are subjected to a drug therapy as compared to PBMCs
CC   isolated from the patient before the drug therapy, and where the patient
CC   has the non-blood disease and is being treated by the drug therapy. The
CC   method, kit, and nucleic acid array are useful for monitoring drug
CC   activities in vivo. The drug is especially CCI-779, an ester analogue of
CC   the immunosuppressant rapamycin which is a potent inhibitor of the
CC   mammalian target of rapamycin (mTOR). This sequence represents a gene
CC   expressed in PBMC altered by the drug therapy. (Note: this sequence does
CC   no form part of the printed specification but was obtained in electronic
CC   format from WIPO at ftp.wipo.int/pub/published_pct_sequences/).
XX
SQ   Sequence 4961 BP; 1628 A; 834 C; 987 G; 1512 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.13e-19        Length:         4961
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     13              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x ADR52788 (1-4961)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 106

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        107 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 166

Qy         41 ValLeuGluTyr 44
```

```
                ||||||||||||
Db        167 GTTCTTGAATAC 178

RESULT 8
ABK12391
ID   ABK12391 standard; cDNA; 4999 BP.
XX
AC   ABK12391;
XX
DT   05-JUN-2002  (first entry)
XX
DE   cDNA encoding human COASTER protein sequence #1.
XX
KW   Human; COASTER; modulation of transcriptional activity; nuclear receptor;
KW   steroid receptor; coactivator; coactivation process; contraception;
KW   hormone-stimulated tumour growth; post-menopausal disorder; aging;
KW   heart disease; contraceptive; cytostatic; cardiant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             226. .3267
FT                   /*tag=  a
FT                   /product= "COASTER protein #1"
FT                   /note= "The ORF is specifically claimed in Claim 11"
XX
PN   WO200216426-A2.
XX
PD   28-FEB-2002.
XX
PF   16-AUG-2001; 2001WO-EP009499.
XX
PR   21-AUG-2000; 2000EP-00202905.
PR   14-MAY-2001; 2001EP-00201771.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Dechering KJ,  Mosselman S;
XX
DR   WPI; 2002-280916/32.
DR   P-PSDB; AAU77898.
XX
PT   Novel COASTER protein useful for producing anti-COASTER antibodies which
PT   are useful in diagnosis of disorders involving changes in the
PT   coactivation process.
XX
PS   Claim 11; Page 47-49; 78pp; English.
XX
CC   The present invention relates to the isolation of a human COASTER protein
CC   with several different amino acid sequences, and the polynucleotide
CC   sequences encoding them. The sequences of the invention can be used in a
CC   method of modulation of transcriptional activity promoted by a responsive
CC   nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a
CC   COASTER protein). The COASTER protein is useful in the diagnosis of
CC   disorders involving changes in the coactivation process. The method of
CC   the invention is useful for determining the action of hormones and
CC   mechanisms for control of transcription of genes in general. The method
CC   is also useful for developing new medicines to specifically influence
CC   physiological processes related to the functioning of a nuclear receptor
CC   for therapeutic, diagnostic, cosmetic and contraceptive purposes. The
CC   method can be used for the treatment of hormone-stimulated tumour growth,
CC   male or female contraception, menopausal and post-menopausal disorders in
CC   women, heart diseases and for the treatment of aging due to reduced
CC   hormonal activity. The present sequence encodes human COASTER protein
CC   sequence #1
XX
SQ   Sequence 4999 BP; 1614 A; 857 C; 1037 G; 1491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.14e-19        Length:         4999
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x ABK12391 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 9
ABK12392
ID   ABK12392 standard; cDNA; 4999 BP.
XX
AC   ABK12392;
XX
DT   05-JUN-2002  (first entry)
XX
DE   cDNA encoding human COASTER protein sequence #2.
XX
KW   Human; COASTER; modulation of transcriptional activity; nuclear receptor;
KW   steroid receptor; coactivator; coactivation process; contraception;
KW   hormone-stimulated tumour growth; post-menopausal disorder; aging;
KW   heart disease; contraceptive; cytostatic; cardiant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             226. .3267
FT                   /*tag=  a
FT                   /product= "COASTER protein #2"
FT                   /note= "The ORF is specifically claimed in Claim 11"
XX
PN   WO200216426-A2.
XX
PD   28-FEB-2002.
XX
PF   16-AUG-2001; 2001WO-EP009499.
XX
PR   21-AUG-2000; 2000EP-00202905.
PR   14-MAY-2001; 2001EP-00201771.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Dechering KJ,  Mosselman S;
XX
DR   WPI; 2002-280916/32.
DR   P-PSDB; AAU77899.
XX
PT   Novel COASTER protein useful for producing anti-COASTER antibodies which
PT   are useful in diagnosis of disorders involving changes in the
PT   coactivation process.
XX
PS   Claim 11; Page 49-51; 78pp; English.
XX
CC   The present invention relates to the isolation of a human COASTER protein
CC   with several different amino acid sequences, and the polynucleotide
CC   sequences encoding them. The sequences of the invention can be used in a
CC   method of modulation of transcriptional activity promoted by a responsive
CC   nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a
CC   COASTER protein). The COASTER protein is useful in the diagnosis of
CC   disorders involving changes in the coactivation process. The method of
CC   the invention is useful for determining the action of hormones and
CC   mechanisms for control of transcription of genes in general. The method
CC   is also useful for developing new medicines to specifically influence
CC   physiological processes related to the functioning of a nuclear receptor
```

```
CC   for therapeutic, diagnostic, cosmetic and contraceptive purposes. The
CC   method can be used for the treatment of hormone-stimulated tumour growth,
CC   male or female contraception, menopausal and post-menopausal disorders in
CC   women, heart diseases and for the treatment of aging due to reduced
CC   hormonal activity. The present sequence encodes human COASTER protein
CC   sequence #2
XX
SQ   Sequence 4999 BP; 1614 A; 858 C; 1037 G; 1490 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.14e-19       Length:          4999
Score:                    223.00         Matches:         43
Percent Similarity:       97.7%          Conservative:    0
Best Local Similarity:    97.7%          Mismatches:      1
Query Match:              97.4%          Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ABK12392 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 10
ADX07795
ID   ADX07795 standard; DNA; 4999 BP.
XX
AC   ADX07795;
XX
DT   21-APR-2005  (first entry)
XX
DE   Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 2360.
XX
KW   cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2005012875-A2.
XX
PD   10-FEB-2005.
XX
PF   29-JUL-2004; 2004WO-US024424.
XX
PR   29-JUL-2003; 2003US-0490890P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Li M,  Rupnow BA,  Webster KR,  Jackson DG,  Wong TW;
XX
DR   WPI; 2005-163068/17.
DR   P-PSDB; ADX07796.
XX
PT   Biomarkers useful for predicting or determining the response of a mammal
PT   to a cancer treatment comprising administration of a modulator of cyclin-
PT   dependent kinase activity.
XX
PS   Claim 5; SEQ ID NO 2360; 141pp; English.
XX
CC   This invention describes a novel method of predicting or determining
CC   whether a mammal will respond or is responding to an anti-cancer agent
CC   that modulates cyclin-dependent kinase (cdk) activity. The method
CC   comprises measuring the level of one or more biomarkers selected from
CC   2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC   NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC   invention is utilized in a kit for determining or predicting whether
CC   patient would be susceptible or resistant to treatment by an agent
CC   modulating cdk activity. The invention also describes a method for
CC   utilizing individualized genetic profiles for treating diseases and
CC   disorders based on patient's response and molecular level, specialized
CC   microarrays comprising the biomarkers described, antibodies directed
CC   against the biomarkers and a cell culture model to identify biomarkers.
CC   The cdk modulator is preferably N-5-[[5-(1,1-Dimethylethyl)-2-
CC   oxazolyl]methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC   tartaric acid salt. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC   sequence encodes a biomarker used in the method of the invention.
XX
SQ   Sequence 4999 BP; 1614 A; 858 C; 1037 G; 1490 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.14e-19       Length:          4999
Score:                    223.00         Matches:         43
Percent Similarity:       97.7%          Conservative:    0
Best Local Similarity:    97.7%          Mismatches:      1
Query Match:              97.4%          Indels:          0
DB:                       14             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ADX07795 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 11
ABK12407
ID   ABK12407 standard; cDNA; 5143 BP.
XX
AC   ABK12407;
XX
DT   05-JUN-2002  (first entry)
XX
DE   cDNA #2 encoding human COASTER protein sequence #4.
XX
KW   Human; COASTER; modulation of transcriptional activity; nuclear receptor;
KW   steroid receptor; coactivator; coactivation process; contraception;
KW   hormone-stimulated tumour growth; post-menopausal disorder; aging;
KW   heart disease; contraceptive; cytostatic; cardiant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             226. .3411
FT                   /*tag= a
FT                   /product= "COASTER protein #4"
XX
PN   WO200216426-A2.
XX
PD   28-FEB-2002.
XX
PF   16-AUG-2001; 2001WO-EP009499.
XX
PR   21-AUG-2000; 2000EP-00202905.
PR   14-MAY-2001; 2001EP-00201771.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Dechering KJ,  Mosselman S;
XX
DR   WPI; 2002-280916/32.
```

```
DR   P-PSDB; AAU77901.
XX
PT   Novel COASTER protein useful for producing anti-COASTER antibodies which
PT   are useful in diagnosis of disorders involving changes in the
PT   coactivation process.
XX
PS   Example; Fig 1; 78pp; English.
XX
CC   The present invention relates to the isolation of a human COASTER protein
CC   with several different amino acid sequences, and the polynucleotide
CC   sequences encoding them. The sequences of the invention can be used in a
CC   method of modulation of transcriptional activity promoted by a responsive
CC   nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a
CC   COASTER protein). The COASTER protein is useful in the diagnosis of
CC   disorders involving changes in the coactivation process. The method of
CC   the invention is useful for determining the action of hormones and
CC   mechanisms for control of transcription of genes in general. The method
CC   is also useful for developing new medicines to specifically influence
CC   physiological processes related to the functioning of a nuclear receptor
CC   for therapeutic, diagnostic, cosmetic and contraceptive purposes. The
CC   method can be used for the treatment of hormone-stimulated tumour growth,
CC   male or female contraception, menopausal and post-menopausal disorders in
CC   women, heart diseases and for the treatment of aging due to reduced
CC   hormonal activity. The present sequence encodes human COASTER protein
CC   sequence #4
XX
SQ   Sequence 5143 BP; 1655 A; 883 C; 1067 G; 1538 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.18e-19        Length:          5143
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ABK12407 (1-5143)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 12
ABK12394
ID   ABK12394 standard; cDNA; 5144 BP.
XX
AC   ABK12394;
XX
DT   05-JUN-2002  (first entry)
XX
DE   cDNA #1 encoding human COASTER protein sequence #4.
XX
KW   Human; COASTER; modulation of transcriptional activity; nuclear receptor;
KW   steroid receptor; coactivator; coactivation process; contraception;
KW   hormone-stimulated tumour growth; post-menopausal disorder; aging;
KW   heart disease; contraceptive; cytostatic; cardiant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             226. .3412
FT                   /*tag= a
FT                   /product= "COASTER protein #4"
FT                   /transl_except= (pos:2974. .2977, aa:Gln)
FT                   /note= "This codon has an apparent 1 nucleotide insertion
FT                   which alters the reading frame"
XX
PN   WO200216426-A2.
XX
PD   28-FEB-2002.
XX
PF   16-AUG-2001; 2001WO-EP009499.
XX
PR   21-AUG-2000; 2000EP-00202905.
PR   14-MAY-2001; 2001EP-00201771.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Dechering KJ,  Mosselman S;
XX
DR   WPI; 2002-280916/32.
DR   P-PSDB; AAU77901.
XX
PT   Novel COASTER protein useful for producing anti-COASTER antibodies which
PT   are useful in diagnosis of disorders involving changes in the
PT   coactivation process.
XX
PS   Claim 11; Page 53-55; 78pp; English.
XX
CC   The present invention relates to the isolation of a human COASTER protein
CC   with several different amino acid sequences, and the polynucleotide
CC   sequences encoding them. The sequences of the invention can be used in a
CC   method of modulation of transcriptional activity promoted by a responsive
CC   nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a
CC   COASTER protein). The COASTER protein is useful in the diagnosis of
CC   disorders involving changes in the coactivation process. The method of
CC   the invention is useful for determining the action of hormones and
CC   mechanisms for control of transcription of genes in general. The method
CC   is also useful for developing new medicines to specifically influence
CC   physiological processes related to the functioning of a nuclear receptor
CC   for therapeutic, diagnostic, cosmetic and contraceptive purposes. The
CC   method can be used for the treatment of hormone-stimulated tumour growth,
CC   male or female contraception, menopausal and post-menopausal disorders in
CC   women, heart diseases and for the treatment of aging due to reduced
CC   hormonal activity. The present sequence encodes human COASTER protein
CC   sequence #4
XX
SQ   Sequence 5144 BP; 1656 A; 883 C; 1067 G; 1538 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.18e-19        Length:          5144
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ABK12394 (1-5144)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 13
ABK12393
ID   ABK12393 standard; cDNA; 5144 BP.
XX
AC   ABK12393;
XX
DT   05-JUN-2002  (first entry)
```

```
XX
DE   cDNA #1 encoding human COASTER protein sequence #3.
XX
KW   Human; COASTER; modulation of transcriptional activity; nuclear receptor;
KW   steroid receptor; coactivator; coactivation process; contraception;
KW   hormone-stimulated tumour growth; post-menopausal disorder; aging;
KW   heart disease; contraceptive; cytostatic; cardiant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             226. .3412
FT                   /*tag=  a
FT                   /product= "COASTER protein #3"
FT                   /transl_except= (pos:2974. .2977, aa:Gln)
FT                   /note= "This codon has an apparent 1 nucleotide insertion
FT                   which alters the reading frame"
XX
PN   WO200216426-A2.
XX
PD   28-FEB-2002.
XX
PF   16-AUG-2001; 2001WO-EP009499.
XX
PR   21-AUG-2000; 2000EP-00202905.
PR   14-MAY-2001; 2001EP-00201771.
XX
PA   (ALKU ) AKZO NOBEL NV.
XX
PI   Dechering KJ,  Mosselman S;
XX
DR   WPI; 2002-280916/32.
DR   P-PSDB; AAU77900.
XX
PT   Novel COASTER protein useful for producing anti-COASTER antibodies which
PT   are useful in diagnosis of disorders involving changes in the
PT   coactivation process.
XX
PS   Claim 11; Page 51-53; 78pp; English.
XX
CC   The present invention relates to the isolation of a human COASTER protein
CC   with several different amino acid sequences, and the polynucleotide
CC   sequences encoding them. The sequences of the invention can be used in a
CC   method of modulation of transcriptional activity promoted by a responsive
CC   nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a
CC   COASTER protein). The COASTER protein is useful in the diagnosis of
CC   disorders involving changes in the coactivation process. The method of
CC   the invention is useful for determining the action of hormones and
CC   mechanisms for control of transcription of genes in general. The method
CC   is also useful for developing new medicines to specifically influence
CC   physiological processes related to the functioning of a nuclear receptor
CC   for therapeutic, diagnostic, cosmetic and contraceptive purposes. The
CC   method can be used for the treatment of hormone-stimulated tumour growth,
CC   male or female contraception, menopausal and post-menopausal disorders in
CC   women, heart diseases and for the treatment of aging due to reduced
CC   hormonal activity. The present sequence encodes human COASTER protein
CC   sequence #3
XX
SQ   Sequence 5144 BP; 1656 A; 882 C; 1067 G; 1539 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.18e-19        Length:          5144
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ABK12393 (1-5144)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 14
ADC32216
ID   ADC32216 standard; cDNA; 5245 BP.
XX
AC   ADC32216;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human novel cDNA contig sequence, SEQ ID NO:2298.
XX
KW   Human; diagnostic; drug screening; forensics; gene mapping;
KW   biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW   neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW   ulcers; osteoporosis; autoimmune disease; cancer;
KW   molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW   neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW   antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW   gene therapy; chromosome 6; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003029271-A2.
XX
PD   10-APR-2003.
XX
PF   24-SEP-2002; 2002WO-US030474.
XX
PR   24-SEP-2001; 2001US-0324631P.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang TY,  Zhang J,  Ren F,  Xue AJ,  Zhao QA,  Wang J,  Wehrman T;
PI   Zhou P,  Ghosh M,  Wang D,  Ma Y,  Asundi V,  Wang Z,  Weng G;
PI   Haley-Vicente D,  Drmanac RT;
XX
DR   WPI; 2003-371981/35.
DR   P-PSDB; ADC32983.
XX
PT   New polynucleotide and polypeptide useful for diagnosing, preventing or
PT   treating conditions such as neurodegenerative diseases, anemias, platelet
PT   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT   cancer.
XX
PS   Example 2; SEQ ID NO 2298; 1185pp; English.
XX
CC   The invention relates to 971 novel human cDNA sequences (ADC29919-
CC   ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC   invention also relates to nucleic acid sequences over 99% identical with
CC   the novel human cDNAs. The invention additionally encompasses expression
CC   vectors and host cells comprising a nucleic acid of the invention; the
CC   recombinant production of a polypeptide of the invention; an antibody
CC   against a polypeptide of the invention; a method of detecting
CC   polynucleotides or polypeptides of the invention; and methods of
CC   identifying a compound which binds to a polypeptide of the invention. The
CC   invention further discloses methods of peventing, treating or
CC   ameliorating a medical condition; kits comprising polynucleotide probes
CC   and/or monoclonal antibodies for carrying out the methods of the
CC   invention; methods for the identification of compounds that modulate the
CC   expression or activity of the polynucleotide and/or polypeptide; and 767
CC   contig sequences corresponding to the cDNA sequences of the invention
CC   (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC   -ADC33394). The nucleic acids and polypeptides of the invention are
CC   useful in diagnostics, drug screening, forensics, gene mapping, in the
CC   identification of mutations responsible for genetic disorders or other
```

```
CC   traits, for assessing biodiversity, and in producing many other types of
CC   data and products dependent on DNA and amino acid sequences. They are
CC   also used for treating diseases such as Parkinson's disease, Alzheimer's
CC   disease and other neurodegenerative diseases, anaemia, platelet
CC   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC   cancer. The nucleic acids may also be used as hybridisation probes or
CC   primers, and in the recombinant production of a protein. The polypeptides
CC   are also useful in generating antibodies, as molecular weight markers,
CC   and as food supplements. The present sequence represents a human contig
CC   sequence used in an example of the invention. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5245 BP; 1697 A; 900 C; 1069 G; 1579 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.21e-19        Length:          5245
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       10              Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ADC32216 (1-5245)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 258

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        259 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 318

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        319 GTTCTTGAATAC 330

RESULT 15
AAS81317
ID   AAS81317 standard; cDNA; 3363 BP.
XX
AC   AAS81317;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #17121.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG17130.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 17121; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3363 BP; 1073 A; 640 C; 755 G; 895 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.59e-19        Length:          3363
Score:                    217.00          Matches:         42
Percent Similarity:       95.5%           Conservative:    0
Best Local Similarity:    95.5%           Mismatches:      2
Query Match:              94.8%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AAS81317 (1-3363)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        181 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATTTGAG 240

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        241 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 300

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        301 GTTCTTGAATAC 312

RESULT 16
ADC30114
ID   ADC30114 standard; cDNA; 2949 BP.
XX
AC   ADC30114;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human novel cDNA sequence, SEQ ID NO:196.
XX
KW   Human; diagnostic; drug screening; forensics; gene mapping;
KW   biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW   neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW   ulcers; osteoporosis; autoimmune disease; cancer;
KW   molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW   neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW   antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW   gene therapy; chromosome 6; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003029271-A2.
XX
PD   10-APR-2003.
```

XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR P-PSDB; ADC31085.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 1; SEQ ID NO 196; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of peventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2949 BP; 1042 A; 450 C; 572 G; 885 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.35e-19	Length:	2949
Score:	216.00	Matches:	44
Percent Similarity:	97.8%	Conservative:	0
Best Local Similarity:	97.8%	Mismatches:	0
Query Match:	94.3%	Indels:	1
DB:	10	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x ADC30114 (1-2949)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 279

Qy         21 SerThrThrAspGluAsnGluAspAspIleGln-PheValSerGluGlyProSerArgPr 40
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        280 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTTGTCAGTGAAGGACCATCGAGACC 339

Qy         40 oValLeuGluTyr 44
              |||||||||||||
Db        340 TGTTCTTGAATAC 352
```

RESULT 17
ADC30333
ID ADC30333 standard; cDNA; 5089 BP.
XX
AC ADC30333;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA sequence, SEQ ID NO:415.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 6; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR P-PSDB; ADC31304.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 1; SEQ ID NO 415; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of peventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet

```
CC   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC   cancer. The nucleic acids may also be used as hybridisation probes or
CC   primers, and in the recombinant production of a protein. The polypeptides
CC   are also useful in generating antibodies, as molecular weight markers,
CC   and as food supplements. The present sequence represents a specifically
CC   claimed human cDNA sequence of the invention. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5089 BP; 1640 A; 872 C; 1042 G; 1535 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.42e-18        Length:          5089
Score:                    215.00          Matches:         41
Percent Similarity:       93.2%           Conservative:    0
Best Local Similarity:    93.2%           Mismatches:      3
Query Match:              93.9%           Indels:          0
DB:                       10              Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ADC30333 (1-5089)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 234

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||||||||||||||||||||||||||   ||||||   ||||||||||||   ||||||
Db        235 TCAACAACGGATGAAAATGAAGACGACTTTCAGTTTGCCAGTGAAGGACCATTACGACCT 294

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        295 GTTCTTGAATAC 306

RESULT 18
AAS81316
ID   AAS81316 standard; cDNA; 394 BP.
XX
AC   AAS81316;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #17120.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG17129.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 17120; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 394 BP; 110 A; 103 C; 111 G; 70 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.69e-14        Length:          394
Score:                    175.00          Matches:         34
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              76.4%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AAS81316 (1-394)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 141

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              ||||||||||||||||||||||||||||||||||||||||||
Db        142 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGT 183

RESULT 19
AAS79405
ID   AAS79405 standard; cDNA; 394 BP.
XX
AC   AAS79405;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #15209.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG15218.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
```

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 15209; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 394 BP; 110 A; 103 C; 111 G; 70 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.69e-14	Length:	394
Score:	175.00	Matches:	34
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	76.4%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x AAS79405 (1-394)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 141

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              ||||||||||||||||||||||||||||||||||||||||||
Db        142 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGT 183
```

RESULT 20
ACH40820
ID ACH40820 standard; cDNA; 448 BP.
XX
AC ACH40820;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human foetal brain cDNA #2187.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 28032; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversities, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 448 BP; 142 A; 82 C; 119 G; 91 T; 0 U; 14 Other;

Alignment Scores:

Pred. No.:	1.97e-14	Length:	448
Score:	175.00	Matches:	34
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	76.4%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x ACH40820 (1-448)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 178

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              ||||||||||||||||||||||||||||||||||||||||||
Db        179 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGT 220
```

RESULT 21
AAS81318
ID AAS81318 standard; cDNA; 2039 BP.
XX
AC AAS81318;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #17122.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.

```
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG17131.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 17122; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2039 BP; 620 A; 472 C; 454 G; 491 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:                1.24e-13        Length:         2039
Score:                    175.00          Matches:        34
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              76.4%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AAS81318 (1-2039)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 130

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              ||||||||||||||||||||||||||||||||||||||||||
Db        131 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGT 172

RESULT 22
AAI91954/c
ID   AAI91954 standard; cDNA; 121 BP.
XX
AC   AAI91954;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 12014.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US004927.
XX
PR   28-FEB-2000; 2000US-00515126.
PR   18-MAY-2000; 2000US-00577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   P-PSDB; AAO12023.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT   and treating e.g. leukemia, inflammation and immune disorders.
XX
PS   Claim 1; SEQ ID NO 12014; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 121 BP; 26 A; 38 C; 33 G; 24 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.00183         Length:         121
Score:                    89.00           Matches:        17
Percent Similarity:       94.7%           Conservative:   1
Best Local Similarity:    89.5%           Mismatches:     1
Query Match:              38.9%           Indels:         0
DB:                       4               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AAI91954 (1-121)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||||||||||||||||||||||||||||||||||||||||||||  :::|||
Db         64 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGTTTCATCC 8

RESULT 23
ADU05801
ID   ADU05801 standard; DNA; 4629 BP.
XX
AC   ADU05801;
XX
DT   27-JAN-2005  (first entry)
XX
DE   Novel bronchial cancer-associated human gene SeqID23.
XX
KW   bronchial cancer; cytostatic; tumour-associated protein;
KW   cancer detection; metastasis; tumour; gene; ds; human.
XX
OS   Homo sapiens.
XX
PN   DE10316701-A1.
XX
PD   04-NOV-2004.
```

```
XX
PF   09-APR-2003; 2003DE-01016701.
XX
PR   09-APR-2003; 2003DE-01016701.
XX
PA   (HINZ/) HINZMANN B.
PA   (HERM/) HERMANN K.
PA   (CAST/) HEIDEN CASTANOS-VELEZ E.
XX
PI   Mennerich D, Bruemmendorf T, Heiden E, Hermann K, Kinnemann H;
PI   Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pilarsky C;
XX
DR   WPI; 2004-786403/78.
DR   P-PSDB; ADU06288.
XX
PT   New nucleic acid, and derived proteins, useful for diagnosis of bronchial
PT   cancer and in screening for therapeutic and diagnostic agents.
XX
PS   Claim 1; SEQ ID NO 23; 1381pp; German.
XX
CC   This invention relates to a novel isolated nucleic acid associated with
CC   bronchial cancer comprising 489 defined sequences given in the
CC   specification. The invention may be useful for the production of
CC   compounds with a cytostatic activity through the inhibition of expression
CC   or activity of tumour-associated proteins. The novel DNA sequences and
CC   the proteins/peptides encoded by them are used for detecting bronchial
CC   cancer or determining the risk of developing it and to screen for
CC   specific binding partners of the DNA or protein sequences, where the
CC   binding partners are potentially useful as agents for treating or
CC   diagnosing bronchial cancer. The DNA or protein sequences can also be
CC   used for prognosis, detection of metastases and for secondary treatment
CC   (of tumours that have been stabilised or are no longer detectable).
CC   Detecting abnormal expression of the DNA sequences provides early
CC   diagnosis of bronchial cancers. The present sequence is that of a novel
CC   bronchial cancer-associated human gene sequence of the invention.
XX
SQ   Sequence 4629 BP; 1359 A; 1022 C; 1100 G; 1137 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.:                0.451          Length:          4629
Score:                    85.50          Matches:         19
Percent Similarity:       53.2%          Conservative:    6
Best Local Similarity:    40.4%          Mismatches:      15
Query Match:              37.3%          Indels:          7
DB:                       13             Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x ADU05801 (1-4629)

Qy          3 ProGlySerGluIleIleGluSerValProProAla--------------------Gly 15
              ||||||              ::::::   |||   |||                    |||
Db       1003 CCTGGAGGACCAGCGGCTCAGGCTGCACCAGCAGCACCAGGGCTTTCCCTGGGTGAGGGC 1062

Qy         16 ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                ||||||:::   |||:::   ||||||   |||||||||:::   |||||||||
Db       1063 CGTGAAGCTGCCACAAGCAGTGACGACGAGGAGGAAGATGATGTTGTTTTTGTTTCCTCT 1122

Qy         36 GlyProSerArgProValLeu 42
                |||       |||:::|||
Db       1123 AAGCCTGGGAGCCCCCTACTC 1143

RESULT 24
ADS57834/c
ID   ADS57834 standard; cDNA; 2688 BP.
XX
AC   ADS57834;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Bacterial polynucleotide #9821.
XX
KW   Recombinant DNA construct; transformed plant; improved plant property;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW   pathogen tolerance; pest tolerance; plant disease resistance;
KW   cell cycle pathway modification; plant growth regulator;
KW   homologous recombination; seed oil yield; protein yield; carbohydrate;
KW   nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW   bacterial polynucleotide; gene; ss.
XX
OS   Bacteria.
XX
PN   US2003233675-A1.
XX
PD   18-DEC-2003.
XX
PF   20-FEB-2003; 2003US-00369493.
XX
PR   21-FEB-2002; 2002US-0360039P.
XX
PA   (CAOY/) CAO Y.
PA   (HINK/) HINKLE G J.
PA   (SLAT/) SLATER S C.
PA   (CHEN/) CHEN X.
PA   (GOLD/) GOLDMAN B S.
XX
PI   Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR   WPI; 2004-061375/06.
XX
PT   New recombinant DNA construct comprising a promoter positioned to provide
PT   for expression of a polynucleotide encoding a polypeptide from a
PT   microbial source, useful for producing plants with improved properties.
XX
PS   Claim 1; SEQ ID NO 33508; 122pp; English.
XX
CC   The invention relates to a recombinant DNA construct comprising a
CC   promoter functional in a plant cell, where the promoter is positioned to
CC   provide for expression of a polynucleotide encoding a polypeptide from a
CC   microbial source. The invention also relates to a transformed plant
CC   comprising the recombinant DNA construct and a method of producing a
CC   transformed plant having an improved property. The plant is a crop plant
CC   such as maize or soybean. The method of producing a transformed plant
CC   having an improved property comprises transforming a plant with the
CC   recombinant DNA construct and growing the transformed plant, where the
CC   polynucleotide or polypeptide is useful for improving plant properties.
CC   The recombinant DNA construct is useful for producing plants with
CC   improved plant properties, e.g. improved cold, heat or drought tolerance,
CC   tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC   increased resistance to plant disease, better growth rate by modification
CC   of the cell cycle pathway with plant growth regulators, increased rate of
CC   homologous recombination, modified seed oil or protein yield and/or
CC   content, improved yield by modification of carbohydrate, nitrogen or
CC   phosphorus use and/or uptake, by modification of photosynthesis or by
CC   providing improved plant growth and development under at least one stress
CC   condition, improved lignin production or improved galactomannan
CC   production. This sequence represents a bacterial polynucleotide used in
CC   the scope of the invention. Note: The sequence data for this patent did
CC   not form part of the printed specification but was obtained in electronic
CC   format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 2688 BP; 582 A; 671 C; 829 G; 606 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                5.3            Length:          2688
Score:                    75.50          Matches:         15
Percent Similarity:       61.1%          Conservative:    7
Best Local Similarity:    41.7%          Mismatches:      13
Query Match:              33.0%          Indels:          1
DB:                       13             Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x ADS57834 (1-2688)

Qy          1 GlyAspProGlySerGlu---IleIleGluSerValProProAlaGlyProGluAlaSer 19
              ||||||||||||   |||   :::||||||:::   ||||||      |||
Db        819 GGAGACCCCGGCCAGGAACATCTTATAGAAGCCTTCCCCCCGCAGAACCCCGGTCAGGAC 760

Qy         20 GluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
```

```
                  :::|||   |||:::|||   |||:::   |||      ::::::
Db        759 TACAATACCGCAGACGATAATACTGATAATGACCAAATATTTGCCCAG 712
```

RESULT 25
ADQ51222/c
ID ADQ51222 standard; DNA; 624 BP.
XX
AC ADQ51222;
XX
DT 21-OCT-2004 (first entry)
XX
DE Novel canine microarray-related DNA sequence SeqID2524.
XX
KW canine microarray; drug screening; toxicity assay;
KW environmental pollutant; cellular response; gene expression profile;
KW toxic response; liver necrosis; fatty liver disease;
KW protein adduct formation; hepatitis; dog; ds.
XX
OS Canis familiaris.
XX
PN WO2004063324-A2.
XX
PD 29-JUL-2004.
XX
PF 05-MAY-2003; 2003WO-US013853.
XX
PR 03-MAY-2002; 2002US-0377240P.
XX
PA (GENE-) GENE LOGIC INC.
PA (PFIZ) PFIZER PROD INC.
XX
PI Diggans JC, Porter M, Wei T;
XX
DR WPI; 2004-561890/54.
XX
PT New isolated nucleic acid molecule, useful for drug screening and
PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX
PS Claim 1; SEQ ID NO 2524; 41pp; English.
XX
CC This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and
CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northerns that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the
CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific
CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein adduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least one
CC of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
CC production of a canine microarray of the invention.
XX
SQ Sequence 624 BP; 123 A; 190 C; 177 G; 130 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.:	3.68	Length:	624
Score:	71.00	Matches:	16
Percent Similarity:	63.4%	Conservative:	10
Best Local Similarity:	39.0%	Mismatches:	13
Query Match:	31.0%	Indels:	2
DB:	13	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x ADQ51222 (1-624)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   ||||||:::   :::        |||||||||   ||||||   :::|||
Db        546 GGGGATGGGGGTAGTCAGGCAGTTTCCTTAGTGCCCCCTCCTGGGCCCCTCAGTTCACAC 487

Qy         21 SerThrThrAspGluAsnGluAspAspIle---GlnPheValSerGluGly---ProSer 38
                :::|||      :::   |||::::::   :::   |||:::   |||   |||:::
Db        486 TGCAGCACCACAAGGCACACAGATGAGGTCCCGAGGCTTGTTGCTGGAGGCATTCCTGCT 427

Qy         39 Arg 39
              |||
Db        426 CGT 424
```

RESULT 26
ACH18197
ID ACH18197 standard; cDNA; 430 BP.
XX
AC ACH18197;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human adult heart cDNA #2511.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 5409; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversities, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of

```
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 430 BP; 102 A; 135 C; 135 G; 57 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                5.98          Length:          430
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ACH18197 (1-430)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::::::|||      |||   |||   |||         |||
Db        197 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 256

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::      |||   :::|||          |||   |||:::
Db        257 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 316

Qy         41 ValLeuGlu 43
                :::|||
Db        317 GAGATGGAG 325

RESULT 27
ACH37587
ID   ACH37587 standard; cDNA; 493 BP.
XX
AC   ACH37587;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human endothelial cell cDNA #5720.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 24799; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 493 BP; 120 A; 145 C; 148 G; 77 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:                7.05          Length:          493
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ACH37587 (1-493)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::::::|||      |||   |||   |||         |||
Db        178 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 237

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::      |||   :::|||          |||   |||:::
Db        238 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 297

Qy         41 ValLeuGlu 43
                :::|||
Db        298 GAGATGGAG 306

RESULT 28
ACH44839
ID   ACH44839 standard; cDNA; 493 BP.
XX
AC   ACH44839;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human foetal brain cDNA #5564.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
```

```
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 32051; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 493 BP; 124 A; 139 C; 163 G; 65 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:                  7.05        Length:          493
Score:                      68.00       Matches:         14
Percent Similarity:         48.8%       Conservative:    7
Best Local Similarity:      32.6%       Mismatches:      22
Query Match:                29.7%       Indels:          0
DB:                         9           Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ACH44839 (1-493)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||:::::::|||       |||   |||   |||       |||
Db        336 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 395

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                |||::::::      |||   :::|||          |||   |||:::
Db        396 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 455

Qy         41 ValLeuGlu 43
                :::|||
Db        456 GAGATGGAG 464

RESULT 29
ACH43217
ID   ACH43217 standard; cDNA; 500 BP.
XX
AC   ACH43217;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human foetal brain cDNA #3942.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 30429; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 500 BP; 129 A; 135 C; 167 G; 66 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:                  7.17        Length:          500
Score:                      68.00       Matches:         14
Percent Similarity:         48.8%       Conservative:    7
Best Local Similarity:      32.6%       Mismatches:      22
Query Match:                29.7%       Indels:          0
DB:                         9           Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ACH43217 (1-500)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||:::::::|||       |||   |||   |||       |||
Db        329 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 388

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                |||::::::      |||   :::|||          |||   |||:::
Db        389 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 448

Qy         41 ValLeuGlu 43
                :::|||
Db        449 GAGATGGAG 457

RESULT 30
ABQ54916
ID   ABQ54916 standard; cDNA; 2438 BP.
XX
AC   ABQ54916;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human ovarian antigen HNBVO53 cDNA, SEQ ID NO:796.
XX
KW   Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW   ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW   infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
```

```
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200200677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US018569.
XX
PR   07-JUN-2000; 2000US-0209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   P-PSDB; ABP41839.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT   useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT   cancer), immune disorders, cardiovascular disorders and neurological
PT   diseases.
XX
PS   Claim 1; SEQ ID NO 796; 2922pp; English.
XX
CC   The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC   ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC   encompasses polypeptides 90% identical and polynucleotides 95% identical
CC   to the sequences of the invention. The invention additionally relates to
CC   recombinant vectors and host cells comprising human ovarian antigen
CC   polynucleotides, antibodies against human ovarian antigens, and the use
CC   of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC   treating, prognosing or preventing various ovary and/or breast-related
CC   disorders. Such conditions include ovarian cancer and breast cancer, and
CC   metastatic tumours of ovarian or breast origin, reproductive system
CC   disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC   polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC   disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC   shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC   vaginitis), immune disorders (e.g., congenital and acquired
CC   immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC   blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC   respiratory disorders, neurological disorders, gastrointestinal disorders
CC   and urinary system disorders. Ovarian antigen polypeptides and
CC   polynucleotides may also be used in screening for compounds which
CC   modulate ovarian antigen expression or activity. The polynucleotides may
CC   further be used for gene therapy, chromosome mapping, in the
CC   identification of individuals and in forensic analysis, and the
CC   polypeptides may be used as food additives or to prepare antibodies
CC   useful in disease diagnosis, drug targeting and phenotyping. The present
CC   sequence represents cDNA encoding a human ovarian antigen of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2438 BP; 561 A; 761 C; 686 G; 427 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:                48.9          Length:          2438
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ABQ54916 (1-2438)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20

                 |||   |||       |||:::::|||       |||   |||   |||          |||
Db       2081 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 2140

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||::::::      |||   :::|||          |||   |||:::
Db       2141 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 2200

Qy         41 ValLeuGlu 43
                 :::|||
Db       2201 GAGATGGAG 2209

RESULT 31
ADC99117
ID   ADC99117 standard; cDNA; 3713 BP.
XX
AC   ADC99117;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human KPP cDNA - SEQ ID 70.
XX
KW   anti-HIV; antiallergic; antiinflammatory; antianaemic; antiparkinsonian;
KW   nootropic; anticonvulsant; antiarteriosclerotic; antiasthmatic;
KW   immunosuppressive; antithyroid; cytostatic; hepatotropic; dermatological;
KW   antidiabetic; nephrotropic; antigout; thyromimetic; neuroprotective;
KW   osteopathic; antiarthritic; antiparasitic; antihelminthic; antipsoriatic;
KW   uropathic; ophthalmological; antirheumatic; haemostatic; antibacterial;
KW   virucide; protozoacide; fungicide; kinase; phosphatase; KPP;
KW   cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;
KW   cancer; developmental; mental retardation; neurological;
KW   Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;
KW   diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;
KW   helminthic infection; transgenic; gene therapy; human; ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO2003033680-A2.
XX
PD   24-APR-2003.
XX
PF   17-OCT-2002; 2002WO-US033723.
XX
PR   19-OCT-2001; 2001US-0345474P.
PR   02-NOV-2001; 2001US-0343910P.
PR   13-NOV-2001; 2001US-0333098P.
PR   16-NOV-2001; 2001US-0332424P.
PR   30-NOV-2001; 2001US-0334288P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Bandman O,  Baughn MR,  Becha SD,  Borowsky ML,  Duggan BM;
PI   Emerling BM,  Forsythe IJ,  Gandhi AR,  Gorvad AE,  Griffin JA;
PI   Gururajan R,  Hafalia AJA,  Khan FA,  Lal PG,  Lee EA,  Lee SY;
PI   Lindquist EA,  Lu DAM,  Lu Y,  Marquis JP,  Nguyen DB,  Arvizu CS;
PI   Ramkumar J,  Recipon SA,  Richardson TW,  Swarnakar A,  Tang YT;
PI   Thornton MB,  Tran UK,  Chawla NK,  Warren BA,  Yang J,  Yao MG,  Yue H;
PI   Zebarjadian Y;
XX
DR   WPI; 2003-403214/38.
DR   P-PSDB; ADC99065.
XX
PT   New human kinases and phosphatases and polynucleotides, useful for
PT   diagnosing, treating or preventing autoimmune or inflammatory disorders
PT   (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
PT   cancer or hepatitis.
XX
PS   Claim 5; SEQ ID NO 70; 424pp; English.
XX
CC   The invention relates to a novel isolated polypeptide which is a human
CC   kinase and phosphatase (KPP). The KPP polypeptides, polynucleotides,
CC   agonists and antagonists are useful for diagnosing, treating or
CC   preventing cell proliferative disorders such as atherosclerosis,
CC   cirrhosis, hepatitis and cancer, developmental disorders e.g. mental
```

CC retardation, neurological disorders including Alzheimer's disease and
CC Parkinson's disease, autoimmune and inflammatory disorders such as
CC Crohn's disease and diabetes mellitus and finally, viral, bacterial,
CC fungal, parasitic, protozoan or helminthic infections. Furthermore, the
CC polynucleotides encoding KPP may be useful for creating transgenic
CC animals to model human disease, as well as during gene therapy
CC procedures. The current sequence is that of the human KPP cDNA of the
CC invention.
XX
SQ Sequence 3713 BP; 735 A; 1174 C; 1089 G; 715 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	81.4	Length:	3713
Score:	68.00	Matches:	14
Percent Similarity:	48.8%	Conservative:	7
Best Local Similarity:	32.6%	Mismatches:	22
Query Match:	29.7%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x ADC99117 (1-3713)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||:::::::|||       |||   |||   |||       |||

Db       3373 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 3432

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                    |||::::::        |||    :::|||           |||   |||:::

Db       3433 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3492

Qy         41 ValLeuGlu 43
                 :::|||

Db       3493 GATGGAG 3501
```

RESULT 32
ADE09676/c
ID ADE09676 standard; DNA; 3818 BP.
XX
AC ADE09676;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel DNA-related contig nucleotide sequence #398.
XX
KW novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; genetic disorder; contig; ds.
XX
OS Unidentified.
XX
PN WO2003054152-A2.
XX
PD 03-JUL-2003.
XX
PF 10-DEC-2002; 2002WO-US039555.
XX
PR 10-DEC-2001; 2001US-0339739P.
PR 11-DEC-2001; 2001US-0339453P.
PR 14-MAR-2002; 2002US-0365091P.
PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372381P.
PR 12-APR-2002; 2002US-0372615P.
PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX
DR WPI; 2003-569235/53.
XX
PT New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX
PS Disclosure; SEQ ID NO 2220; 1177pp; English.
XX
CC The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present DNA sequence was used in the
CC exemplification of the invention.
XX
SQ Sequence 3818 BP; 800 A; 1092 C; 1152 G; 774 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	84.2	Length:	3818
Score:	68.00	Matches:	14
Percent Similarity:	48.8%	Conservative:	7
Best Local Similarity:	32.6%	Mismatches:	22
Query Match:	29.7%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x ADE09676 (1-3818)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||:::::|||       |||   |||   |||       |||

Db        316 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 257

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                    |||::::::        |||    :::|||           |||   |||:::

Db        256 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 197

Qy         41 ValLeuGlu 43
                 :::|||

Db        196 GATGGAG 188
```

RESULT 33
ADM87570/c
ID ADM87570 standard; cDNA; 3818 BP.
XX
AC ADM87570;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human EST derived nucleotide sequence SEQ ID NO:663.
XX
KW respiratory; cytostatic; antiarthritic; antiinflammatory;
KW gastrointestinal; antibacterial; immunosuppressive; antidiabetic;
KW antirheumatic; gene therapy; molecular weight marker; chromosome marker;
KW chromosome tag; genetic fingerprinting; nutritional supplement; cancer;
KW inflammatory condition; arthritis; inflammatory bowel disease;
KW Crohn's disease; sepsis; rheumatoid arthritis; diabetes mellitus type 1;
KW graft versus host disease; human; expressed sequence tag; EST; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004009834-A2.
XX
PD 29-JAN-2004.
XX
PF 19-JUL-2002; 2002WO-US022858.
XX
PR 21-JUL-2001; 2001US-0306971P.
PR 28-MAR-2002; 2002US-00112944.
XX
PA (NUVE-) NUVELO INC.
XX
PI Tang YT, Yang Y, Weng G, Zhang J, Ren F, Xue A, Wang J;
PI Wehrman T, Ghosh MJ, Wang D, Zhao QA, Wang Z;
XX
DR WPI; 2004-143291/14.
DR P-PSDB; ADM87788.

```
XX
PT   New isolated polynucleotides and polypeptides, useful for treating, e.g.
PT   cancer, lung or liver fibrosis, arthritis, inflammatory bowel disease,
PT   Crohn's disease, rheumatoid arthritis, diabetes mellitus type 1 or graft
PT   versus host disease.
XX
PS   Example 2; SEQ ID NO 663; 591pp; English.
XX
CC   The present invention describes an isolated polynucleotide (I): (a)
CC   comprising a nucleotide sequence selected from SEQ ID NO:1-244; or (b)
CC   which encodes a polypeptide with biological activity, where the
CC   polynucleotide hybridises to (I) under stringent hybridisation conditions
CC   or has greater than 99% sequence identity with (I). (I) has respiratory,
CC   cytostatic, antiarthritic, antiinflammatory, gastrointestinal,
CC   antibacterial, immunosuppressive, antidiabetic and antirheumatic
CC   activities, and can be used in gene therapy. (I) can be used for
CC   generating polynucleotides encoding chimeric or fusion proteins and
CC   heterologous protein sequences. The polynucleotides can be used to
CC   express recombinant protein for analysis, characterisation or therapeutic
CC   use; as markers for tissues in which the corresponding protein is
CC   preferentially expressed; as molecular weight markers on gels; as
CC   chromosome markers or tags to identify chromosomes or to map related gene
CC   positions; to compare with endogenous DNA sequences in patients to
CC   identify potential genetic disorders; as probes to hybridise and discover
CC   genes, related DNA sequences; as a source of information to derive PCR
CC   primers for genetic fingerprinting; as a probe to subtract-out known
CC   sequences in the process of discovering other novel polynucleotides; for
CC   selecting and making oligomers for attachment to a gene chip or other
CC   support, including for examination of expression patterns; to raise anti-
CC   protein antibodies using DNA immunisation techniques; and as an antigen
CC   to raise anti-DNA antibodies or elicit another immune response. The
CC   polynucleotides and polypeptides can also be used as nutritional sources
CC   or supplements, e.g. as a protein or amino acid supplement, as a carbon
CC   source, as a nitrogen source or as a source of carbohydrates. The
CC   polynucleotides and polypeptides can also be used treat cancer. The
CC   compositions are useful for promoting better or faster closure of non-
CC   healing wounds, for the generation and regeneration of tissues, for gut
CC   protection or regeneration and treatment of lung or liver fibrosis,
CC   reperfusion injury in various tissues, and conditions resulting from
CC   systemic cytokine damage. The compositions can also be used to treat
CC   inflammatory conditions (e.g. arthritis, inflammatory bowel disease or
CC   Crohn's disease), sepsis, rheumatoid arthritis, diabetes mellitus type 1
CC   or graft versus host disease. The present sequence represents an
CC   expressed sequence tag (EST) derived nucleotide sequence from the present
CC   invention. N.B. The sequences for this patent were obtained from the
CC   USPTO web site from an equivalent US patent US20040048249A1.
XX
SQ   Sequence 3818 BP; 800 A; 1092 C; 1152 G; 774 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                84.2          Length:          3818
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
DB:                       12            Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ADM87570 (1-3818)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::::|||      |||   |||   |||          |||
Db        316 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 257

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||:::::      |||   :::|||         |||   |||:::
Db        256 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 197

Qy         41 ValLeuGlu 43
                :::|||
Db        196 GAGATGGAG 188

RESULT 34
ACC45139
ID   ACC45139 standard; DNA; 7399 BP.
XX
AC   ACC45139;
XX
DT   11-JUN-2003  (first entry)
XX
DE   Alphavirus replication unit NS1-NS4 gene related sequence.
XX
KW   Vaccine; alpha virus replication unit; NS1-NS4; gene; ds.
XX
OS   Alphavirus.
XX
PN   CN1367020-A.
XX
PD   04-SEP-2002.
XX
PF   19-OCT-2001; 2001CN-00128979.
XX
PR   19-OCT-2001; 2001CN-00128979.
XX
PA   (UYHU-) UNIV HUAXI HOSPITAL SICHUAN.
XX
PI   Zhao L,  Qin S;
XX
DR   WPI; 2003-168348/17.
XX
PT   Nucleic acid vaccine capable of being high-effectively and safely
PT   expressed comprises a conventional nucleic acid vaccine plasmid and alpha
PT   virus replication unit NS1-NS4.
XX
PS   Example; Fig 1; 14pp; Chinese.
XX
CC   The present invention describes a nucleic acid vaccine which can be
CC   effectively safely expressed. The nucleic acid vaccine plasmid forming
CC   the nucleic acid vaccine is formed from a conventional nucleic acid
CC   vaccine plasmid and an alphavirus replication unit NS1-NS4 gene
CC   recombination. After having been injected into body, under the action of
CC   a CMV promoter, the nucleic acid vaccine can transcribe the NS1-NS4
CC   replication unit and RNA of its downstream target gene with high
CC   replication, and it possess high expression efficiency of its target
CC   antigen, and after its transcription in the cell and expression level has
CC   reach a certain high level, it can induce the cell transfected by
CC   performing a death process to form a self-destroying mechanism so as to
CC   effectively solve the safety problem of nucleic acid vaccines. The
CC   present sequence represents an alphavirus replication unit NS1-NS4 gene
CC   related nucleotide sequence, which is used in an example from the present
CC   invention
XX
SQ   Sequence 7399 BP; 1966 A; 1915 C; 2025 G; 1493 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                256           Length:          7399
Score:                    67.00         Matches:         14
Percent Similarity:       51.2%         Conservative:    7
Best Local Similarity:    34.1%         Mismatches:      20
Query Match:              29.3%         Indels:          0
DB:                       8             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ACC45139 (1-7399)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||                |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                     |||      |||   |||      |||   |||||||||:::   |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100

RESULT 35
```

```
AAX78129
ID   AAX78129 standard; DNA; 8010 BP.
XX
AC   AAX78129;
XX
DT   20-AUG-1999  (first entry)
XX
DE   Semliki Forest virus replicon fragment.
XX
KW   Alphavirus; replicon; expression vector; heterologous splice site;
KW   RNA splicing; viral RNA replication; antigen; vaccination; therapy; ss.
XX
OS   Semliki forest virus.
XX
PN   WO9925859-A1.
XX
PD   27-MAY-1999.
XX
PF   13-NOV-1998;   98WO-CA001065.
XX
PR   14-NOV-1997;   97US-0065793P.
XX
PA   (CONN-) CONNAUGHT LAB LTD.
XX
PI   Parrington M;
XX
DR   WPI; 1999-385140/32.
XX
PT   New vector (EV) used to express an antigen for vaccination, e.g. against
PT   human immune deficiency virus.
XX
PS   Disclosure; Fig 3A-F; 60pp; English.
XX
CC   This invention describes the construction of a novel expression vector
CC   (EV) containing an alphavirus replicon having a heterologous splice site.
CC   The expression vector contains (i) a DNA molecule (I) complementary to at
CC   least part of an alphavirus RNA genome, and (ii) inserted into a region
CC   of (I) that is not essential for its replication, a heterologous DNA
CC   (II), under control of a promoter. (I) includes at least one heterologous
CC   splice site (HSS) to prevent aberrant RNA splicing of the alphavirus and
CC   is the complement of the complete alphavirus genomic region essential for
CC   replication of viral RNA. The expression vector is used to express (II)
CC   in humans or animals, e.g. to express an antigen for vaccination (against
CC   human immune deficiency virus) or to produce a therapeutically active
CC   protein or peptide. Introducing an HSS: (1) makes it more likely that any
CC   splicing will occur at this site, rather than at a cryptic splice site;
CC   (2) restores function of the alphavirus when eliminated, and (3) may
CC   improve transport of RNA from the nucleus
XX
SQ   Sequence 8010 BP; 2213 A; 2036 C; 2158 G; 1603 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              282           Length:         8010
Score:                  67.00         Matches:        14
Percent Similarity:     51.2%         Conservative:   7
Best Local Similarity:  34.1%         Mismatches:     20
Query Match:            29.3%         Indels:         0
DB:                     2             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AAX78129 (1-8010)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::|||           |||:::
Db       2701 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 2760

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||   |||      |||   |||||||||:::   |||
Db       2761 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 2820

Qy         41 Val 41
              |||
Db       2821 GTG 2823
```

```
RESULT 36
AAX76588
ID   AAX76588 standard; DNA; 8100 BP.
XX
AC   AAX76588;
XX
DT   11-AUG-1999  (first entry)
XX
DE   SFV EcoRV-SpeI fragment ligated to the hepatitis delta ribozyme.
XX
KW   Plasmid pMP44; hepatitis delta ribozyme; alphavirus vector; vaccine;
KW   paramyxovirus; respiratory syncytial virus fusion protein; RSV F;
KW   detection; infection; immune response; immunisation; ss.
XX
OS   Synthetic.
OS   Hepatitis D virus.
XX
PN   WO9925858-A1.
XX
PD   27-MAY-1999.
XX
PF   13-NOV-1998;   98WO-CA001064.
XX
PR   14-NOV-1997;   97US-0065791P.
XX
PA   (CONN-) CONNAUGHT LAB LTD.
XX
PI   Parrington M,  Li X;
XX
DR   WPI; 1999-370764/31.
XX
PT   Improved alphavirus vector containing DNA encoding a paramyxovirus
PT   protein.
XX
PS   Example 2; Fig 6; 61pp; English.
XX
CC   The present invention describes an improved alphavirus vector containing
CC   DNA encoding an immunogenic paramyxovirus protein and DNA to enhance the
CC   immunoprotective ability of the paramyxovirus protein in vivo. The vector
CC   comprises: (1) a first DNA sequence which is complementary to at least
CC   part of an alphavirus RNA genome and including the complement of the
CC   complete alphavirus RNA genome replication regions to permit in vivo
CC   replication; and (2) a second DNA sequence encoding a paramyxovirus
CC   protein or protein fragment that generates antibodies that specifically
CC   react with the paramyxovirus protein; where the second DNA sequence is
CC   inserted into a region of the first DNA sequence which is non-essential
CC   for replication, the first and second DNA sequences being under
CC   transcriptional control of a promoter. The vector is used to immunize a
CC   host against disease caused by infection with paramyxovirus, especially
CC   respiratory syncytial virus. Human RSV is a major pathogen responsible
CC   for severe respiratory tract infections in infants, young children and
CC   the institutionalized elderly. The vector provides a protective immune
CC   response using a lower dose and less time than vectors in US8476397 and
CC   US8896500. Additionally the vectors using native RSV F produce protective
CC   immune responses in the absence of pretreatment of the animal model with
CC   cardiotoxin, a material known to increase the uptake of DNA and enhance
CC   the immune response. The present sequence represents the SFV EcoRV-SpeI
CC   fragment ligated to the hepatitis delta ribozyme from an example from the
CC   present invention
XX
SQ   Sequence 8100 BP; 2231 A; 2063 C; 2189 G; 1617 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              286           Length:         8100
Score:                  67.00         Matches:        14
Percent Similarity:     51.2%         Conservative:   7
Best Local Similarity:  34.1%         Mismatches:     20
Query Match:            29.3%         Indels:         0
DB:                     2             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AAX76588 (1-8100)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
```

```
                ||||||||||      :::::::::   ::::::|||                 |||:::
Db       2701 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 2760

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                       |||      |||   |||      |||   |||||||||:::   |||
Db       2761 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 2820

Qy         41 Val 41
              |||
Db       2821 GTG 2823
```

RESULT 37
AAX78130
ID AAX78130 standard; DNA; 11060 BP.
XX
AC AAX78130;
XX
DT 20-AUG-1999 (first entry)
XX
DE pSFV link DNA.
XX
KW Alphavirus; replicon; expression vector; heterologous splice site;
KW RNA splicing; viral RNA replication; antigen; vaccination; therapy; ss.
XX
OS Semliki forest virus.
OS Synthetic.
XX
PN WO9925859-A1.
XX
PD 27-MAY-1999.
XX
PF 13-NOV-1998; 98WO-CA001065.
XX
PR 14-NOV-1997; 97US-0065793P.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
PI Parrington M;
XX
DR WPI; 1999-385140/32.
XX
PT New vector (EV) used to express an antigen for vaccination, e.g. against
PT human immune deficiency virus.
XX
PS Disclosure; Fig 4A-H; 60pp; English.
XX
CC This invention describes the construction of a novel expression vector
CC (EV) containing an alphavirus replicon having a heterologous splice site.
CC The expression vector contains (i) a DNA molecule (I) complementary to at
CC least part of an alphavirus RNA genome, and (ii) inserted into a region
CC of (I) that is not essential for its replication, a heterologous DNA
CC (II), under control of a promoter. (I) includes at least one heterologous
CC splice site (HSS) to prevent aberrant RNA splicing of the alphavirus and
CC is the complement of the complete alphavirus genomic region essential for
CC replication of viral RNA. The expression vector is used to express (II)
CC in humans or animals, e.g. to express an antigen for vaccination (against
CC human immune deficiency virus) or to produce a therapeutically active
CC protein or peptide. Introducing an HSS: (1) makes it more likely that any
CC splicing will occur at this site, rather than at a cryptic splice site;
CC (2) restores function of the alphavirus when eliminated, and (3) may
CC improve transport of RNA from the nucleus
XX
SQ Sequence 11060 BP; 2978 A; 2826 C; 2926 G; 2330 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	417	Length:	11060
Score:	67.00	Matches:	14
Percent Similarity:	51.2%	Conservative:	7
Best Local Similarity:	34.1%	Mismatches:	20
Query Match:	29.3%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x AAX78130 (1-11060)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||                 |||:::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                       |||      |||   |||      |||   |||||||||:::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101
```

RESULT 38
ABN86687
ID ABN86687 standard; DNA; 11489 BP.
XX
AC ABN86687;
XX
DT 05-NOV-2002 (first entry)
XX
DE Nucleotide sequence of a pSCA1 suicide DNA vector.
XX
KW Major histocompatibility complex; MHC; antigen presenting cell; APC;
KW antigen; cytostatic; virucide; gene therapy; CD8; vaccine; therapeutic;
KW cancer; viral infection; ds.
XX
OS Synthetic.
XX
PN WO200261113-A2.
XX
PD 08-AUG-2002.
XX
PF 01-FEB-2002; 2002WO-US002598.
XX
PR 01-FEB-2001; 2001US-0265334P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Wu T, Hung C;
XX
DR WPI; 2002-619261/66.
XX
PT Nucleic acid molecule encoding a fusion polypeptide that promotes
PT processing via the Major Histocompatibility Complex class I pathway
PT and/or promotes activity of an antigen presenting cell, useful as vaccine
PT for cancer and viral infections.
XX
PS Claim 32; Page 25-27; 127pp; English.
XX
CC The invention relates to a new nucleic acid molecule (I) encoding a
CC fusion polypeptide useful as a vaccine composition. (I) comprises a first
CC nucleic acid sequence encoding a first polypeptide or peptide that
CC promotes processing via the Major Histocompatibility Complex (MHC) class
CC I pathway (MHC-I-PP) and/or promotes development or activity of an
CC antigen presenting cell (APC). The second nucleic acid sequence is linked
CC in frame to the first nucleic acid sequence or to a linker nucleic acid
CC sequence and encodes an antigenic polypeptide or peptide. The methods and
CC compositions of the present invention are useful as therapeutic vaccine
CC for cancer and for major viral infections, such as hepatoma and cervical
CC cancer, that cause morbidity and mortality. They can also be used in
CC treating animal diseases, such as equine herpesvirus, bovine viruses,
CC Marek's disease, retroviral and lentiviral diseases and rabies, in the
CC veterinary medicine context. The present sequence represents the
CC nucleotide sequence of a pSCA1 suicide DNA vector
XX
SQ Sequence 11489 BP; 3105 A; 2908 C; 2988 G; 2488 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	437	Length:	11489
Score:	67.00	Matches:	14
Percent Similarity:	51.2%	Conservative:	7
Best Local Similarity:	34.1%	Mismatches:	20

```
Query Match:            29.3%          Indels:        0
DB:                     6              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x ABN86687 (1-11489)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::|||                 |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||       |||   |||       |||   ||||||||||:::   |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100

RESULT 39
AEA40812
ID   AEA40812 standard; DNA; 11489 BP.
XX
AC   AEA40812;
XX
DT   28-JUL-2005  (first entry)
XX
DE   Anti-apoptotic pSCA1 suicide vector DNA.
XX
KW   immunogenicity; immunogenicity-potentiating polypeptide; IPP; T-cell;
KW   vaccine; immune stimulation; tumor; cytostatic; ds.
XX
OS   Synthetic.
OS   Unidentified.
XX
PN   WO2005047501-A1.
XX
PD   26-MAY-2005.
XX
PF   24-FEB-2004; 2004WO-US005292.
XX
PR   24-FEB-2003; 2003US-0449429P.
PR   18-JUL-2003; 2003US-0488527P.
PR   31-DEC-2003; 2003US-0533792P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Wu T,  Hung CF,  Kim T;
XX
DR   WPI; 2005-367009/37.
XX
PT   New nucleic acid composition comprising a first nucleic acid vector
PT   encoding an antigenic polypeptide and a second nucleic acid vector
PT   encoding an anti-apoptotic polypeptide, useful as an immunogen for
PT   inhibiting the growth of a tumor.
XX
PS   Disclosure; SEQ ID NO 2; 158pp; English.
XX
CC   The invention relates to a novel nucleic acid composition useful as an
CC   immunogen. The composition comprises a combination of: a first nucleic
CC   acid vector comprising a first sequence encoding an antigenic polypeptide
CC   or peptide, and optionally, a second sequence linked to the first
CC   sequence and encoding an immunogenicity-potentiating polypeptide (IPP);
CC   and a second nucleic acid vector encoding an anti-apoptotic polypeptide.
CC   When the second vector is administered with the first vector to a
CC   subject, a T cell mediated immune response to the antigenic polypeptide
CC   or peptide is induced that is greater in magnitude and/or duration than
CC   an immune response induced by administration of the first vector alone.
CC   The invention further includes: a particle comprising a material that is
CC   suitable for introduction into a cell or an animal by particle
CC   bombardment, bound to which is the first and second vectors or
CC   composition; a pharmaceutical composition capable of inducing or
CC   enhancing an antigen specific immune response, comprising the particle
CC   and a carrier or an excipient; inducing or enhancing an antigen specific
CC   immune response in a subject; increasing the numbers of CD8+ CTLs
CC   specific for a selected desired antigen in a subject; and inhibiting the
CC   growth of a tumor in a subject. The nucleic acid composition is useful as
CC   an immunogen for inhibiting the growth of a tumor, hence it has
CC   cytostatic activity. This polynucleotide sequence represents an anti-
CC   apoptotic pSCA1 suicide vector DNA of the invention.
XX
SQ   Sequence 11489 BP; 3105 A; 2908 C; 2988 G; 2488 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              437            Length:        11489
Score:                  67.00          Matches:       14
Percent Similarity:     51.2%          Conservative:  7
Best Local Similarity:  34.1%          Mismatches:    20
Query Match:            29.3%          Indels:        0
DB:                     14             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x AEA40812 (1-11489)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::|||                 |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||       |||   |||       |||   ||||||||||:::   |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100

RESULT 40
AAQ26021
ID   AAQ26021 standard; DNA; 11517 BP.
XX
AC   AAQ26021;
XX
DT   25-MAR-2003  (revised)
DT   05-JAN-1993  (first entry)
XX
DE   pSP6-SFV4 RNA transcript as DNA.
XX
KW   Semliki forest virus; SFV; SP6 expression vector; RNA polymerase; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             87. .7379
FT                   /*tag=  a
FT                   /label= Non-structural_polyprotein
FT   CDS             7421. .11182
FT                   /*tag=  b
FT                   /label= Structural_polyprotein
XX
PN   WO9210578-A1.
XX
PD   25-JUN-1992.
XX
PF   12-DEC-1991;   91WO-SE000855.
XX
PR   13-DEC-1990;   90SE-00003978.
XX
PA   (BIOP-) BIOPTION AB.
XX
PI   Garoff H,  Liljestroem P;
XX
DR   WPI; 1992-234633/28.
DR   P-PSDB; AAR25138, AAR28337.
XX
PT   RNA mol. derived from alphavirus RNA genome - chimeric alphavirus antigen
PT   and vaccine for immunisation against viral infections.
XX
PS   Disclosure; Fig 5; 94pp; English.
XX
```

```
CC   The sequence given contains a full length Semliki forest virus (SFV) cDNA
CC   clone within an SP6 expression vector. The SP6 RNA polymerase promoter
CC   allows in vitro transcription of full length and infectious transcripts.
CC   (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to
CC   correct PI field.)
XX
SQ   Sequence 11517 BP; 3130 A; 2985 C; 3091 G; 2311 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                438           Length:        11517
Score:                    67.00         Matches:       14
Percent Similarity:       51.2%         Conservative:  7
Best Local Similarity:    34.1%         Mismatches:    20
Query Match:              29.3%         Indels:        0
DB:                       2             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x AAQ26021 (1-11517)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||                |||::: 
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||   |||      |||   ||||||||||:::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101

RESULT 41
AEA40832
ID   AEA40832 standard; DNA; 12347 BP.
XX
AC   AEA40832;
XX
DT   28-JUL-2005  (first entry)
XX
DE   Anti-apoptotic suicidal vector pSCA1-BCL-xL DNA.
XX
KW   immunogenicity; immunogenicity-potentiating polypeptide; IPP; T-cell;
KW   vaccine; immune stimulation; tumor; cytostatic; ds.
XX
OS   Synthetic.
OS   Unidentified.
XX
PN   WO2005047501-A1.
XX
PD   26-MAY-2005.
XX
PF   24-FEB-2004; 2004WO-US005292.
XX
PR   24-FEB-2003; 2003US-0449429P.
PR   18-JUL-2003; 2003US-0488527P.
PR   31-DEC-2003; 2003US-0533792P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Wu T,  Hung CF,  Kim T;
XX
DR   WPI; 2005-367009/37.
XX
PT   New nucleic acid composition comprising a first nucleic acid vector
PT   encoding an antigenic polypeptide and a second nucleic acid vector
PT   encoding an anti-apoptotic polypeptide, useful as an immunogen for
PT   inhibiting the growth of a tumor.
XX
PS   Disclosure; SEQ ID NO 24; 158pp; English.
XX
CC   The invention relates to a novel nucleic acid composition useful as an
CC   immunogen. The composition comprises a combination of: a first nucleic
CC   acid vector comprising a first sequence encoding an antigenic polypeptide
CC   or peptide, and optionally, a second sequence linked to the first
CC   sequence and encoding an immunogenicity-potentiating polypeptide (IPP);
CC   and a second nucleic acid vector encoding an anti-apoptotic polypeptide.
CC   When the second vector is administered with the first vector to a
CC   subject, a T cell mediated immune response to the antigenic polypeptide
CC   or peptide is induced that is greater in magnitude and/or duration than
CC   an immune response induced by administration of the first vector alone.
CC   The invention further includes: a particle comprising a material that is
CC   suitable for introduction into a cell or an animal by particle
CC   bombardment, bound to which is the first and second vectors or
CC   composition; a pharmaceutical composition capable of inducing or
CC   enhancing an antigen specific immune response, comprising the particle
CC   and a carrier or an excipient; inducing or enhancing an antigen specific
CC   immune response in a subject; increasing the numbers of CD8+ CTLs
CC   specific for a selected desired antigen in a subject; and inhibiting the
CC   growth of a tumor in a subject. The nucleic acid composition is useful as
CC   an immunogen for inhibiting the growth of a tumor, hence it has
CC   cytostatic activity. This polynucleotide sequence represents an anti-
CC   apoptotic suicidal vector pSCA1-BCL-xL DNA of the invention.
XX
SQ   Sequence 12347 BP; 3308 A; 3127 C; 3249 G; 2663 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                476           Length:        12347
Score:                    67.00         Matches:       14
Percent Similarity:       51.2%         Conservative:  7
Best Local Similarity:    34.1%         Mismatches:    20
Query Match:              29.3%         Indels:        0
DB:                       14            Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x AEA40832 (1-12347)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||                |||::: 
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||   |||      |||   ||||||||||:::   |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100

RESULT 42
AAC55473
ID   AAC55473 standard; DNA; 12464 BP.
XX
AC   AAC55473;
XX
DT   11-JAN-2001  (first entry)
XX
DE   Destination vector pDEST9 nucleotide sequence.
XX
KW   Bacteriophage lambda; att; recombination site; attB; attP; attR; attL;
KW   mutant; recombinational cloning; entry vector; destination vector;
KW   gene product targeting; fusion tag cleavage; ds.
XX
OS   Bacteriophage lambda.
OS   Synthetic.
XX
PN   WO200052027-A1.
XX
PD   08-SEP-2000.
XX
PF   02-MAR-2000; 2000WO-US005432.
XX
PR   02-MAR-1999;   99US-0122389P.
PR   23-MAR-1999;   99US-0126049P.
PR   28-MAY-1999;   99US-0136744P.
XX
PA   (LIFE-) LIFE TECHNOLOGIES INC.
XX
PI   Hartley JL,  Brasch MA,  Temple GF,  Cheo D;
```

```
XX
DR   WPI; 2000-543948/49.
XX
PT   Isolated nucleic acid molecules encoding an attB1, attB2, attP1, attP2,
PT   attL1, attL2, attR1, and attR2 nucleotide sequence useful for the
PT   recombinational cloning of polypeptides.
XX
PS   Disclosure; Fig 29; 459pp; English.
XX
CC   The present invention describes isolated nucleic acid molecules (I)
CC   encoding an attB1, attB2, attP1, attP2, attL1, attL2, attR1, and attR2
CC   nucleotide sequence. Also described are: (1) an isolated nucleic acid
CC   molecule (II) comprising one or more att recombination sites comprising
CC   at least one mutation in its core region that increases the specificity
CC   of interaction between the recombination site and a second att
CC   recombination site; and (2) an isolated nucleic acid molecule (III)
CC   comprising one or more mutated att recombination sites comprising at
CC   least one mutation in its core region that enhances the efficiency of
CC   recombination between a first nucleic acid molecule comprising the
CC   mutated att recombination site and a second nucleic acid molecule
CC   comprising a second recombination site that interacts with the mutated
CC   att recombination site. (I), (II), (III), primers, vectors and methods
CC   from the present invention are used for the recombinational cloning of
CC   nucleic acid molecules. They can be used for changing vectors, targeting
CC   gene products to intracellular locations, cleaving fusion tags from
CC   desired proteins, operably linking nucleic acid molecules of interest to
CC   regulatory genetic sequences, constructing genes for fusion proteins,
CC   changing copy number, changing replicons, cloning into phages and
CC   cloning. (I), (II), (III), host cells and vectors can be used in the
CC   production of polypeptides and antibodies. The present sequence is used
CC   in the exemplification of the present invention
XX
SQ   Sequence 12464 BP; 3328 A; 3120 C; 3257 G; 2759 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                482             Length:         12464
Score:                    67.00           Matches:        14
Percent Similarity:       51.2%           Conservative:   7
Best Local Similarity:    34.1%           Mismatches:     20
Query Match:              29.3%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AAC55473 (1-12464)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||              |||:::
Db       8244 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 8303

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||      |||  |||      |||  ||||||||||:::  |||
Db       8304 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 8363

Qy         41 Val 41
              |||
Db       8364 GTG 8366

RESULT 43
AEA40833
ID   AEA40833 standard; DNA; 12612 BP.
XX
AC   AEA40833;
XX
DT   28-JUL-2005  (first entry)
XX
DE   Anti-apoptotic vector pSCA1-E7/BCL-xL DNA.
XX
KW   immunogenicity; immunogenicity-potentiating polypeptide; IPP; T-cell;
KW   vaccine; immune stimulation; tumor; cytostatic; ds.
XX
OS   Human papillomavirus.
OS   Synthetic.
OS   Unidentified.
XX
PN   WO2005047501-A1.
XX
PD   26-MAY-2005.
XX
PF   24-FEB-2004; 2004WO-US005292.
XX
PR   24-FEB-2003; 2003US-0449429P.
PR   18-JUL-2003; 2003US-0488527P.
PR   31-DEC-2003; 2003US-0533792P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Wu T,  Hung CF,  Kim T;
XX
DR   WPI; 2005-367009/37.
XX
PT   New nucleic acid composition comprising a first nucleic acid vector
PT   encoding an antigenic polypeptide and a second nucleic acid vector
PT   encoding an anti-apoptotic polypeptide, useful as an immunogen for
PT   inhibiting the growth of a tumor.
XX
PS   Disclosure; SEQ ID NO 25; 158pp; English.
XX
CC   The invention relates to a novel nucleic acid composition useful as an
CC   immunogen. The composition comprises a combination of: a first nucleic
CC   acid vector comprising a first sequence encoding an antigenic polypeptide
CC   or peptide, and optionally, a second sequence linked to the first
CC   sequence and encoding an immunogenicity-potentiating polypeptide (IPP);
CC   and a second nucleic acid vector encoding an anti-apoptotic polypeptide.
CC   When the second vector is administered with the first vector to a
CC   subject, a T cell mediated immune response to the antigenic polypeptide
CC   or peptide is induced that is greater in magnitude and/or duration than
CC   an immune response induced by administration of the first vector alone.
CC   The invention further includes: a particle comprising a material that is
CC   suitable for introduction into a cell or an animal by particle
CC   bombardment, bound to which is the first and second vectors or
CC   composition; a pharmaceutical composition capable of inducing or
CC   enhancing an antigen specific immune response, comprising the particle
CC   and a carrier or an excipient; inducing or enhancing an antigen specific
CC   immune response in a subject; increasing the numbers of CD8+ CTLs
CC   specific for a selected desired antigen in a subject; and inhibiting the
CC   growth of a tumor in a subject. The nucleic acid composition is useful as
CC   an immunogen for inhibiting the growth of a tumor, hence it has
CC   cytostatic activity. This polynucleotide sequence represents an anti-
CC   apoptotic vector pSCA1-E7/BCL-xL DNA of the invention.
XX
SQ   Sequence 12612 BP; 3392 A; 3178 C; 3309 G; 2733 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                489             Length:         12612
Score:                    67.00           Matches:        14
Percent Similarity:       51.2%           Conservative:   7
Best Local Similarity:    34.1%           Mismatches:     20
Query Match:              29.3%           Indels:         0
DB:                       14              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AEA40833 (1-12612)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||              |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||      |||  |||      |||  ||||||||||:::  |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100

RESULT 44
ADL06634
ID   ADL06634 standard; DNA; 12837 BP.
```

```
XX
AC   ADL06634;
XX
DT   06-MAY-2004  (first entry)
XX
DE   SFV DNA containing sTNFRI:IgG1 Fc construct.
XX
KW   human; tumour necrosis factor; TNF; immunoglobulin G1; heavy chain;
KW   constant region; chimeric protein; husTNFR:IgG1 Fc; rheumatoid arthritis;
KW   pyaemia; heart failure; systemic lupus erythematosus; SLE;
KW   autoimmune disease; ds.
XX
OS   Homo sapiens.
OS   Semliki forest virus.
OS   Synthetic.
XX
PN   CN1408875-A.
XX
PD   09-APR-2003.
XX
PF   29-SEP-2001; 2001CN-00136032.
XX
PR   29-SEP-2001; 2001CN-00136032.
XX
PA   (VIRO-) VIROLOGY INST CHINA ACAD PREVENTIVE MEDI.
XX
PI   Xu C,  Zhang Z;
XX
DR   WPI; 2003-664079/63.
XX
PT   DNA type semliki forest cerebritis virus carrier for expressing IgG1 Fc
PT   chimeric protein as TNF receptor.
XX
PS   Disclosure; Fig 3; 20pp; Chinese.
XX
CC   This invention describes a novel soluble acceptor encoding a human tumour
CC   necrosis factor (TNF), immunoglobulin G1 heavy chain constant region
CC   chimeric protein (husTNFR:IgG1 Fc). The invention also describes a
CC   Semliki forest cerebritis virus expression carrier which expresses the
CC   recombinant protein in high efficiency. The IgG1 Fc chimeric protein of
CC   the invention may be used to treat rheumatoid arthritis, pyaemia, heart
CC   failure, systemic lupus erythematosus (SLE) and other autoimmune
CC   diseases.
XX
SQ   Sequence 12837 BP; 3443 A; 3296 C; 3357 G; 2741 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              499          Length:        12837
Score:                  67.00        Matches:       14
Percent Similarity:     51.2%        Conservative:  7
Best Local Similarity:  34.1%        Mismatches:    20
Query Match:            29.3%        Indels:        0
DB:                     10           Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x ADL06634 (1-12837)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||               |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||  |||      |||  |||||||||:::  |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100

RESULT 45
ADL06635
ID   ADL06635 standard; DNA; 12855 BP.
XX
AC   ADL06635;
XX
DT   06-MAY-2004  (first entry)
XX
DE   SFV DNA containing sTNFRII:IgG1 Fc construct.
XX
KW   human; tumour necrosis factor; TNF; immunoglobulin G1; heavy chain;
KW   constant region; chimeric protein; husTNFR:IgG1 Fc; rheumatoid arthritis;
KW   pyaemia; heart failure; systemic lupus erythematosus; SLE;
KW   autoimmune disease; ds.
XX
OS   Homo sapiens.
OS   Semliki forest virus.
OS   Synthetic.
XX
PN   CN1408875-A.
XX
PD   09-APR-2003.
XX
PF   29-SEP-2001; 2001CN-00136032.
XX
PR   29-SEP-2001; 2001CN-00136032.
XX
PA   (VIRO-) VIROLOGY INST CHINA ACAD PREVENTIVE MEDI.
XX
PI   Xu C,  Zhang Z;
XX
DR   WPI; 2003-664079/63.
XX
PT   DNA type semliki forest cerebritis virus carrier for expressing IgG1 Fc
PT   chimeric protein as TNF receptor.
XX
PS   Disclosure; Fig 4; 20pp; Chinese.
XX
CC   This invention describes a novel soluble acceptor encoding a human tumour
CC   necrosis factor (TNF), immunoglobulin G1 heavy chain constant region
CC   chimeric protein (husTNFR:IgG1 Fc). The invention also describes a
CC   Semliki forest cerebritis virus expression carrier which expresses the
CC   recombinant protein in high efficiency. The IgG1 Fc chimeric protein of
CC   the invention may be used to treat rheumatoid arthritis, pyaemia, heart
CC   failure, systemic lupus erythematosus (SLE) and other autoimmune
CC   diseases.
XX
SQ   Sequence 12855 BP; 3410 A; 3347 C; 3382 G; 2716 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              500          Length:        12855
Score:                  67.00        Matches:       14
Percent Similarity:     51.2%        Conservative:  7
Best Local Similarity:  34.1%        Mismatches:    20
Query Match:            29.3%        Indels:        0
DB:                     10           Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x ADL06635 (1-12855)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||               |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||  |||      |||  |||||||||:::  |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100

RESULT 46
ABN86690
ID   ABN86690 standard; DNA; 13599 BP.
XX
AC   ABN86690;
XX
DT   05-NOV-2002  (first entry)
```

XX
DE Nucleotide sequence of vector pSCA1-E7-Hsp70.
XX
KW Major histocompatibility complex; MHC; antigen presenting cell; APC;
KW antigen; cytostatic; virucide; gene therapy; CD8; vaccine; therapeutic;
KW cancer; viral infection; HPV; E7; heat shock protein 70; Hsp70; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 7405. .9509
FT /*tag= a
FT /note= "E7-Hsp70 fusion sequence"
XX
PN WO200261113-A2.
XX
PD 08-AUG-2002.
XX
PF 01-FEB-2002; 2002WO-US002598.
XX
PR 01-FEB-2001; 2001US-0265334P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Wu T, Hung C;
XX
DR WPI; 2002-619261/66.
XX
PT Nucleic acid molecule encoding a fusion polypeptide that promotes
PT processing via the Major Histocompatibility Complex class I pathway
PT and/or promotes activity of an antigen presenting cell, useful as vaccine
PT for cancer and viral infections.
XX
PS Claim 37; Page 34-37; 127pp; English.
XX
CC The invention relates to a new nucleic acid molecule (I) encoding a
CC fusion polypeptide useful as a vaccine composition. (I) comprises a first
CC nucleic acid sequence encoding a first polypeptide or peptide that
CC promotes processing via the Major Histocompatibility Complex (MHC) class
CC I pathway (MHC-I-PP) and/or promotes development or activity of an
CC antigen presenting cell (APC). The second nucleic acid sequence is linked
CC in frame to the first nucleic acid sequence or to a linker nucleic acid
CC sequence and encodes an antigenic polypeptide or peptide. The methods and
CC compositions of the present invention are useful as therapeutic vaccine
CC for cancer and for major viral infections, such as hepatoma and cervical
CC cancer, that cause morbidity and mortality. They can also be used in
CC treating animal diseases, such as equine herpesvirus, bovine viruses,
CC Marek's disease, retroviral and lentiviral diseases and rabies, in the
CC veterinary medicine context. The present sequence represents the
CC nucleotide sequence of vector pSCA1-E7-Hsp70 comprising the human
CC papillomavirus (HPV) E7 antigenic protein and M. tuberculosis heat shock
CC protein 70 (Hsp70) fusion sequence
XX
SQ Sequence 13599 BP; 3582 A; 3516 C; 3647 G; 2854 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	536	Length:	13599
Score:	67.00	Matches:	14
Percent Similarity:	51.2%	Conservative:	7
Best Local Similarity:	34.1%	Mismatches:	20
Query Match:	29.3%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x ABN86690 (1-13599)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||           |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||  |||      |||  ||||||||||:::  |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
                |||
Db       3098 GTG 3100
```

RESULT 47
AEA40871
ID AEA40871 standard; DNA; 13599 BP.
XX
AC AEA40871;
XX
DT 28-JUL-2005 (first entry)
XX
DE Anti-apoptotic vector pSCA-E7-Hsp70 fusion DNA.
XX
KW immunogenicity; immunogenicity-potentiating polypeptide; IPP; T-cell;
KW vaccine; immune stimulation; tumor; cytostatic; ds.
XX
OS Mycobacterium tuberculosis.
OS Human papillomavirus.
OS Synthetic.
XX
PN WO2005047501-A1.
XX
PD 26-MAY-2005.
XX
PF 24-FEB-2004; 2004WO-US005292.
XX
PR 24-FEB-2003; 2003US-0449429P.
PR 18-JUL-2003; 2003US-0488527P.
PR 31-DEC-2003; 2003US-0533792P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Wu T, Hung CF, Kim T;
XX
DR WPI; 2005-367009/37.
XX
PT New nucleic acid composition comprising a first nucleic acid vector
PT encoding an antigenic polypeptide and a second nucleic acid vector
PT encoding an anti-apoptotic polypeptide, useful as an immunogen for
PT inhibiting the growth of a tumor.
XX
PS Disclosure; SEQ ID NO 68; 158pp; English.
XX
CC The invention relates to a novel nucleic acid composition useful as an
CC immunogen. The composition comprises a combination of: a first nucleic
CC acid vector comprising a first sequence encoding an antigenic polypeptide
CC or peptide, and optionally, a second sequence linked to the first
CC sequence and encoding an immunogenicity-potentiating polypeptide (IPP);
CC and a second nucleic acid vector encoding an anti-apoptotic polypeptide.
CC When the second vector is administered with the first vector to a
CC subject, a T cell mediated immune response to the antigenic polypeptide
CC or peptide is induced that is greater in magnitude and/or duration than
CC an immune response induced by administration of the first vector alone.
CC The invention further includes: a particle comprising a material that is
CC suitable for introduction into a cell or an animal by particle
CC bombardment, bound to which is the first and second vectors or
CC composition; a pharmaceutical composition capable of inducing or
CC enhancing an antigen specific immune response, comprising the particle
CC and a carrier or an excipient; inducing or enhancing an antigen specific
CC immune response in a subject; increasing the numbers of CD8+ CTLs
CC specific for a selected desired antigen in a subject; and inhibiting the
CC growth of a tumor in a subject. The nucleic acid composition is useful as
CC an immunogen for inhibiting the growth of a tumor, hence it has
CC cytostatic activity. This polynucleotide sequence represents an anti-
CC apoptotic vector pSCA-E7-Hsp70 fusion DNA of the invention.
XX
SQ Sequence 13599 BP; 3582 A; 3516 C; 3647 G; 2854 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	536	Length:	13599
Score:	67.00	Matches:	14
Percent Similarity:	51.2%	Conservative:	7

```
Best Local Similarity:  34.1%          Mismatches:    20
Query Match:            29.3%          Indels:        0
DB:                     14             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x AEA40871 (1-13599)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||              |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||   |||      |||   ||||||||||:::   |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100

RESULT 48
AAX76582
ID   AAX76582 standard; DNA; 15538 BP.
XX
AC   AAX76582;
XX
DT   11-AUG-1999  (first entry)
XX
DE   Plasmid pMP44 nucleotide sequence.
XX
KW   Plasmid pMP44; hepatitis delta ribozyme; alphavirus vector; vaccine;
KW   paramyxovirus; respiratory syncytial virus fusion protein; RSV F;
KW   detection; infection; immune response; immunisation; ss.
XX
OS   Synthetic.
XX
PN   WO9925858-A1.
XX
PD   27-MAY-1999.
XX
PF   13-NOV-1998;   98WO-CA001064.
XX
PR   14-NOV-1997;   97US-0065791P.
XX
PA   (CONN-) CONNAUGHT LAB LTD.
XX
PI   Parrington M,  Li X;
XX
DR   WPI; 1999-370764/31.
XX
PT   Improved alphavirus vector containing DNA encoding a paramyxovirus
PT   protein.
XX
PS   Claim 19; Fig 3; 61pp; English.
XX
CC   The present invention describes an improved alphavirus vector containing
CC   DNA encoding an immunogenic paramyxovirus protein and DNA to enhance the
CC   immunoprotective ability of the paramyxovirus protein in vivo. The vector
CC   comprises: (1) a first DNA sequence which is complementary to at least
CC   part of an alphavirus RNA genome and including the complement of the
CC   complete alphavirus RNA genome replication regions to permit in vivo
CC   replication; and (2) a second DNA sequence encoding a paramyxovirus
CC   protein or protein fragment that generates antibodies that specifically
CC   react with the paramyxovirus protein; where the second DNA sequence is
CC   inserted into a region of the first DNA sequence which is non-essential
CC   for replication, the first and second DNA sequences being under
CC   transcriptional control of a promoter. The vector is used to immunize a
CC   host against disease caused by infection with paramyxovirus, especially
CC   respiratory syncytial virus. Human RSV is a major pathogen responsible
CC   for severe respiratory tract infections in infants, young children and
CC   the institutionalized elderly. The vector provides a protective immune
CC   response using a lower dose and less time than vectors in US8476397 and
CC   US8896500. Additionally the vectors using native RSV F produce protective
CC   immune responses in the absence of pretreatment of the animal model with
CC   cardiotoxin, a material known to increase the uptake of DNA and enhance
CC   the immune response. The present sequence represents a specifically
CC   claimed plasmid pMP44 nucleotide sequence from the present invention
XX
SQ   Sequence 15538 BP; 4313 A; 3774 C; 3837 G; 3614 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              629            Length:        15538
Score:                  67.00          Matches:       14
Percent Similarity:     51.2%          Conservative:  7
Best Local Similarity:  34.1%          Mismatches:    20
Query Match:            29.3%          Indels:        0
DB:                     2              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x AAX76582 (1-15538)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||              |||:::
Db       5525 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 5584

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||   |||      |||   ||||||||||:::   |||
Db       5585 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 5644

Qy         41 Val 41
              |||
Db       5645 GTG 5647

RESULT 49
AAA14651/c
ID   AAA14651 standard; DNA; 77536 BP.
XX
AC   AAA14651;
XX
DT   08-AUG-2000  (first entry)
XX
DE   Nucleotide sequence of the FK-520 biosynthetic gene cluster.
XX
KW   FK-520; polyketide synthase; PKS; gene cluster; immunosuppressant;
KW   Streptomyces hygroscopicus var. ascomyceticus; immunophilin;
KW   FK-506 binding protein; polyketide compound; transplant rejection;
KW   graft-versus-host disease; uveitis; alopecia universalis;
KW   autoimmune chronic active hepatitis; inflammatory bowel disease;
KW   multiple sclerosis; primary biliary cirrhosis; scleroderma;
KW   neurite outgrowth; nerve regrowth; Parkinson's disease;
KW   Alzheimer's disease; stroke; traumatic spinal cord; brain injury;
KW   peripheral neuropathy; ss.
XX
OS   Streptomyces hygroscopicus.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement(412. .1836)
FT                   /*tag=  a
FT                   /note= "fkbW gene"
FT   CDS             complement(2020. .3579)
FT                   /*tag=  b
FT                   /note= "fkbV gene"
FT   CDS             3969. .4496
FT                   /*tag=  c
FT                   /note= "fkbR2 gene"
FT   CDS             complement(4595. .5488)
FT                   /*tag=  d
FT                   /note= "fkbR1 gene"
FT   CDS             5601. .6818
FT                   /*tag=  e
FT                   /note= "fkbE gene"
FT   CDS             6808. .8052
FT                   /*tag=  f
FT                   /note= "fkbF gene"
FT   CDS             8156. .8824
FT                   /*tag=  g
FT                   /note= "fkbG gene"
FT   CDS             complement(9122. .9883)
FT                   /*tag=  h
```

```
FT                   /note= "fkbH gene"
FT   CDS             complement(9894. .10994)
FT                   /*tag= i
FT                   /note= "fkbI gene"
FT   CDS             complement(10987. .11247)
FT                   /*tag= j
FT                   /note= "fkbJ gene"
FT   CDS             complement(11244. .12092)
FT                   /*tag= k
FT                   /note= "fkbK gene"
FT   CDS             complement(12113. .13150)
FT                   /*tag= l
FT                   /note= "fkbL gene"
FT   CDS             complement(13212. .23988)
FT                   /*tag= m
FT                   /note= "fkbC gene"
FT   misc_feature    complement(13452. .13662)
FT                   /*tag= n
FT                   /note= "ACP6"
FT   misc_feature    complement(13761. .14394)
FT                   /*tag= o
FT                   /note= "KR6"
FT   misc_feature    complement(14517. .15294)
FT                   /*tag= p
FT                   /note= "ER6"
FT   misc_feature    complement(15438. .16587)
FT                   /*tag= q
FT                   /note= "dehydratase domain (DH) 6"
FT   misc_feature    complement(16587. .17820)
FT                   /*tag= r
FT                   /note= "acyltransferase domain (AT) 6"
FT   misc_feature    complement(17820. .19053)
FT                   /*tag= s
FT                   /note= "KS6"
FT   misc_feature    complement(19116. .19326)
FT                   /*tag= t
FT                   /note= "ACP5"
FT   misc_feature    complement(19464. .20097)
FT                   /*tag= u
FT                   /note= "KR5"
FT   misc_feature    complement(20241. .21420)
FT                   /*tag= v
FT                   /note= "DH5"
FT   misc_feature    complement(21420. .22653)
FT                   /*tag= w
FT                   /note= "AT5"
FT   misc_feature    complement(22653. .23892)
FT                   /*tag= x
FT                   /note= "KS5"
FT   CDS             complement(23992. .46573)
FT                   /*tag= y
FT                   /note= "fkbB gene"
FT   misc_feature    complement(24163. .24373)
FT                   /*tag= z
FT                   /note= "ACP4"
FT   misc_feature    complement(24997. .26146)
FT                   /*tag= aa
FT                   /note= "DH4 (inactive)"
FT   misc_feature    complement(26146. .27430)
FT                   /*tag= ab
FT                   /note= "AT4"
FT   misc_feature    complement(27430. .28684)
FT                   /*tag= ac
FT                   /note= "KS4"
FT   misc_feature    complement(28750. .28960)
FT                   /*tag= ad
FT                   /note= "ACP3"
FT   misc_feature    complement(28750. .28960)
FT                   /*tag= al
FT                   /note= "DH2 (inactive)"
FT   misc_feature    complement(29092. .29740)
FT                   /*tag= ae
FT                   /note= "KR3"
FT   misc_feature    complement(29869. .31018)
FT                   /*tag= af
FT                   /note= "DH3 (inactive)"
FT   misc_feature    complement(31018. .32185)
FT                   /*tag= ag
FT                   /note= "AT3"
FT   misc_feature    complement(32185. .33439)
FT                   /*tag= ah
FT                   /note= "KS3"
FT   misc_feature    complement(33505. .33715)
FT                   /*tag= ai
FT                   /note= "ACP2"
FT   misc_feature    complement(33823. .34480)
FT                   /*tag= aj
FT                   /note= "KR2"
FT   misc_feature    complement(34606. .35749)
FT                   /*tag= ak
FT   misc_feature    complement(35749. .37144)
FT                   /*tag= am
FT                   /note= "AT2"
FT   misc_feature    complement(37145. .38296)
FT                   /*tag= an
FT                   /note= "KS2"
FT   misc_feature    complement(38371. .38581)
FT                   /*tag= ao
FT                   /note= "ACP1"
FT   misc_feature    complement(38677. .39307)
FT                   /*tag= ap
FT                   /note= "KR1"
FT   misc_feature    complement(39442. .40609)
FT                   /*tag= aq
FT                   /note= "DH1"
FT   misc_feature    complement(40609. .41842)
FT                   /*tag= ar
FT                   /note= "AT1"
FT   misc_feature    complement(41842. .43093)
FT                   /*tag= as
FT                   /note= "KS of extender module 1 (KS1)"
FT   misc_feature    complement(43144. .43660)
FT                   /*tag= at
FT                   /note= "ACP of loading domain"
FT   misc_feature    complement(43777. .44629)
FT                   /*tag= au
FT                   /note= "ER of loading domain"
FT   misc_feature    complement(44974. .46573)
FT                   /*tag= av
FT                   /note= "CoA ligase of laoding domain"
FT   CDS             46754. .47788
FT                   /*tag= aw
FT                   /note= "fkbO gene"
FT   CDS             47785. .52272
FT                   /*tag= ax
FT                   /note= "fkbP gene"
FT   CDS             52275. .71465
FT                   /*tag= ay
FT                   /note= "fkbA gene"
FT   misc_feature    52362. .53576
FT                   /*tag= az
FT                   /note= "KS7"
FT   misc_feature    53577. .54716
FT                   /*tag= bb
FT                   /note= "AT7"
FT   misc_feature    54717. .55871
FT                   /*tag= ba
FT                   /note= "DH7"
FT   misc_feature    56019. .56819
FT                   /*tag= bc
FT                   /note= "ER7"
FT   misc_feature    56943. .57575
FT                   /*tag= bd
FT                   /note= "KR7"
FT   misc_feature    57710. .57920
FT                   /*tag= be
```

```
FT                   /note= "ACP7"
FT   misc_feature    57990. .59243
FT                   /*tag= bf
FT                   /note= "KS8"
FT   misc_feature    59244. .60398
FT                   /*tag= bg
FT                   /note= "AT8"
FT   misc_feature    60399. .61412
FT                   /*tag= bh
FT                   /note= "DH8 (inactive)"
```

Alignment Scores:

Pred. No.:	4.41e+03	Length:	77536
Score:	67.00	Matches:	17
Percent Similarity:	46.9%	Conservative:	6
Best Local Similarity:	34.7%	Mismatches:	16
Query Match:	29.3%	Indels:	10
DB:	3	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x AAA14651 (1-77536)

```
Qy          2 AspProGlySerGluIleIleGluSerValProProAla--------------GlyPro 16
              :::|||      |||   |||        ||||||                    |||
Db      52933 AACCCCGCCGGCGAGGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCG 52874

Qy         17 GluAlaSerGlu--------------SerThrThrAspGluAsnGluAspAspIleGln 31
              |||   |||:::              |||   |||   ::::::|||      :::
Db      52873 GAGCGCAGCGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCG 52814

Qy         32 PheValSerGluGlyProSerArgPro 40
                |||:::   ||||||||||||||
Db      52813 ACCGTGACCGCCGGACCCTCCAGACCG 52787
```

```
RESULT 50
ADB69911
ID   ADB69911 standard; DNA; 2436 BP.
XX
AC   ADB69911;
XX
DT   04-DEC-2003  (first entry)
XX
DE   C. neoformans open reading frame SEQ ID NO:2316.
XX
KW   ds; gene; fungicide; gene therapy; infection.
XX
OS   Cryptococcus neoformans.
XX
PN   WO2003052076-A2.
XX
PD   26-JUN-2003.
XX
PF   17-DEC-2002; 2002WO-US040225.
XX
PR   17-DEC-2001; 2001US-0341261P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Zamudio C,  Eroshkin AM;
XX
DR   WPI; 2003-533017/50.
DR   P-PSDB; ADB70272.
XX
PT   New nucleic acid, useful for preparing a composition for treating an
PT   infection caused by Cryptococcus neoformans.
XX
PS   Claim 2; SEQ ID NO 2316; 136pp; English.
XX
CC   The invention relates to a novel purified or isolated Cryptococcus
CC   neoformans nucleic acid molecule comprising a sequence encoding a
CC   polypeptide comprising a sequence not given in the specification. A
CC   polynucleotide of the invention has fungicide activity, and may have a
CC   use in gene therapy. The nucleic acid is useful for preparing a
CC   composition for treating an infection caused by Cryptococcus neoformans.
CC   The present sequence represents a C. neoformans sequence of the
CC   invention. Note: The sequence data for this patent is not represented in
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ   Sequence 2436 BP; 653 A; 598 C; 686 G; 499 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	125	Length:	2436
Score:	65.00	Matches:	16
Percent Similarity:	57.1%	Conservative:	8
Best Local Similarity:	38.1%	Mismatches:	12
Query Match:	28.4%	Indels:	6
DB:	10	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x ADB69911 (1-2436)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   |||:::   ::::::              :::|||   |||   |||
Db        799 GGTGATGAAGGTACTTCCGTAGTT--------------AGCGGCTCAGAAGAAAGC--- 840

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||   ||||||||||||||:::::   :::          |||   |||   :::|||
Db        841 AGCAACACGGATGAAAATGAAGAGGAGGAGGAGGAGGAGGAAGAAGAACCTGTCAAGCCT 900

Qy         41 ValLeu 42
                |||
Db        901 ACCCTC 906
```

```
RESULT 51
ADB69550
ID   ADB69550 standard; DNA; 3346 BP.
XX
AC   ADB69550;
XX
DT   04-DEC-2003  (first entry)
XX
DE   C. neoformans coding sequence with introns SEQ ID NO:1316.
XX
KW   ds; gene; fungicide; gene therapy; infection.
XX
OS   Cryptococcus neoformans.
XX
PN   WO2003052076-A2.
XX
PD   26-JUN-2003.
XX
PF   17-DEC-2002; 2002WO-US040225.
XX
PR   17-DEC-2001; 2001US-0341261P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Zamudio C,  Eroshkin AM;
XX
DR   WPI; 2003-533017/50.
DR   P-PSDB; ADB70272.
XX
PT   New nucleic acid, useful for preparing a composition for treating an
PT   infection caused by Cryptococcus neoformans.
XX
PS   Claim 3; SEQ ID NO 1316; 136pp; English.
XX
CC   The invention relates to a novel purified or isolated Cryptococcus
CC   neoformans nucleic acid molecule comprising a sequence encoding a
CC   polypeptide comprising a sequence not given in the specification. A
CC   polynucleotide of the invention has fungicide activity, and may have a
CC   use in gene therapy. The nucleic acid is useful for preparing a
CC   composition for treating an infection caused by Cryptococcus neoformans.
CC   The present sequence represents a C. neoformans sequence of the
CC   invention. Note: The sequence data for this patent is not represented in
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pat_sequences.
```

```
XX
SQ   Sequence 3346 BP; 924 A; 793 C; 911 G; 718 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                183           Length:        3346
Score:                    65.00         Matches:       16
Percent Similarity:       57.1%         Conservative:  8
Best Local Similarity:    38.1%         Mismatches:    12
Query Match:              28.4%         Indels:        6
DB:                       10            Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x ADB69550 (1-3346)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   |||:::   ::::::               :::|||   |||   |||
Db       1357 GGTGATGAAGGTACTTCCGTAGTT--------------AGCGGCTCAGAAGAAAGC--- 1398

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||   ||||||||||||||::::::   :::          |||   |||   :::|||
Db       1399 AGCAACACGGATGAAAATGAAGAGGAGGAGGAGGAGGAGGAAGAAGAACCTGTCAAGCCT 1458

Qy         41 ValLeu 42
                |||
Db       1459 ACCCTC 1464

RESULT 52
ADB69189
ID   ADB69189 standard; DNA; 5346 BP.
XX
AC   ADB69189;
XX
DT   04-DEC-2003  (first entry)
XX
DE   C. neoformans genomic DNA sequence SEQ ID NO:316.
XX
KW   ds; gene; fungicide; gene therapy; infection.
XX
OS   Cryptococcus neoformans.
XX
PN   WO2003052076-A2.
XX
PD   26-JUN-2003.
XX
PF   17-DEC-2002; 2002WO-US040225.
XX
PR   17-DEC-2001; 2001US-0341261P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Zamudio C,  Eroshkin AM;
XX
DR   WPI; 2003-533017/50.
DR   P-PSDB; ADB70272.
XX
PT   New nucleic acid, useful for preparing a composition for treating an
PT   infection caused by Cryptococcus neoformans.
XX
PS   Claim 3; SEQ ID NO 316; 136pp; English.
XX
CC   The invention relates to a novel purified or isolated Cryptococcus
CC   neoformans nucleic acid molecule comprising a sequence encoding a
CC   polypeptide comprising a sequence not given in the specification. A
CC   polynucleotide of the invention has fungicide activity, and may have a
CC   use in gene therapy. The nucleic acid is useful for preparing a
CC   composition for treating an infection caused by Cryptococcus neoformans.
CC   The present sequence represents a C. neoformans sequence of the
CC   invention. Note: The sequence data for this patent is not represented in
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ   Sequence 5346 BP; 1449 A; 1242 C; 1367 G; 1287 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                323           Length:        5346
Score:                    65.00         Matches:       16
Percent Similarity:       57.1%         Conservative:  8
Best Local Similarity:    38.1%         Mismatches:    12
Query Match:              28.4%         Indels:        6
DB:                       10            Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x ADB69189 (1-5346)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   |||:::   ::::::               :::|||   |||   |||
Db       2357 GGTGATGAAGGTACTTCCGTAGTT--------------AGCGGCTCAGAAGAAAGC--- 2398

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||   ||||||||||||||::::::   :::          |||   |||   :::|||
Db       2399 AGCAACACGGATGAAAATGAAGAGGAGGAGGAGGAGGAGGAAGAAGAACCTGTCAAGCCT 2458

Qy         41 ValLeu 42
                |||
Db       2459 ACCCTC 2464

RESULT 53
AAA14651
ID   AAA14651 standard; DNA; 77536 BP.
XX
AC   AAA14651;
XX
DT   08-AUG-2000  (first entry)
XX
DE   Nucleotide sequence of the FK-520 biosynthetic gene cluster.
XX
KW   FK-520; polyketide synthase; PKS; gene cluster; immunosuppressant;
KW   Streptomyces hygroscopicus var. ascomyceticus; immunophilin;
KW   FK-506 binding protein; polyketide compound; transplant rejection;
KW   graft-versus-host disease; uveitis; alopecia universalis;
KW   autoimmune chronic active hepatitis; inflammatory bowel disease;
KW   multiple sclerosis; primary biliary cirrhosis; scleroderma;
KW   neurite outgrowth; nerve regrowth; Parkinson's disease;
KW   Alzheimer's disease; stroke; traumatic spinal cord; brain injury;
KW   peripheral neuropathy; ss.
XX
OS   Streptomyces hygroscopicus.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement(412. .1836)
FT                   /*tag=  a
FT                   /note= "fkbW gene"
FT   CDS             complement(2020. .3579)
FT                   /*tag=  b
FT                   /note= "fkbV gene"
FT   CDS             3969. .4496
FT                   /*tag=  c
FT                   /note= "fkbR2 gene"
FT   CDS             complement(4595. .5488)
FT                   /*tag=  d
FT                   /note= "fkbR1 gene"
FT   CDS             5601. .6818
FT                   /*tag=  e
FT                   /note= "fkbE gene"
FT   CDS             6808. .8052
FT                   /*tag=  f
FT                   /note= "fkbF gene"
FT   CDS             8156. .8824
FT                   /*tag=  g
FT                   /note= "fkbG gene"
FT   CDS             complement(9122. .9883)
FT                   /*tag=  h
FT                   /note= "fkbH gene"
FT   CDS             complement(9894. .10994)
FT                   /*tag=  i
FT                   /note= "fkbI gene"
FT   CDS             complement(10987. .11247)
FT                   /*tag=  j
```

```
FT                   /note= "fkbJ gene"
FT   CDS             complement(11244. .12092)
FT                   /*tag= k
FT                   /note= "fkbK gene"
FT   CDS             complement(12113. .13150)
FT                   /*tag= l
FT                   /note= "fkbL gene"
FT   CDS             complement(13212. .23988)
FT                   /*tag= m
FT                   /note= "fkbC gene"
FT   misc_feature    complement(13452. .13662)
FT                   /*tag= n
FT                   /note= "ACP6"
FT   misc_feature    complement(13761. .14394)
FT                   /*tag= o
FT                   /note= "KR6"
FT   misc_feature    complement(14517. .15294)
FT                   /*tag= p
FT                   /note= "ER6"
FT   misc_feature    complement(15438. .16587)
FT                   /*tag= q
FT                   /note= "dehydratase domain (DH) 6"
FT   misc_feature    complement(16587. .17820)
FT                   /*tag= r
FT                   /note= "acyltransferase domain (AT) 6"
FT   misc_feature    complement(17820. .19053)
FT                   /*tag= s
FT                   /note= "KS6"
FT   misc_feature    complement(19116. .19326)
FT                   /*tag= t
FT                   /note= "ACP5"
FT   misc_feature    complement(19464. .20097)
FT                   /*tag= u
FT                   /note= "KR5"
FT   misc_feature    complement(20241. .21420)
FT                   /*tag= v
FT                   /note= "DH5"
FT   misc_feature    complement(21420. .22653)
FT                   /*tag= w
FT                   /note= "AT5"
FT   misc_feature    complement(22653. .23892)
FT                   /*tag= x
FT                   /note= "KS5"
FT   CDS             complement(23992. .46573)
FT                   /*tag= y
FT                   /note= "fkbB gene"
FT   misc_feature    complement(24163. .24373)
FT                   /*tag= z
FT                   /note= "ACP4"
FT   misc_feature    complement(24997. .26146)
FT                   /*tag= aa
FT                   /note= "DH4 (inactive)"
FT   misc_feature    complement(26146. .27430)
FT                   /*tag= ab
FT                   /note= "AT4"
FT   misc_feature    complement(27430. .28684)
FT                   /*tag= ac
FT                   /note= "KS4"
FT   misc_feature    complement(28750. .28960)
FT                   /*tag= ad
FT                   /note= "ACP3"
FT   misc_feature    complement(28750. .28960)
FT                   /*tag= al
FT                   /note= "DH2 (inactive)"
FT   misc_feature    complement(29092. .29740)
FT                   /*tag= ae
FT                   /note= "KR3"
FT   misc_feature    complement(29869. .31018)
FT                   /*tag= af
FT                   /note= "DH3 (inactive)"
FT   misc_feature    complement(31018. .32185)
FT                   /*tag= ag
FT                   /note= "AT3"
FT   misc_feature    complement(32185. .33439)
FT                   /*tag= ah
FT                   /note= "KS3"
FT   misc_feature    complement(33505. .33715)
FT                   /*tag= ai
FT                   /note= "ACP2"
FT   misc_feature    complement(33823. .34480)
FT                   /*tag= aj
FT                   /note= "KR2"
FT   misc_feature    complement(34606. .35749)
FT                   /*tag= ak
FT   misc_feature    complement(35749. .37144)
FT                   /*tag= am
FT                   /note= "AT2"
FT   misc_feature    complement(37145. .38296)
FT                   /*tag= an
FT                   /note= "KS2"
FT   misc_feature    complement(38371. .38581)
FT                   /*tag= ao
FT                   /note= "ACP1"
FT   misc_feature    complement(38677. .39307)
FT                   /*tag= ap
FT                   /note= "KR1"
FT   misc_feature    complement(39442. .40609)
FT                   /*tag= aq
FT                   /note= "DH1"
FT   misc_feature    complement(40609. .41842)
FT                   /*tag= ar
FT                   /note= "AT1"
FT   misc_feature    complement(41842. .43093)
FT                   /*tag= as
FT                   /note= "KS of extender module 1 (KS1)"
FT   misc_feature    complement(43144. .43660)
FT                   /*tag= at
FT                   /note= "ACP of loading domain"
FT   misc_feature    complement(43777. .44629)
FT                   /*tag= au
FT                   /note= "ER of loading domain"
FT   misc_feature    complement(44974. .46573)
FT                   /*tag= av
FT                   /note= "CoA ligase of laoding domain"
FT   CDS             46754. .47788
FT                   /*tag= aw
FT                   /note= "fkbO gene"
FT   CDS             47785. .52272
FT                   /*tag= ax
FT                   /note= "fkbP gene"
FT   CDS             52275. .71465
FT                   /*tag= ay
FT                   /note= "fkbA gene"
FT   misc_feature    52362. .53576
FT                   /*tag= az
FT                   /note= "KS7"
FT   misc_feature    53577. .54716
FT                   /*tag= bb
FT                   /note= "AT7"
FT   misc_feature    54717. .55871
FT                   /*tag= ba
FT                   /note= "DH7"
FT   misc_feature    56019. .56819
FT                   /*tag= bc
FT                   /note= "ER7"
FT   misc_feature    56943. .57575
FT                   /*tag= bd
FT                   /note= "KR7"
FT   misc_feature    57710. .57920
FT                   /*tag= be
FT                   /note= "ACP7"
FT   misc_feature    57990. .59243
FT                   /*tag= bf
FT                   /note= "KS8"
FT   misc_feature    59244. .60398
FT                   /*tag= bg
```

```
FT                   /note= "AT8"
FT   misc_feature    60399. .61412
FT                   /*tag= bh
FT                   /note= "DH8 (inactive)"

Alignment Scores:
Pred. No.:              8.23e+03        Length:         77536
Score:                  65.00           Matches:        17
Percent Similarity:     45.8%           Conservative:   5
Best Local Similarity:  35.4%           Mismatches:     16
Query Match:            28.4%           Indels:         10
DB:                     3               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x AAA14651 (1-77536)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------GlyProGlu 17
              |||      |||   |||          ||||||                  ||||||
Db      28077 CCGGACGGCGAGGCCATCACCGTGACGCCGCCGACCAGGGCGAGCGAGCATTCGCCGGAG 28136

Qy         18 AlaSerGlu---------------SerThrThrAspGluAsnGluAspAspIleGlnPhe 32
                |||:::                |||   |||   ::::::|||      :::
Db      28137 CGCAGTGACTGCCCGGCCTGGTGCAGCGCCACCAGCGACGACGAACACGCCGTGTCGACC 28196

Qy         33 ValSerGluGlyProSerArgPro 40
              |||:::   |||||||||||||||
Db      28197 GTGACCGCCGGACCCTCCAGACCG 28220

RESULT 54
ADE55089
ID   ADE55089 standard; DNA; 2225 BP.
XX
AC   ADE55089;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Rat gene AF052042, SEQ ID NO 894.
XX
KW   Rat; ds; gene; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; AF052042.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat DNA (shown in Table 2 of the
CC   specification) which encodes one of the polypeptides of the invention
CC   which is differentially expressed during pain. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2225 BP; 593 A; 554 C; 612 G; 466 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              130             Length:         2225
Score:                  64.50           Matches:        13
Percent Similarity:     66.7%           Conservative:   7
Best Local Similarity:  43.3%           Mismatches:     7
Query Match:            28.2%           Indels:         3
DB:                     10              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x ADE55089 (1-2225)

Qy          6 GluIleIleGluSerValPro---------ProAlaGlyProGluAlaSerGluSerThr 22
              |||:::   |||::::::|||         |||:::|||||||||   :::||||||
Db        205 GAAGTGGAGGAGACTCTCCCGGAGTCAGGACCCTCCGGTCCGGAGATTGCAGAATCCGAA 264

Qy         23 ThrAspGluAsnGluAspAspIleGlnPhe 32
              |||   ||||||          |||::::::
Db        265 ACTTCTGAAAACAGTCAGACAATTCGCTAC 294

RESULT 55
ADI61727
ID   ADI61727 standard; cDNA; 7390 BP.
XX
AC   ADI61727;
XX
DT   22-APR-2004  (first entry)
XX
DE   Human cDNA downregulated in Alzheimer's disease, INCYTE 337896.11.
XX
KW   Human; ss; Alzheimer's disease; differential display; neuroprotective;
KW   brain disorder.
XX
OS   Homo sapiens.
XX
PN   US6682888-B1.
XX
PD   27-JAN-2004.
XX
PF   05-MAY-2000; 2000US-00566921.
XX
PR   05-MAY-2000; 2000US-00566921.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Loring JF,  Tingley DW,  Edwards CM;
XX
DR   WPI; 2004-118572/12.
XX
PT   New composition comprising cDNAs that are differentially expressed in
PT   brain disorders, useful for diagnosing or treating Alzheimer's disease.
```

```
XX
PS   Claim 1; SEQ ID NO 95; 223pp; English.
XX
CC   The invention relates to a new composition comprising ADI61633-
CC   ADI61770and their complements that are cDNAs differentially expressed in
CC   brain disorders. Also included are a high throughput method for detecting
CC   differential expression of one or more cDNAs in a sample containing
CC   nucleic acids and a high throughput method for screening a library of
CC   molecules or compounds to identify a ligand that specifically binds a
CC   cDNA. The expression of the each of the cDNAs is downregulated at least
CC   two-fold in the brain of the subjects with Alzheimer's disease (ADI61633-
CC   ADI61727) or upregulated at least two fold in Alzheimer's disease
CC   (ADI61728-ADI61770). The composition is useful for diagnosing or treating
CC   Alzheimer's disease. The present sequence is a cDNA downregulated at
CC   least two-fold in the brain of the subjects with Alzheimer's disease.
XX
SQ   Sequence 7390 BP; 1545 A; 1962 C; 2026 G; 1537 T; 0 U; 320 Other;

Alignment Scores:
Pred. No.:                653          Length:          7390
Score:                    64.00        Matches:         17
Percent Similarity:       45.7%        Conservative:    4
Best Local Similarity:    37.0%        Mismatches:      17
Query Match:              27.9%        Indels:          8
DB:                       12           Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x ADI61727 (1-7390)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        332 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 391

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||       ||||||||||||   |||   |||:::   |||            :::
Db        392 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 451

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        452 GAGTCGCCCTCGCACCCG 469

RESULT 56
AEA43892
ID   AEA43892 standard; cDNA; 7390 BP.
XX
AC   AEA43892;
XX
DT   25-AUG-2005  (first entry)
XX
DE   Human cDNA differentially expressed in brain tissue SEQ ID NO:95.
XX
KW   diagnosis; treatment; brain disease; neuroprotective; cerebroprotective;
KW   muscular-gen.; cytostatic; neuroleptic; nootropic; antidepressant;
KW   anticonvulsant; analgesic; antiparkinsonian; ophthalmological;
KW   immunotherapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   US2005130171-A1.
XX
PD   16-JUN-2005.
XX
PF   26-JAN-2004; 2004US-00765700.
XX
PR   05-MAY-2000; 2000US-00566921.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Loring JF,  Tingley DW,  Edwards CM;
XX
DR   WPI; 2005-456990/46.
XX
PT   Composition useful for diagnosis, staging, treating or monitoring
PT   treatment of a subject with a brain disorder, comprises several cDNAs
PT   that are differentially expressed in brain disorders.
XX
PS   Claim 1; SEQ ID NO 95; 231pp; English.
XX
CC   The invention relates to a composition (I) for the diagnosis, staging,
CC   treatment or for the monitoring of treatment of a subject with a brain
CC   disorder. (I) comprises several cDNAs that are differentially expressed
CC   in brain disorders and chosen from any one of 138 nucleotide sequences of
CC   AEA43798-AEA43935, or their complements. Also described: (1) a high
CC   throughput method for detecting differential expression of one or more
CC   cDNAs in a sample containing nucleic acids; (2) an isolated cDNA (II)
CC   selected from AEA43812, AEA43813, AEA43830, AEA43831, AEA43855, AEA43856,
CC   AEA43883 and AEA43923; (3) an expression vector (III) containing (II);
CC   (4) a host cell (IV) containing (III); (5) a protein (V) produced using
CC   (IV); and (6) a pharmaceutical composition comprising (V). (I) is useful
CC   for a high throughput method of screening a library of molecules or
CC   compounds to identify a ligand which specifically binds a cDNA, where the
CC   method involves combining (I) with the library of molecules or compound
CC   under conditions to allow specific binding and detecting specific binding
CC   between each cDNA and a molecule or compound. (IV) is useful for
CC   producing a protein, which involves culturing (IV) under conditions for
CC   the expression of the protein and recovering the protein from the
CC   culture. (V) is useful for high throughput method for screening a library
CC   of molecules or compounds to identify a ligand which specifically binds
CC   (V), where the method involves combining (V) or its portion with the
CC   library of molecules or compound under conditions to allow specific
CC   binding and detecting specific binding between (V) and a molecule or
CC   compound. (V) is useful for purifying a ligand from a sample, which
CC   involves combining (V) or its portion with the sample under conditions to
CC   allow specific binding, recovering the bound protein and separating the
CC   protein from ligand. (V) is also useful for producing an antibody, which
CC   involves immunizing an animal with (V) or its portion under conditions to
CC   elicit an antibody response, isolating animal antibodies and screening
CC   the isolated antibodies with the protein. (II) is useful in gene therapy
CC   for the treatment or prevention of conditions and disorders associated
CC   with immune response. The present sequence represents a human cDNA
CC   sequence which is differentially expressed in brain tissues, which is
CC   used in the exemplification of the present invention.
XX
SQ   Sequence 7390 BP; 1545 A; 1962 C; 2026 G; 1537 T; 0 U; 320 Other;

Alignment Scores:
Pred. No.:                653          Length:          7390
Score:                    64.00        Matches:         17
Percent Similarity:       45.7%        Conservative:    4
Best Local Similarity:    37.0%        Mismatches:      17
Query Match:              27.9%        Indels:          8
DB:                       14           Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x AEA43892 (1-7390)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        332 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 391

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||       ||||||||||||   |||   |||:::   |||            :::
Db        392 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 451

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        452 GAGTCGCCCTCGCACCCG 469

RESULT 57
ABQ93451
ID   ABQ93451 standard; cDNA; 7666 BP.
XX
AC   ABQ93451;
XX
DT   14-OCT-2002  (first entry)
XX
DE   Human cDNA SEQ ID NO 164.
XX
```

```
KW   Human; vulnerary; dermatological; neuroprotective; nootropic; cancer;
KW   antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
KW   antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
KW   burn; central nervous system disorder; Alzheimer's disease;
KW   Parkinson's disease; Huntington's disease; immune disorder;
KW   autoimmune disorder; multiple sclerosis; diabetes; allergy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200218424-A2.
XX
PD   07-MAR-2002.
XX
PF   31-AUG-2001; 2001WO-US027093.
XX
PR   01-SEP-2000; 2000US-00654935.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Asundi V,  Zhou P,  Xue AJ,  Ren F,  Zhang J,  Wang J;
PI   Zhao QA,  Wang D,  Liu C,  Drmanac RT,  Wehrman T;
XX
DR   WPI; 2002-583321/62.
DR   P-PSDB; ABP62972.
XX
PT   New polynucleotide and polypeptides, useful for treatment and diagnosis
PT   of Alzheimer's, Parkinson's, Huntington's, amylotrophic lateral
PT   sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
PT   sclerosis, diabetes and allergies.
XX
PS   Claim 22; SEQ ID NO 164; 284pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (I) comprising one of
CC   245 sequences (ABQ93288-ABQ93532). Treating a condition comprising
CC   administering to a mammalian subject a composition comprising the protein
CC   (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
CC   (I), (II) and (III) are useful for diagnostic evaluation of disorders.
CC   (I) is useful for gene therapy of diseases and (II) can be used for
CC   therapeutic treatment. Diseases that may be treated include wound healing
CC   and tissue repair, burns, central nervous system disorders (e.g.
CC   Alzheimer's, Parkinson's, Huntington's and amylotrophic lateral
CC   sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
CC   sclerosis, diabetes and allergies. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 7666 BP; 1641 A; 2138 C; 2176 G; 1711 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                682             Length:          7666
Score:                    64.00           Matches:         17
Percent Similarity:       45.7%           Conservative:    4
Best Local Similarity:    37.0%           Mismatches:      17
Query Match:              27.9%           Indels:          8
DB:                       6               Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x ABQ93451 (1-7666)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||      ||||||||||   |||   |||:::   |||         :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        806 GAGTCGCCCTCGCACCCG 823

RESULT 58
ABK83454
ID   ABK83454 standard; cDNA; 7758 BP.
XX
AC   ABK83454;
XX
DT   14-AUG-2002  (first entry)
XX
DE   Human cDNA differentially expressed in granulocytic cells #25.
XX
KW   Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW   viral infection; parasitic infection; protozoal infection;
KW   fungal infection; sterile inflammatory disease; psoriasis;
KW   rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW   cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW   adult respiratory distress syndrome; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; periodontal disease;
KW   granulocyte activation; chronic inflammation; allergy.
XX
OS   Homo sapiens.
XX
PN   WO200228999-A2.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US030821.
XX
PR   03-OCT-2000; 2000US-0237189P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Beazer-Barclay Y,  Weissman SM,  Yamaga S,  Vockley J;
XX
DR   WPI; 2002-435328/46.
XX
PT   Detecting granulocyte activation by detecting differential expression of
PT   genes associated with granulocyte activation, which serves as diagnostic
PT   markers that is useful for monitoring disease states and drug toxicity.
XX
PS   Claim 1; SEQ ID NO 25; 114pp; English.
XX
CC   The invention relates to detecting (M1) granulocyte (GC) activation
CC   (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC   DNA chip analysis as given in the specification, and comparing the
CC   expression level to an expression level in an unactivated GC, where
CC   differential expression of Gs is indicative of GCA. Also included are
CC   modulating (M2) GA by contacting GC with an agent that alters the
CC   expression of at least one gene in Gs; (2) screening (M3) for an agent
CC   capable of modulating GCA or an inflammation (especially chronic) in a
CC   tissue, an allergic response in a subject, exposure of a subject to a
CC   pathogen or sterile inflammatory disease using the gene expression
CC   profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC   tissue, an allergic response in a subject, exposure of a subject to a
CC   pathogen or sterile inflammatory disease, by detecting the level of
CC   expression in a sample of the tissue of gene(s) from Gs, where the level
CC   of expression of the gene is indicative of inflammation; (4) treating
CC   (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC   response in a subject, exposure of a subject to a pathogen or sterile
CC   inflammatory disease, by contacting a tissue having inflammation with an
CC   agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC   is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC   for screening an agent capable of modulating GCA preferably in an
CC   inflammation in a tissue; M4 is useful for detecting an inflammation
CC   (especially chronic) in a tissue, an allergic response in a subject,
CC   exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC   psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC   cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC   respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC   disease, ulcerative colitis, periodontal disease; also bacterial
CC   infection, viral infection, parasitic infection, protozoal infection,
CC   fungal infection and M5 is useful for treating one of the above
CC   conditions. The present sequence represents a gene differentially
CC   expressed in granulocytes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
```

```
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 7758 BP; 1666 A; 2155 C; 2209 G; 1728 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              692          Length:         7758
Score:                  64.00        Matches:        17
Percent Similarity:     45.7%        Conservative:   4
Best Local Similarity:  37.0%        Mismatches:     17
Query Match:            27.9%        Indels:         8
DB:                     6            Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x ABK83454 (1-7758)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||      ||||||||||||   |||   |||:::   |||            :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        806 GAGTCGCCCTCGCACCCG 823

RESULT 59
ADH29027
ID   ADH29027 standard; DNA; 7758 BP.
XX
AC   ADH29027;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human chronic myclogenous leukaemia (CML) gene marker #295.
XX
KW   ds; chronic phase chronic myelogenous leukaemia; CP-CML;
KW   blast crisis CML; BC-CML; human; chronic myclogenous leukaemia;
KW   gene marker.
XX
OS   Homo sapiens.
XX
PN   US2003104426-A1.
XX
PD   05-JUN-2003.
XX
PF   14-JUN-2002; 2002US-00171581.
XX
PR   18-JUN-2001; 2001US-0298914P.
XX
PA   (LINS/) LINSLEY P S.
PA   (MAOM/) MAO M.
PA   (DAIH/) DAI H.
PA   (HEYY/) HE Y.
PA   (RADI/) RADICH J P.
XX
PI   Linsley PS,  Mao M,  Dai H,  He Y,  Radich JP;
XX
DR   WPI; 2003-787046/74.
XX
PT   Classifying cell sample as chronic phase chronic myelogenous leukemia or
PT   blast crisis chronic myelogenous leukemia by detecting difference in
PT   expression of genes corresponding to the markers such as X15415, U89436.
XX
PS   Disclosure; SEQ ID NO 295; 31pp; English.
XX
CC   The invention relates to a method of classifying a cell sample as chronic
CC   phase chronic myelogenous leukaemia (CP-CML) or blast crisis CML (BC-
CC   CML). The method is useful for classifying a sample as CP-CML or BC-CML.
CC   The present sequence represents a human chronic myclogenous leukaemia
CC   (CML) gene marker used to distinguish blast crisis CML from chronic phase
CC   CML.
XX
SQ   Sequence 7758 BP; 1666 A; 2155 C; 2209 G; 1728 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              692          Length:         7758
Score:                  64.00        Matches:        17
Percent Similarity:     45.7%        Conservative:   4
Best Local Similarity:  37.0%        Mismatches:     17
Query Match:            27.9%        Indels:         8
DB:                     10           Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x ADH29027 (1-7758)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||      ||||||||||||   |||   |||:::   |||            :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        806 GAGTCGCCCTCGCACCCG 823

RESULT 60
ADN95314
ID   ADN95314 standard; DNA; 7758 BP.
XX
AC   ADN95314;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Human BEC/LEC-related gene sequence SeqID236.
XX
KW   growth; differentiation; blood endothelial cell; BEC;
KW   lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
KW   lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
KW   vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;
KW   inflammatory disease; cancer metastasis; lymphatic system; gene; ds;
KW   human.
XX
OS   Homo sapiens.
XX
PN   WO2003080640-A1.
XX
PD   02-OCT-2003.
XX
PF   07-MAR-2003; 2003WO-US006900.
XX
PR   07-MAR-2002; 2002US-0363019P.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (LICN ) LICENTIA LTD.
XX
PI   Alitalo K,  Makinen T,  Petrova T,  Saharinen P,  Saharinen J;
XX
DR   WPI; 2003-876899/81.
DR   P-PSDB; ADN95313.
XX
PS   Claim 23; SEQ ID NO 236; 176pp; English.
XX
CC   This invention relates to a method of differentially modulating the
CC   growth or differentiation of blood endothelial cells (BEC) or lymphatic
CC   endothelial cells (LEC) comprises contacting endothelial cells with a
CC   composition comprising an agent that differentially modulates blood or
CC   lymphatic endothelial cells. Treating hereditary lymphoedema comprises
CC   identifying a human subject with lymphoedema and with a mutation in at
CC   least one allele of a gene encoding a LEC protein, where the mutation
CC   correlates with lymphoedema in human subjects, and with the proviso that
CC   the LEC protein is not VEGFR-3; and administering to the subject a
CC   composition comprising a lymphatic growth agent selected from VEGF-C or
CC   VEGF-D polypeptides and polynucleotides. The invention may be useful for
CC   the development of compounds with an antiangiogenic, cytostatic,
```

```
CC   vasotropic or antiinflammatory activity or for gene therapy. The method
CC   is useful in modulating the growth or differentiation of blood
CC   endothelial cells or lymphatic endothelial cells, in treating hereditary
CC   lymphoedema, in screening for an endothelial cell disorder or
CC   predisposition to the disorder or in monitoring the efficacy or toxicity
CC   of a drug on endothelial cells. The agent is useful in manufacturing a
CC   medicament for the differential modulation of blood vessel endothelial
CC   cell or lymphatic vessel endothelial cell growth or differentiation. The
CC   lymphatic growth agent may also be used in manufacturing a medicament for
CC   the treatment of hereditary lymphoedema resulting from a mutation in a
CC   LEC gene or of other diseases involving the lymphatic vessels, such as
CC   various inflammatory diseases and cancer metastasis via the lymphatic
CC   system. The present sequence is that of a human LEC/BEC differentially
CC   expressed gene which is related to the method of the invention.Note: This
CC   sequence does not appear in the specification but was obtained by the
CC   indexer using the source data given in table 14 of the specification.
XX
SQ   Sequence 7758 BP; 1666 A; 2155 C; 2209 G; 1728 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                692             Length:         7758
Score:                    64.00           Matches:        17
Percent Similarity:       45.7%           Conservative:   4
Best Local Similarity:    37.0%           Mismatches:     17
Query Match:              27.9%           Indels:         8
DB:                       11              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x ADN95314 (1-7758)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||       ||||||||||||   |||   |||:::   |||            :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        806 GAGTCGCCCTCGCACCCG 823

RESULT 61
ADQ18420
ID   ADQ18420 standard; DNA; 7758 BP.
XX
AC   ADQ18420;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 1239.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 1239; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 7758 BP; 1666 A; 2155 C; 2209 G; 1728 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                692             Length:         7758
Score:                    64.00           Matches:        17
Percent Similarity:       45.7%           Conservative:   4
Best Local Similarity:    37.0%           Mismatches:     17
Query Match:              27.9%           Indels:         8
DB:                       12              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x ADQ18420 (1-7758)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||       ||||||||||||   |||   |||:::   |||            :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        806 GAGTCGCCCTCGCACCCG 823

RESULT 62
ADQ22923
ID   ADQ22923 standard; DNA; 7792 BP.
XX
AC   ADQ22923;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 5743.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
```

```
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 5743; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 7792 BP; 1693 A; 2149 C; 2203 G; 1721 T; 0 U; 26 Other;

Alignment Scores:
Pred. No.:                696          Length:          7792
Score:                    64.00        Matches:         17
Percent Similarity:       45.7%        Conservative:    4
Best Local Similarity:    37.0%        Mismatches:      17
Query Match:              27.9%        Indels:          8
DB:                       12           Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x ADQ22923 (1-7792)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||       ||||||||||||   |||   |||:::   |||           :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        806 GAGTCGCCCTCGCACCCG 823

RESULT 63
AAA58471/c
ID   AAA58471 standard; DNA; 58857 BP.
XX
AC   AAA58471;
XX
DT   31-OCT-2000  (first entry)
XX
DE   Nucleotide sequence of the bleomycin (BLM) gene cluster ORFs 8-30.
XX
KW   BLM gene cluster; bleomycin gene cluster; polyketide metabolite;
KW   bleomycin; bleomycin analogue; holo-carrier protein; thiazolidine;
KW   thiazoline; bithiazoline; microbial metabolite; sugar; ss.
XX
OS   Streptomyces verticillus.
XX
FH   Key             Location/Qualifiers
FT   CDS             223. .564
FT                   /*tag=  a
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 30; encodes AAB07556"
FT   CDS             561. .2309
FT                   /*tag=  b
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 29; encodes AAB07557"
FT   CDS             2767. .3486
FT                   /*tag=  c
FT                   /note= "ORF 28; encodes AAB07558"
FT   CDS             3527. .5593
FT                   /*tag=  d
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 27; encodes AAB07559"
FT   CDS             5806. .12294
FT                   /*tag=  e
FT                   /note= "ORF 26; encodes AAB07560"
FT   CDS             12291. .15491
FT                   /*tag=  f
FT                   /note= "ORF 25; encodes AAB07561"
FT   CDS             15488. .21013
FT                   /*tag=  g
FT                   /note= "ORF 24; encodes AAB07562"
FT   CDS             21010. .24666
FT                   /*tag=  h
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 23; encodes AAB07563"
FT   CDS             24663. .32690
FT                   /*tag=  i
FT                   /note= "ORF 22; encodes AAB07564"
FT   CDS             32893. .34830
FT                   /*tag=  j
FT                   /note= "ORF 21; encodes AAB07565"
FT   CDS             34827. .35804
FT                   /*tag=  k
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 20; encodes AAB07566"
FT   CDS             35818. .37302
FT                   /*tag=  l
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 19; encodes AAB07567"
FT   CDS             37299. .39215
FT                   /*tag=  m
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 18; encodes AAB07568"
FT   CDS             39301. .47181
FT                   /*tag=  n
FT                   /note= "ORF 17; encodes AAB07569"
FT   CDS             47178. .49985
FT                   /*tag=  o
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 16; encodes AAB07570"
FT   CDS             49982. .51001
FT                   /*tag=  p
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 15; encodes AAB07571"
FT   CDS             50998. .52386
FT                   /*tag=  q
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 14; encodes AAB07572"
FT   CDS             52383. .52946
FT                   /*tag=  r
FT                   /note= "ORF 13; encodes AAB07573"
FT   CDS             53018. .54190
FT                   /*tag=  s
FT                   /note= "ORF 12; encodes AAB07574"
FT   CDS             54187. .55824
FT                   /*tag=  t
FT                   /note= "ORF 11; encodes AAB07575"
FT   CDS             55821. .56093
FT                   /*tag=  u
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 10; encodes AAB07576"
FT   CDS             56090. .57586
FT                   /*tag=  v
FT                   /transl_except= (pos: 1. .3, aa: Met)
FT                   /note= "ORF 9; encodes AAB07577"
FT   CDS             57583. .58857
FT                   /*tag=  w
FT                   /note= "ORF 8; encodes AAB07578"
XX
PN   WO200040704-A1.
XX
PD   13-JUL-2000.
XX
```

```
PF   06-JAN-2000; 2000WO-US000445.
XX
PR   06-JAN-1999;   99US-0115435P.
PR   05-FEB-1999;   99US-0118848P.
PR   05-JAN-2000; 2000US-00477962.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Shen B,  Du L,  Sanchez C,  Chen M,  Edwards DJ;
XX
DR   WPI; 2000-465974/40.
DR   P-PSDB; AAB07556, AAB07557, AAB07558, AAB07559, AAB07560, AAB07561,
DR   AAB07562, AAB07563, AAB07564, AAB07565, AAB07566, AAB07567, AAB07568,
DR   AAB07569, AAB07570, AAB07571, AAB07572, AAB07573, AAB07574, AAB07575,
DR   AAB07576, AAB07577, AAB07578.
XX
PT   New bleomycin gene cluster components useful for peptide and/or
PT   polyketide metabolites, especially bleomycin, production and for
PT   chemically modifying biological molecules.
XX
PS   Claim 8; Page 97-136; 162pp; English.
XX
CC   The present sequence represents the BLM (Bleomycin) gene cluster,
CC   containing open reading frames (ORFs) 8-30. The proteins encoded by the
CC   gene cluster are useful for producing peptides and/or polyketide
CC   metabolites, especially bleomycin or bleomycin analogues. They are also
CC   useful for chemically modifying biological molecules to produce branched
CC   methyl groups, and for coupling amino acids and fatty acids. They may be
CC   reacted with an apo-carrier protein and coenzyme A to produce a holo-
CC   carrier protein. The BLM gene cluster or catalytic domains can be used
CC   individually or collectively to produce thiazolidine, thiazoline,
CC   bithiazoline and bithiazoline-containing microbial metabolites. The BLM
CC   gene cluster may also be used to produce sugars
XX
SQ   Sequence 58857 BP; 7256 A; 25139 C; 19353 G; 7109 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                8.06e+03        Length:          58857
Score:                    64.00           Matches:         14
Percent Similarity:       52.5%           Conservative:    7
Best Local Similarity:    35.0%           Mismatches:      15
Query Match:              27.9%           Indels:          4
DB:                       3               Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x AAA58471 (1-58857)

Qy          1 GlyAspProGlySerGluIleIleGluSer------------ValProProAlaGlyPro 16
              ||||||   |||   |||:::::::::                ||||||   |||
Db      34476 GGCGACGGCGGCGGGGAAGTTCTCGACGACCATGATGCTGTCGAACAGCCCGCCGCCCCG 34417

Qy         17 GluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGly 36
                 ||||||:::      |||      |||   |||::::::   |||   :::|||
Db      34416 GCCCAGTCCGAGACCGCGCTCGACGCGGCTGAGCGGGACGTGCGCGTGGTCCCACAGGGC 34357

RESULT 64
ADX56092/c
ID   ADX56092 standard; DNA; 104096 BP.
XX
AC   ADX56092;
XX
DT   21-APR-2005  (first entry)
XX
DE   Streptomyces nanchangensis nanoligomycin gene cluster SeqID1.
XX
KW   nanoligomycin; agriculture; pharmaceutical; protein production; gene; ds.
XX
OS   Streptomyces nanchangensis.
XX
FH   Key             Location/Qualifiers
FT   CDS             255. .1208
FT                   /*tag=  a
FT                   /product= "nlmTI"
FT   CDS             2168. .4990
FT                   /*tag=  b
FT                   /product= "nlmRI"
FT                   /note= "This sequence contains translation exceptions"
FT   CDS             5469. .23915
FT                   /*tag=  c
FT                   /product= "nlmA1"
FT                   /note= "This sequence contains numerous translation
FT                   exceptions"
FT   CDS             23938. .38337
FT                   /*tag=  d
FT                   /product= "nlmA2"
FT   CDS             38626. .50133
FT                   /*tag=  e
FT                   /product= "nlmA3"
FT   CDS             50289. .64196
FT                   /*tag=  f
FT                   /product= "nlmA6"
FT                   /note= "This sequence contains translation exceptions"
FT   CDS             64243. .75024
FT                   /*tag=  g
FT                   /product= "nlmA7"
FT                   /note= "This sequence contains translation exceptions"
FT   CDS             complement(75191. .76408)
FT                   /*tag=  h
FT                   /product= "nlmB"
FT   CDS             complement(76564. .82170)
FT                   /*tag=  i
FT                   /product= "nlmA5"
FT   CDS             complement(82242. .93935)
FT                   /*tag=  j
FT                   /product= "nlmA4"
FT                   /partial
FT                   /note= "No start codon"
FT   CDS             94138. .94641
FT                   /*tag=  k
FT                   /product= "nlmOI"
FT   CDS             94705. .96048
FT                   /*tag=  l
FT                   /product= "ccrA"
FT   CDS             complement(96412. .98193)
FT                   /*tag=  m
FT                   /product= "nlmTII"
FT                   /partial
FT                   /note= "No start codon"
XX
PN   CN1523034-A.
XX
PD   25-AUG-2004.
XX
PF   11-SEP-2003; 2003CN-00150923.
XX
PR   11-SEP-2003; 2003CN-00150923.
XX
PA   (UYSH-) UNIV SHANGHAI JIAOTONG.
XX
PI   Deng Z,  Sun Y,  Zhou X;
XX
DR   WPI; 2004-797302/79.
DR   P-PSDB; ADX56093, ADX56094, ADX56095, ADX56096, ADX56097, ADX56098,
DR   ADX56099, ADX56100, ADX56101, ADX56102, ADX56103, ADX56104, ADX56105.
XX
PT   Southern oligomycin biosynthetic gene cluster.
XX
PS   Claim 1; SEQ ID NO 1; 132pp; Chinese.
XX
CC   The present invention relates to a novel nanoligomycin biological
CC   synthetic gene cluster, belonging to the field of gene technology. The
CC   whole nanoligomycin biological synthetic gene cluster has 13 genes
CC   including polyketone synthase genes (7 genes of nlm A1, nlmA2, nlmA3,
CC   nlmA4, nlm A5, nlmA6 and nlmA7), modification genes of nanoligomycin (2
CC   genes of nlm B and nlmOI), nanoligomycin transposase genes (2 genes of
CC   nlmTI and nlmTII), a regulating gene of nanoligomycin (nlmRI) and a
CC   nanoligomycin precursor synthetic gene (ccrA). The genes, proteins and
```

```
CC   antibodies provided by the invention also can be used to seek and develop
CC   compounds and proteins which can be used in medicine, industry and
CC   agriculture. The present sequence is that of the Streptomyces
CC   nanchangensis nanoligomycin gene cluster of the invention.
XX
SQ   Sequence 104096 BP; 13946 A; 36244 C; 38265 G; 15641 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.61e+04      Length:         104096
Score:                  64.00         Matches:        16
Percent Similarity:     52.3%         Conservative:   7
Best Local Similarity:  36.4%         Mismatches:     11
Query Match:            27.9%         Indels:         10
DB:                     13            Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x ADX56092 (1-104096)

Qy          7 IleIleGluSerValProProAla--------------GlyProGluAlaSerGluSer 21
              :::|||   :::   |||||||||               |||      |||:::|||
Db      45290 GTCATCACGGCGACACCGCCGGCGAGGGCAAGGTCGCAGTCGCCCGCGCGCAGTGACAGC 45231

Qy         22 --------------ThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGly 36
                           |||:::|||:::|||     :::      |||:::   |||
Db      45230 GCGGCGGTGTGCAGGGCGACCAACGAGGACGAGCAGGCCGTGTCCACGGTGACGGCGGGC 45171

Qy         37 ProSerArgPro 40
              ||||||||||||
Db      45170 CCCTCCAGGCCC 45159

RESULT 65
ADT05874
ID   ADT05874 standard; DNA; 171096 BP.
XX
AC   ADT05874;
XX
DT   13-JAN-2005  (first entry)
XX
DE   Cercopithecine herpesvirus 15 complete genome.
XX
KW   Epstein-Barr virus; ds; African green Monkey herpes virus;
KW   infectious mononucleosis; glandular fever; Burkitt's lymphoma;
KW   Hodgkin's disease; post-transplant lymphoproliferative disorder; BZLF1;
KW   ZEBRA; viral load; Epstein-Barr virus infection; immune selection.
XX
OS   Cercopithecine herpesvirus 15; Strain LCL8664.
XX
PN   US2004214161-A1.
XX
PD   28-OCT-2004.
XX
PF   02-JUL-2003; 2003US-00612869.
XX
PR   24-APR-2003; 2003GB-00009311.
XX
PA   (SMIT/) SMITH M.
XX
PI   Smith M;
XX
DR   WPI; 2004-765595/75.
XX
PT   Detection of Epstein Barr virus nucleic acid in isolated sample, by
PT   contacting sample with probe capable of binding to BZLF 1 gene sequence
PT   of Epstein Barr virus and producing signal, and detecting signal
PT   produced.
XX
PS   Disclosure; SEQ ID NO 8; 96pp; English.
XX
CC   The invention relates to detection of Epstein Barr virus nucleic acid in
CC   an isolated sample, involving contacting the sample with a probe, where
CC   the probe binds to a target region having a BZLF1 gene (encoding the
CC   ZEBRA protein), or its homologue or its complementary strand, which
CC   binding provides a detectable signal, and detecting the signal. Also
CC   included are a probe suitable for use in the method, a pair of
CC   oligonucleotide primers comprising forward and reverse primers for use in
CC   the method, a kit for detection of Epstein Barr virus nucleic acid,
CC   comprising the primers and probes and a DNA array comprising an
CC   immobilised nucleic acid probe that binds to the target region. The
CC   method is useful for detecting/and or quantifying Epstein Barr virus
CC   nucleic acid in an isolated sample. The probe is useful for quantifying
CC   Epstein Barr viral load in a first isolated sample and in an in vitro
CC   method of monitoring drug efficacy for alleviating Epstein Barr virus
CC   infection or an Epstein Barr virus induced medical condition (e.g.
CC   infectious mononucleosis (glandular fever), Burkitt's lymphoma, Hodgkin's
CC   disease, and post-transplant lymphoproliferative disorders). In the
CC   method, the target nucleic acid sequence proves to be beneficial as it
CC   does not encode a surface protein and thus are not susceptible to immune
CC   selection pressure. The present sequence is the genome of African Green
CC   monkey herpes virus (an Epstein Barr-type virus) which contains a
CC   homologue of BZLF1.
XX
SQ   Sequence 171096 BP; 31572 A; 53211 C; 52770 G; 33543 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.93e+04      Length:         171096
Score:                  64.00         Matches:        19
Percent Similarity:     53.1%         Conservative:   7
Best Local Similarity:  38.8%         Mismatches:     14
Query Match:            27.9%         Indels:         9
DB:                     13            Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x ADT05874 (1-171096)

Qy          1 GlyAspProGlySerGluIleIle------GluSerValProProAlaGlyProGluAla 18
              ||||||||||||||||              |||:::|||   |||          |||
Db      76730 GGTGATCCCGGGTCTAATCCCAGTCCACGTGAAACTGTGTCCCCGGGACCGTCTACTGCC 76789

Qy         19 SerGluSer-----------------ThrThrAspGluAsnGluAspAspIleGlnPhe 32
              ||||||                  :::   ::::::||||||::::::
Db      76790 AGCGAGCAAATGTTGCAACACGAGCCCATGGAAGAGAATTCAGATGATGACGTGGAAGAA 76849

Qy         33 ValSerGlu-GlyProSerArgPro 40
              ||||||::: ||||||   ||||||
Db      76850 GTGAGCCAACGGGCCAGAACGCCCC 76874

RESULT 66
ADV99887
ID   ADV99887 standard; DNA; 132544 BP.
XX
AC   ADV99887;
XX
DT   24-FEB-2005  (first entry)
XX
DE   Nanchangmycin biosynthetic gene cluster.
XX
KW   gene cluster; gene; ds.
XX
OS   Streptomyces nanchangensis.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement(682. .2052)
FT                   /*tag=  a
FT                   /product= "NanT1"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             2064. .2879
FT                   /*tag=  b
FT                   /product= "NanT2"
FT   CDS             3117. .3818
FT                   /*tag=  c
FT                   /product= "NanT3"
FT   CDS             complement(3724. .4923)
FT                   /*tag=  d
FT                   /product= "NanT4"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(4916. .6106)
FT                   /*tag=  e
FT                   /product= "NanT5"
```

```
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(6768. .7493)
FT                   /*tag= f
FT                   /product= "NanR1"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(7573. .8334)
FT                   /*tag= g
FT                   /product= "NanR2"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             8919. .17627
FT                   /*tag= h
FT                   /product= "NanA1"
FT   CDS             17642. .24313
FT                   /*tag= i
FT                   /product= "NanA2"
FT   CDS             24310. .36408
FT                   /*tag= j
FT                   /product= "NanA3"
FT   CDS             36429. .48299
FT                   /*tag= k
FT                   /product= "NanA4"
FT   CDS             48345. .60284
FT                   /*tag= l
FT                   /product= "NanA5"
FT   CDS             60305. .65302
FT                   /*tag= m
FT                   /product= "NanA6"
FT   CDS             complement(65365. .66747)
FT                   /*tag= n
FT                   /product= "NanG5"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             66881. .67798
FT                   /*tag= o
FT                   /product= "NanM"
FT   CDS             67823. .68863
FT                   /*tag= p
FT                   /product= "NanG4"
FT   CDS             68860. .70164
FT                   /*tag= q
FT                   /product= "NanG3"
FT   CDS             complement(70145. .71140)
FT                   /*tag= r
FT                   /product= "NanG2"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(71137. .72045)
FT                   /*tag= s
FT                   /product= "NanG1"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(72050. .76990)
FT                   /*tag= t
FT                   /product= "NanA7"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(76992. .77864)
FT                   /*tag= u
FT                   /product= "NanE"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(77949. .78263)
FT                   /*tag= v
FT                   /product= "NanA10"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(78327. .79763)
FT                   /*tag= w
FT                   /product= "NanO"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(79828. .80769)
FT                   /*tag= x
FT                   /product= "NanI"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(80784. .83192)
FT                   /*tag= y
FT                   /product= "NanA9"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(83196. .89759)
FT                   /*tag= z
FT                   /product= "NanA11"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(89771. .100138)
FT                   /*tag= aa
FT                   /product= "NanA8"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             complement(100196. .101467)
FT                   /*tag= ab
FT                   /product= "NanP"
FT                   /note= "Encoded on the reverse complement strand"
FT   CDS             101593. .102600
FT                   /*tag= ac
FT                   /product= "NanR3"
FT   CDS             102600. .103541
FT                   /*tag= ad
FT                   /product= "NanR4"
XX
PN   CN1431220-A.
XX
PD   23-JUL-2003.
XX
PF   16-JAN-2003; 2003CN-00114920.
XX
PR   16-JAN-2003; 2003CN-00114920.
XX
PA   (UYSH-) UNIV SHANGHAI JIAOTONG.
XX
PI   Deng Z,  Sun Y,  Zhou X;
XX
DR   WPI; 2004-110066/12.
DR   P-PSDB; ADV99888, ADV99889, ADV99890, ADV99891, ADV99892, ADV99893,
DR   ADV99894, ADV99895, ADV99896, ADV99897, ADV99898, ADV99899, ADV99900,
DR   ADV99901, ADV99902, ADV99903, ADV99904, ADV99905, ADV99906, ADV99907,
DR   ADV99908, ADV99909, ADV99910, ADV99911, ADV99912, ADV99913, ADV99914,
DR   ADV99915, ADV99916, ADV99917.
XX
PT   Biosynthesis gene cluster of Nanchang.
XX
PS   Disclosure; SEQ ID NO 1; 338pp; Chinese.
XX
CC   The invention describes a biosynthetic gene cluster of Nanchangmycin
CC   containing 30 genes including 11 polyketone synthase genes, 4 modifying
CC   genes, 6 biosynthetic genes of its deoxysugar, and 9 regulating genes.
CC   This sequence represents the Nanchangmycin biosynthetic gene cluster of
CC   the invention.
XX
SQ   Sequence 132544 BP; 18270 A; 49105 C; 47228 G; 17941 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.52e+04        Length:          132544
Score:                    63.50           Matches:         14
Percent Similarity:       51.3%           Conservative:    6
Best Local Similarity:    35.9%           Mismatches:      18
Query Match:              27.7%           Indels:          1
DB:                       13              Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x ADV99887 (1-132544)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu---Ser 21
              |||   :::          |||       ||||||       |||       :::|||   |||
Db      94749 CCACCCGCCAACGCCAGCGAGCACTCACCGCCGCGCAGCCCCCGCACCGCCGAGTGAAGC 94808

Qy         22 ThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                |||   ::::::|||      :::      |||:::   |||||||||   |||
Db      94809 GCCACCAGCGACGACGAACACCCCGTGTCCACCGTCACCGCGGGACCCTCCAGCCCC 94865

RESULT 67
ADM87098
ID   ADM87098 standard; cDNA; 3120 BP.
XX
AC   ADM87098;
XX
```

DT 03-JUN-2004 (first entry)
XX
DE Human protein encoding cDNA SEQ ID NO:191.
XX
KW respiratory; cytostatic; antiarthritic; antiinflammatory;
KW gastrointestinal; antibacterial; immunosuppressive; antidiabetic;
KW antirheumatic; gene therapy; molecular weight marker; chromosome marker;
KW chromosome tag; genetic fingerprinting; nutritional supplement; cancer;
KW inflammatory condition; arthritis; inflammatory bowel disease;
KW Crohn's disease; sepsis; rheumatoid arthritis; diabetes mellitus type 1;
KW graft versus host disease; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004009834-A2.
XX
PD 29-JAN-2004.
XX
PF 19-JUL-2002; 2002WO-US022858.
XX
PR 21-JUL-2001; 2001US-0306971P.
PR 28-MAR-2002; 2002US-00112944.
XX
PA (NUVE-) NUVELO INC.
XX
PI Tang YT, Yang Y, Weng G, Zhang J, Ren F, Xue A, Wang J;
PI Wehrman T, Ghosh MJ, Wang D, Zhao QA, Wang Z;
XX
DR WPI; 2004-143291/14.
DR P-PSDB; ADM87342.
XX
PT New isolated polynucleotides and polypeptides, useful for treating, e.g.
PT cancer, lung or liver fibrosis, arthritis, inflammatory bowel disease,
PT Crohn's disease, rheumatoid arthritis, diabetes mellitus type 1 or graft
PT versus host disease.
XX
PS Claim 1; SEQ ID NO 191; 591pp; English.
XX
CC The present invention describes an isolated polynucleotide (I): (a)
CC comprising a nucleotide sequence selected from SEQ ID NO:1-244; or (b)
CC which encodes a polypeptide with biological activity, where the
CC polynucleotide hybridises to (I) under stringent hybridisation conditions
CC or has greater than 99% sequence identity with (I). (I) has respiratory,
CC cytostatic, antiarthritic, antiinflammatory, gastrointestinal,
CC antibacterial, immunosuppressive, antidiabetic and antirheumatic
CC activities, and can be used in gene therapy. (I) can be used for
CC generating polynucleotides encoding chimeric or fusion proteins and
CC heterologous protein sequences. The polynucleotides can be used to
CC express recombinant protein for analysis, characterisation or therapeutic
CC use; as markers for tissues in which the corresponding protein is
CC preferentially expressed; as molecular weight markers on gels; as
CC chromosome markers or tags to identify chromosomes or to map related gene
CC positions; to compare with endogenous DNA sequences in patients to
CC identify potential genetic disorders; as probes to hybridise and discover
CC genes, related DNA sequences; as a source of information to derive PCR
CC primers for genetic fingerprinting; as a probe to subtract-out known
CC sequences in the process of discovering other novel polynucleotides; for
CC selecting and making oligomers for attachment to a gene chip or other
CC support, including for examination of expression patterns; to raise anti-
CC protein antibodies using DNA immunisation techniques; and as an antigen
CC to raise anti-DNA antibodies or elicit another immune response. The
CC polynucleotides and polypeptides can also be used as nutritional sources
CC or supplements, e.g. as a protein or amino acid supplement, as a carbon
CC source, as a nitrogen source or as a source of carbohydrates. The
CC polynucleotides and polypeptides can also be used treat cancer. The
CC compositions are useful for promoting better or faster closure of non-
CC healing wounds, for the generation and regeneration of tissues, for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, and conditions resulting from
CC systemic cytokine damage. The compositions can also be used to treat
CC inflammatory conditions (e.g. arthritis, inflammatory bowel disease or
CC Crohn's disease), sepsis, rheumatoid arthritis, diabetes mellitus type 1
CC or graft versus host disease. The present sequence represents a novel
CC human polynucleotide sequence from the present invention. N.B. The
CC sequences for this patent were obtained from the USPTO web site from an
CC equivalent US patent US20040048249A1.
XX
SQ Sequence 3120 BP; 642 A; 962 C; 908 G; 608 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	314	Length:	3120
Score:	63.00	Matches:	13
Percent Similarity:	46.5%	Conservative:	7
Best Local Similarity:	30.2%	Mismatches:	23
Query Match:	27.5%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x ADM87098 (1-3120)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||::::::|||       |||           |||         |||
Db       2812 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTCCTCCCCCAACCCTGTTGGAA 2871

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::       |||   :::|||           |||   |||:::
Db       2872 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 2931

Qy         41 ValLeuGlu 43
                :::|||
Db       2932 GAGATGGAG 2940
```

RESULT 68
AAT89346
ID AAT89346 standard; cDNA; 3211 BP.
XX
AC AAT89346;
XX
DT 11-MAR-1998 (first entry)
XX
DE Human p160 cDNA 160.2.
XX
KW p160; p62; cytoplasmic; T cell; B cell; development; activation;
KW modulation; cellular response; cell proliferation; autoimmune disease;
KW p56-lck; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 439. .3156
FT /*tag= a
FT /product= "p160"
FT /note= "160.2"
XX
PN WO9722255-A1.
XX
PD 26-JUN-1997.
XX
PF 11-DEC-1996; 96WO-US019944.
XX
PR 19-DEC-1995; 95US-00574959.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Shin J, Joung I, Vadlamudi RK, Strominger JL;
XX
DR WPI; 1997-341351/31.
DR P-PSDB; AAW31186.
XX
PT cDNA encoding p62 and p160 and corresponding proteins - used in the
PT treatment of autoimmune disease and for T and B cell proliferation, e.g.
PT for treatment of tumours.
XX
PS Claim 82; Fig 10; 175pp; English.
XX
CC This cDNA sequence encodes a novel p160 (160.2) which is capable of
CC activating transcription of a variety of genes upon activation of p62 and

```
CC   is capable of binding to the p62/p56lpk complex to modulate Lck function
CC   in a manner similar to p62. The genes transcribed in response to p160
CC   activation likely include those of which are involved in T or B cell
CC   development/differentiation, T or B cell activation or production of T or
CC   B cell specific factors e.g. lymphokines or antibodies. This p160
CC   polypeptide is also a substrate for serine/threonine kinase activity.
CC   p160 polypeptides can modulate degradation of cellular proteins e.g. cell
CC   cycle regulatory proteins stimulating expression of cell cycle dependent
CC   kinase inhibitors and arresting cell cycle progression at specific
CC   boundaries to thereby modulate cell proliferation. As p160 boosts B cell
CC   response it may be used to treat disorders where this is beneficial, e.g.
CC   infections by pathogenic microorganisms. p160 can be used to expand T
CC   cell populations for treating infectious diseases or cancer and p160
CC   inhibitors could reduce B or T cell responses and may be used to treat a
CC   variety of autoimmune diseases, e.g. diabetes mellitus, arthritis,
CC   multiple sclerosis allergic reactions, Crohn's diseases etc
XX
SQ   Sequence 3211 BP; 649 A; 990 C; 948 G; 624 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                325             Length:         3211
Score:                    63.00           Matches:        13
Percent Similarity:       46.5%           Conservative:   7
Best Local Similarity:    30.2%           Mismatches:     23
Query Match:              27.5%           Indels:         0
DB:                       2               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AAT89346 (1-3211)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||     |||:::::::|||     |||         |||         |||
Db       2902 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTCCTCCCCCAACCCTGTTGGAA 2961

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||||::::::     |||   :::|||         |||   |||:::
Db       2962 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3021

Qy         41 ValLeuGlu 43
                :::|||
Db       3022 GAGATGGAG 3030

RESULT 69
AAT89345
ID   AAT89345 standard; cDNA; 3901 BP.
XX
AC   AAT89345;
XX
DT   11-MAR-1998  (first entry)
XX
DE   Human p160 cDNA 160.1.
XX
KW   p160; p62; cytoplasmic; T cell; B cell; development; activation;
KW   modulation; cellular response; cell proliferation; autoimmune disease;
KW   p56-lck; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             439. .3846
FT                   /*tag=  a
FT                   /product= "p160"
FT                   /note= "160.1"
XX
PN   WO9722255-A1.
XX
PD   26-JUN-1997.
XX
PF   11-DEC-1996;   96WO-US019944.
XX
PR   19-DEC-1995;   95US-00574959.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Shin J,  Joung I,  Vadlamudi RK,  Strominger JL;
XX
DR   WPI; 1997-341351/31.
DR   P-PSDB; AAW31185.
XX
PT   cDNA encoding p62 and p160 and corresponding proteins - used in the
PT   treatment of autoimmune disease and for T and B cell proliferation, e.g.
PT   for treatment of tumours.
XX
PS   Claim 82; Fig 8; 175pp; English.
XX
CC   This cDNA sequence encodes a novel p160 (160.1) which is capable of
CC   activating transcription of a variety of genes upon activation of p62 and
CC   is capable of binding to the p62/p56lpk complex to modulate Lck function
CC   in a manner similar to p62. The genes transcribed in response to p160
CC   activation likely include those of which are involved in T or B cell
CC   development/differentiation, T or B cell activation or production of T or
CC   B cell specific factors e.g. lymphokines or antibodies. This p160
CC   polypeptide is also a substrate for serine/threonine kinase activity.
CC   p160 polypeptides can modulate degradation of cellular proteins e.g. cell
CC   cycle regulatory proteins stimulating expression of cell cycle dependent
CC   kinase inhibitors and arresting cell cycle progression at specific
CC   boundaries to thereby modulate cell proliferation. As p160 boosts B cell
CC   response it may be used to treat disorders where this is beneficial, e.g.
CC   infections by pathogenic microorganisms. p160 can be used to expand T
CC   cell populations for treating infectious diseases or cancer and p160
CC   inhibitors could reduce B or T cell responses and may be used to treat a
CC   variety of autoimmune diseases, e.g. diabetes mellitus, arthritis,
CC   multiple sclerosis allergic reactions, Crohn's diseases etc
XX
SQ   Sequence 3901 BP; 781 A; 1183 C; 1128 G; 809 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                411             Length:         3901
Score:                    63.00           Matches:        13
Percent Similarity:       46.5%           Conservative:   7
Best Local Similarity:    30.2%           Mismatches:     23
Query Match:              27.5%           Indels:         0
DB:                       2               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AAT89345 (1-3901)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||     |||:::::::|||     |||         |||         |||
Db       3592 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTCCTCCCCCAACCCTGTTGGAA 3651

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||||::::::     |||   :::|||         |||   |||:::
Db       3652 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3711

Qy         41 ValLeuGlu 43
                :::|||
Db       3712 GAGATGGAG 3720

RESULT 70
ADY72595/c
ID   ADY72595 standard; DNA; 9795 BP.
XX
AC   ADY72595;
XX
DT   02-JUN-2005  (first entry)
XX
DE   A. orientalis polyene polyketide ORF 21 DNA.
XX
KW   antibiotic; antibacterial; cytostatic; antiinflammatory;
KW   bacterial infection; cancer; inflammation; gene; ds.
XX
OS   Amycolatopsis orientalis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .9795
FT                   /*tag=  a
FT                   /product= "Polyene polyketide ORF21"
```

XX
PN WO2005024025-A1.
XX
PD 17-MAR-2005.
XX
PF 10-SEP-2004; 2004WO-CA001603.
XX
PR 11-SEP-2003; 2003US-0501821P.
PR 28-MAY-2004; 2004US-0574922P.
PR 23-JUN-2004; 2004US-0581707P.
XX
PA (ECOP-) ECOPIA BIOSCIENCES INC.
XX
PI Mcalpine JB, Farnet CM, Zazopoulos E, Sorensen D;
XX
DR WPI; 2005-233289/24.
DR P-PSDB; ADY72594.
XX
PT Novel polyene polyketide or its salt useful for treating bacterial
PT infection, cancer or acute or chronic inflammation.
XX
PS Claim 40; SEQ ID NO 44; 374pp; English.
XX
CC The invention relates to a novel polyene polyketide compound, its salt
CC having a structural formula (I), prodrugs, and derivatives which are
CC shown to have antibiotic activity. The invention further comprises: a
CC composition comprising the compound or its salt, together with a carrier;
CC a method of producing the compound, which involves culturing
CC Amycolatopsis orientalis strain in a nutrient medium comprising one or
CC more source of carbon and nitrogen atoms, and isolating the compound; an
CC A. orientalis having the IDAC Accession No. 220604-01; an isolated
CC purified or enriched nucleic acid encoding the polyketide synthase for
CC producing the compound; an expression vector comprising the nucleic acid;
CC an isolated host cell transformed with the vector; an isolated or
CC purified polyketide synthase for producing the compound; an isolated or
CC purified ketosynthase domain for producing the compound; an isolated or
CC purified acyl transferase domain for producing the compound; an isolated
CC or purified acyl carrier protein domain for producing the compound; an
CC isolated or purified dehydratase domain for producing the compound; an
CC isolated or purified ketoreductase domain for producing the compound; an
CC isolated or purified enoyl reductase domain for producing the compound;
CC an isolated or purified thioesterase domain for producing the compound; a
CC method for producing PS, involves introducing into a host cell a nucleic
CC acid encoding the polypeptide, where the nucleic acid is operably linked
CC to a promoter; and a cosmid deposited under IDAC Accession No. 050903-01,
CC 050903-02, 050903-03 or 050903-04. The novel compound and its composition
CC have antibacterial, cytostatic and antiinflammatory activities. The novel
CC compound or its composition is useful for treating bacterial infection
CC and for treating cancer or acute or chronic inflammation. This sequence
CC represents an A. orientalis polyene polyketide open reading frame derived
CC protein of the invention.
XX
SQ Sequence 9795 BP; 1225 A; 3550 C; 3615 G; 1405 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.25e+03	Length:	9795
Score:	63.00	Matches:	17
Percent Similarity:	46.7%	Conservative:	4
Best Local Similarity:	37.8%	Mismatches:	14
Query Match:	27.5%	Indels:	10
DB:	14	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x ADY72595 (1-9795)

```
Qy          6 GluIleIleGluSerValProProAla--------------GlyProGluAlaSerGlu 20
              |||   |||         |||||||||                   |||      |||:::
Db        722 GAGGCCATCACGGTCACCCCGCCCGCCAGCGCGAGCGTGGACTCCCCGCGCCGGAGCGAC 663

Qy         21 --------------SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                           |||   |||   |||:::|||     :::      |||:::
Db        662 TCGGCAGCGAGGTGTAGCGCGACCAGCGAAGACGAACACGCGGTGTCGATGGTGACGGCC 603

Qy         36 GlyProSerArgPro 40
```

```
                 ||||||||||||||
Db        602 GGGCCTTCGAGGCCG 588
```

RESULT 71
ADY72575/c
ID ADY72575 standard; DNA; 73599 BP.
XX
AC ADY72575;
XX
DT 02-JUN-2005 (first entry)
XX
DE Polyene polyketide A. orientalis contiguous DNA, contig 2.
XX
KW antibiotic; antibacterial; cytostatic; antiinflammatory;
KW bacterial infection; cancer; inflammation; gene; ds.
XX
OS Amycolatopsis orientalis.
XX

FH	Key	Location/Qualifiers
FT	CDS	complement(32. .520)
FT		/*tag= a
FT		/partial
FT		/product= "Polyene polyketide ORF12"
FT		/note= "No start codon"
FT	CDS	840. .2165
FT		/*tag= b
FT		/partial
FT		/product= "Polyene polyketide ORF13"
FT		/note= "No start codon"
FT	CDS	2201. .3424
FT		/*tag= c
FT		/product= "Polyene polyketide ORF14"
FT	CDS	3429. .4994
FT		/*tag= d
FT		/partial
FT		/product= "Polyene polyketide ORF15"
FT		/note= "No start codon"
FT	CDS	4991. .6199
FT		/*tag= e
FT		/product= "Polyene polyketide ORF16"
FT	CDS	6389. .7924
FT		/*tag= f
FT		/partial
FT		/product= "Polyene polyketide ORF17"
FT		/note= "No start codon"
FT	CDS	8404. .19908
FT		/*tag= g
FT		/product= "Polyene polyketide ORF18"
FT	CDS	19910. .37081
FT		/*tag= h
FT		/partial
FT		/product= "Polyene polyketide ORF19"
FT		/note= "No start codon"
FT	CDS	37085. .42292
FT		/*tag= i
FT		/partial
FT		/product= "Polyene polyketide ORF20"
FT		/note= "No start codon"
FT	CDS	42617. .52411
FT		/*tag= j
FT		/product= "Polyene polyketide ORF21"
FT	CDS	52438. .67737
FT		/*tag= k
FT		/partial
FT		/product= "Polyene polyketide ORF22"
FT		/note= "No start codon"
FT	CDS	67751. .73516
FT		/*tag= l
FT		/product= "Polyene polyketide ORF23"

XX
PN WO2005024025-A1.
XX
PD 17-MAR-2005.

```
XX
PF   10-SEP-2004; 2004WO-CA001603.
XX
PR   11-SEP-2003; 2003US-0501821P.
PR   28-MAY-2004; 2004US-0574922P.
PR   23-JUN-2004; 2004US-0581707P.
XX
PA   (ECOP-) ECOPIA BIOSCIENCES INC.
XX
PI   Mcalpine JB,  Farnet CM,  Zazopoulos E,  Sorensen D;
XX
DR   WPI; 2005-233289/24.
DR   P-PSDB; ADY72576, ADY72578, ADY72580, ADY72582, ADY72584, ADY72586,
DR   ADY72588, ADY72590, ADY72592, ADY72594, ADY72596, ADY72598.
XX
PT   Novel polyene polyketide or its salt useful for treating bacterial
PT   infection, cancer or acute or chronic inflammation.
XX
PS   Example 1; SEQ ID NO 24; 374pp; English.
XX
CC   The invention relates to a novel polyene polyketide compound, its salt
CC   having a structural formula (I), prodrugs, and derivatives which are
CC   shown to have antibiotic activity. The invention further comprises: a
CC   composition comprising the compound or its salt, together with a carrier;
CC   a method of producing the compound, which involves culturing
CC   Amycolatopsis orientalis strain in a nutrient medium comprising one or
CC   more source of carbon and nitrogen atoms, and isolating the compound; an
CC   A. orientalis having the IDAC Accession No. 220604-01; an isolated
CC   purified or enriched nucleic acid encoding the polyketide synthase for
CC   producing the compound; an expression vector comprising the nucleic acid;
CC   an isolated host cell transformed with the vector; an isolated or
CC   purified polyketide synthase for producing the compound; an isolated or
CC   purified ketosynthase domain for producing the compound; an isolated or
CC   purified acyl transferase domain for producing the compound; an isolated
CC   or purified acyl carrier protein domain for producing the compound; an
CC   isolated or purified dehydratase domain for producing the compound; an
CC   isolated or purified ketoreductase domain for producing the compound; an
CC   isolated or purified enoyl reductase domain for producing the compound;
CC   an isolated or purified thioesterase domain for producing the compound; a
CC   method for producing PS, involves introducing into a host cell a nucleic
CC   acid encoding the polypeptide, where the nucleic acid is operably linked
CC   to a promoter; and a cosmid deposited under IDAC Accession No. 050903-01,
CC   050903-02, 050903-03 or 050903-04. The novel compound and its composition
CC   have antibacterial, cytostatic and antiinflammatory activities. The novel
CC   compound or its composition is useful for treating bacterial infection
CC   and for treating cancer or acute or chronic inflammation. This
CC   polynucleotide sequence represents A. orientalis contiguous DNA used in
CC   the production of the novel polyene polyketide compound of the invention.
XX
SQ   Sequence 73599 BP; 9645 A; 26041 C; 26967 G; 10946 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.44e+04        Length:          73599
Score:                    63.00           Matches:         17
Percent Similarity:       46.7%           Conservative:    4
Best Local Similarity:    37.8%           Mismatches:      14
Query Match:              27.5%           Indels:          10
DB:                       14              Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x ADY72575 (1-73599)

Qy          6 GluIleIleGluSerValProProAla--------------GlyProGluAlaSerGlu 20
              |||   |||          |||||||||              |||       |||:::
Db      43338 GAGGCCATCACGGTCACCCCGCCCGCCAGCGCGAGCGTGGACTCCCCGCGCCGGAGCGAC 43279

Qy         21 --------------SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                          |||   |||   |||:::|||     :::       |||:::
Db      43278 TCGGCAGCGAGGTGTAGCGCGACCAGCGAAGACGAACACGCGGTGTCGATGGTGACGGCC 43219

Qy         36 GlyProSerArgPro 40
              |||||||||||||
Db      43218 GGGCCTTCGAGGCCG 43204
```

```
RESULT 72
ABL82056/c
ID   ABL82056 standard; cDNA; 557 BP.
XX
AC   ABL82056;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Human ovarian cancer related cDNA clone SEQ ID NO:5034.
XX
KW   Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200192581-A2.
XX
PD   06-DEC-2001.
XX
PF   29-MAY-2001; 2001WO-US017756.
XX
PR   26-MAY-2000; 2000US-0207484P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA,  Harlocker SL,  Jones R;
XX
DR   WPI; 2002-122075/16.
XX
PT   Composition for therapy and diagnosis of ovarian cancer comprising
PT   polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT   polypeptide, antibody specific to polypeptide or T cell expressing
PT   polypeptide.
XX
PS   Claim 1; SEQ ID NO 5034; 489pp; English.
XX
CC   The present invention describes a composition (I) comprising: carriers
CC   and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC   polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC   from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC   (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC   or antigen presenting cells that express (II). (I) has cytostatic
CC   activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC   detecting ovarian cancer in a patient's biological sample preferably
CC   serum or ovarian tissue. The method comprises contacting a biological
CC   sample from a patient with (IV), detecting the amount of polynucleotide
CC   hybridising to (IV) and comparing the amount to a predetermined cutoff
CC   value and thereby detecting ovarian cancer in the patient, where the
CC   amount of polynucleotide hybridising to (IV) is detected preferably by
CC   polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC   useful for stimulating and/or expanding T cells specific for an ovarian
CC   tumour protein comprising contacting T cells with (III) or (II). (III) is
CC   useful in design and preparation of ribozyme molecules for inhibiting
CC   expression of the tumour polypeptides and proteins in tumour cells; and
CC   to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC   library using well known techniques
XX
SQ   Sequence 557 BP; 121 A; 145 C; 185 G; 105 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                53.2            Length:          557
Score:                    62.00           Matches:         17
Percent Similarity:       50.0%           Conservative:    6
Best Local Similarity:    37.0%           Mismatches:      11
Query Match:              27.1%           Indels:          12
DB:                       6               Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x ABL82056 (1-557)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn------ 26
              :::::::::|||||||||||||||||||||||||                  ||||||
Db        180 AAAGCCATTCCACCGGCCGGCCCCGAAGCCTCGCTTTTTCATGCCCAGGAGAACTCGCTG 121

Qy         27 ------GluAspAspIleGlnPheValSerGluGly---------------------- 36
```

```
                    :::|||      |||   :::      |||
Db        120 AGGCTGCAGGACCAGCACCAGGGTATAGAGCCTGGAGGCGCCCATGGTCCCTGGGTTTCT 61

Qy         37 ProSerArgProValLeu 42
              |||||||||   |||:::
Db         60 CCAAGCCGCCGTGTGATC 43

RESULT 73
ACL61290/c
ID   ACL61290 standard; cDNA; 574 BP.
XX
AC   ACL61290;
XX
DT   24-MAR-2005  (first entry)
XX
DE   Human colon cancer differentially expressed polynucleotide, SEQ ID:7425.
XX
KW   Differential expression; diagnosis; therapy; drug screening; cancer;
KW   neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO2005000087-A2.
XX
PD   06-JAN-2005.
XX
PF   13-MAY-2004; 2004WO-US015421.
XX
PR   03-JUN-2003; 2003US-0475872P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Randazzo F,  Moler E,  Escobedo J,  Garcia PD;
XX
DR   WPI; 2005-075421/08.
XX
PT   New isolated polynucleotides, which are differentially expressed in colon
PT   cancer cell, useful for treating cancer, e.g. colon cancer, breast
PT   cancer, or pancreatic cancer.
XX
PS   Claim 1; SEQ ID NO 7425; 97pp; English.
XX
CC   The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
CC   are differentially expressed in colon cancer cells. The invention also
CC   relates to vectors and host cells comprising a differentially expressed
CC   polynucleotide of the invention; a method for detecting a cancerous cell
CC   by detection of a gene product of the polynucleotides; a method for
CC   inhibiting a cancerous phenotype of a cell by inhibiting a gene product
CC   of the polynucleotides; a method of treating an individual with cancer by
CC   administration of a modulator of a gene product of the polynucleotides;
CC   and an isolated antibody that specifically binds to a polypeptide encoded
CC   by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
CC   antibodies, and methods are useful for the detection of cancerous cells;
CC   for the diagnosis, prognosis and management of cancer; for the
CC   identification of agents that modulate the phenotype of cancerous cells;
CC   for the identification of therapeutic targets for cancer chemotherapy;
CC   and for the treatment of cancer, especially colon cancer and metastasized
CC   colon cancer, but also breast or pancreatic cancer. The polynucleotides
CC   are also useful as a source of probes or primers for use in diagnostic
CC   methods. The differentially expressed polynucleotides or their encoded
CC   proteins can additionally be used as vaccines to modulate primary immune
CC   responses for the prevention or treatment of cancer. The present sequence
CC   represents a specifically claimed polynucleotide which is differentially
CC   expressed in colon cancer. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 574 BP; 121 A; 154 C; 191 G; 108 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              55.2          Length:        574
Score:                  62.00         Matches:       17
Percent Similarity:     50.0%         Conservative:  6
Best Local Similarity:  37.0%         Mismatches:    11
Query Match:            27.1%         Indels:        12
DB:                     14            Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x ACL61290 (1-574)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn------ 26
              :::::::::|||||||||||||||||||||||||                  ||||||
Db        182 AAAGCCATTCCACCGGCCGGCCCCGAAGCCTCGCTTTTTCATGCCCAGGAGAACTCGCTG 123

Qy         27 ------GluAspAspIleGlnPheValSerGluGly---------------------- 36
                  :::|||      |||   :::      |||
Db        122 AGGCTGCAGGACCAGCACCAGGGTATAGAGCCTGGAGGCGCCCATGGTCCCTGGGTTTCT 63

Qy         37 ProSerArgProValLeu 42
              |||||||||   |||:::
Db         62 CCAAGCCGCCGTGTGATC 45

RESULT 74
ADX83075
ID   ADX83075 standard; cDNA; 762 BP.
XX
AC   ADX83075;
XX
DT   19-MAY-2005  (first entry)
XX
DE   DNA encoding human (PELP1), C-terminus.
XX
KW   Proline; glutamic acid and leucine rich protein 1; PELP1; Cytostatic;
KW   RNA Interference; antibody production; pharmaceutical; tumor; neoplasm;
KW   antiestrogenic; ss; gene.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .762
FT                   /*tag=  a
FT                   /product= "Human PELP1, C-terminus"
FT                   /partial
FT                   /note= "No start codon"
XX
PN   WO2005018575-A2.
XX
PD   03-MAR-2005.
XX
PF   26-AUG-2004; 2004WO-US027834.
XX
PR   26-AUG-2003; 2003US-0498118P.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Vadlamudi RK,  Kumar R,  Balasenthil S;
XX
DR   WPI; 2005-202443/21.
DR   P-PSDB; ADX83074.
XX
PT   New Proline, Glutamic acid and Leucine rich protein 1 (PELP1)
PT   polypeptide, useful in preparing a composition for diagnosing or treating
PT   tumorigenesis in estrogen-responsive cell types.
XX
PS   Claim 4; Page 37; 50pp; English.
XX
CC   The invention relates to isolated polypeptides comprising PELP1 (Proline,
CC   glutamic acid and leucine rich protein 1) and encoding nucleic acids.
CC   Also described are: (i) a PELP1-specific antibody; (ii) a siRNA (short
CC   interfering RNA) capable of disrupting PELP1 activity; (iii) a
CC   peptidomimetic capable of disrupting PELP1 activity; (iv) a
CC   pharmaceutical composition comprising the isolated polypeptide, siRNA,
CC   peptidomimetic or PELP1-specific antibody and a carrier. The polypeptide
CC   is useful in preparing a composition for diagnosing or treating
CC   tumorigenesis in estrogen-responsive cell types. The composition
```

```
CC   comprising the isolated polypeptide, siRNA, peptidomimetic or PELP1-
CC   specific antibody and a carrier capable of all disrupting PELP1 (Proline,
CC   glutamic acid and leucine rich protein 1) activity is used to slow or
CC   halt tumorigenesis. The present sequence represents DNA encoding the C-
CC   terminus of human PELP1.
XX
SQ   Sequence 762 BP; 235 A; 167 C; 262 G; 98 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                77.7          Length:          762
Score:                    62.00         Matches:         13
Percent Similarity:       46.5%         Conservative:    7
Best Local Similarity:    30.2%         Mismatches:      23
Query Match:              27.1%         Indels:          0
DB:                       14            Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ADX83075 (1-762)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::::|||      |||   |||   |||         |||
Db        508 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 567

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 :::::     |||  :::|||        |||   |||:::
Db        568 GAGGAGCCTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 627

Qy         41 ValLeuGlu 43
                ::|||
Db        628 GAGATGGAG 636

RESULT 75
ABS65666
ID   ABS65666 standard; DNA; 1056 BP.
XX
AC   ABS65666;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Mouse JNK gene (EMBL No. q9r010) exon 22.
XX
KW   Mouse; vertebrate; anorexia/anorexia genotype; anx/anx;
KW   control of food intake; feed; body weight; human; weight gain;
KW   ANKT protein; nucleolar protein; proliferating cell; anorectic;
KW   chromosome 2; gene; ds.
XX
OS   Mus musculus.
XX
PN   WO200262835-A1.
XX
PD   15-AUG-2002.
XX
PF   08-FEB-2002; 2002WO-SE000222.
XX
PR   08-FEB-2001; 2001US-00778844.
XX
PA   (APPE-) APPETITE CONTROL AB.
XX
PI   Johansen J,  Schalling M;
XX
DR   WPI; 2002-657520/70.
XX
PT   New nucleic acid molecule encoding a polypeptide capable of at least
PT   partially complementing the phenotype of an animal having the anx/anx
PT   genotype, useful for controlling and regulating weight gain and food
PT   intake in humans.
XX
PS   Disclosure; Fig 13; 314pp; English.
XX
CC   The present invention relates to an isolated nucleic acid molecule
CC   comprising a nucleotide sequence derived from a vertebrate animal (e.g.
CC   mouse), and encodes a polypeptide capable of at least partially
CC   complementing the phenotype of an animal having the anorexia/anorexia
CC   (anx/anx) genotype. The nucleic acid molecule and the polypeptide encoded
CC   by the nucleotide sequence is useful for controlling food or feed intake,
CC   or body weight in animals including humans. The polynucleotide and
CC   polypeptide sequences of the invention are useful for regulating and
CC   controlling weight gain and food intake in animals particularly humans.
CC   The polypeptide sequences are also useful for generating monoclonal or
CC   polyclonal antibodies. The antibodies are useful for diagnostic and
CC   therapeutic purposes. These antibodies can be used in an immunoassay for
CC   identification or detection of the polypeptides. The polynucleotide
CC   sequences are useful for recombinant production of large quantities of
CC   the receptor, for further studies into the mechanism of regulation of
CC   food intake and/or body weight and for obtaining agonists and
CC   antagonists. ABS65557-ABS65762 represent mouse gene sequences relating to
CC   the present invention. The gene sequences are located on mouse chromosome
CC   2
XX
SQ   Sequence 1056 BP; 249 A; 262 C; 297 G; 248 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                115           Length:          1056
Score:                    62.00         Matches:         13
Percent Similarity:       56.2%         Conservative:    5
Best Local Similarity:    40.6%         Mismatches:      14
Query Match:              27.1%         Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x ABS65666 (1-1056)

Qy         11 ValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIle 30
              |||||||||:::|||||||      |||      :::      |||   |||||:::
Db        191 GTACCCCCTTCTGGACCAGCTCTTTCCTCAGACAGTGACAAGGAGGGAGAAGATGAGGGT 250

Qy         31 GlnPheValSerGluGlyProSerArgProValLeu 42
                        |||   |||:::   |||:::|||
Db        251 ACTGAAGAAGAAGAACTGCCAGCTCTGCCCATCCTT 286


Search completed: March 10, 2006, 15:40:14
Job time : 240.261 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:40:54 ; Search time 1450.68 Seconds
(without alignments)
1724.091 Million cell updates/sec

Title: US-10-757-745-6_COPY_2_45
Perfect score: 229
Sequence: 1 GDPGSEIIESVPPAGPEASE.....ENEDDIQFVSEGPSRPVLEY 44

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170327_4011/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_5954_@runat_09032006_170327_4011 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	229	100.0	274	6	BD025398	BD025398 Sequence
2	229	100.0	274	6	AX885788	AX885788 Sequence
3	229	100.0	1377	8	BC012011	BC012011 Homo sapi
4	229	100.0	1383	8	BC029372	BC029372 Homo sapi
5	229	100.0	1536	6	BD205492	BD205492 CD40-Inte
6	229	100.0	1536	6	AR594296	AR594296 Sequence
7	229	100.0	1536	6	AX011603	AX011603 Sequence
8	229	100.0	2878	8	BC041125	BC041125 Homo sapi
9	223	97.4	2022	6	CQ730647	CQ730647 Sequence
10	223	97.4	3042	6	AX419140	AX419140 Sequence
11	223	97.4	3042	6	AX419141	AX419141 Sequence
12	223	97.4	3183	6	AX419144	AX419144 Sequence
13	223	97.4	3183	6	AX419145	AX419145 Sequence
14	223	97.4	3247	8	BC098382	BC098382 Homo sapi
15	223	97.4	3420	8	BC021712	BC021712 Homo sapi
16	223	97.4	4961	6	CQ861506	CQ861506 Sequence
17	223	97.4	4961	8	AB011148	AB011148 Homo sapi
18	223	97.4	4999	6	AX419136	AX419136 Sequence
19	223	97.4	4999	6	AX419137	AX419137 Sequence
20	223	97.4	4999	8	AY055204	AY055204 Homo sapi
21	223	97.4	5100	8	HSM808227	BX648081 Homo sapi
22	223	97.4	5144	6	AX419138	AX419138 Sequence
23	223	97.4	5144	6	AX419139	AX419139 Sequence
24	218	95.2	1928	9	BC019546	BC019546 Mus muscu
25	218	95.2	3597	9	BC062154	BC062154 Mus muscu
26	218	95.2	5165	9	AK122541	AK122541 Mus muscu
27	142	62.0	238881	14	AC084103	AC084103 Mus muscu
c 28	142	62.0	247243	14	AC163668	AC163668 Mus muscu
29	141	61.6	101666	8	AL450489	AL450489 Human DNA
c 30	131	57.2	235278	14	AC097601	AC097601 Rattus no
31	131	57.2	267870	14	AC099359	AC099359 Rattus no
c 32	127	55.5	246801	14	AC163142	AC163142 Bos tauru
33	91	39.7	1300	5	BX932889	BX932889 Gallus ga
34	88	38.4	824	5	BX932270	BX932270 Gallus ga
35	85.5	37.3	2628	8	BC030058	BC030058 Homo sapi
36	85.5	37.3	3593	6	CQ728000	CQ728000 Sequence
37	85.5	37.3	3612	8	AF073771	AF073771 Homo sapi
38	85.5	37.3	4629	6	CQ981168	CQ981168 Sequence
39	85.5	37.3	4629	8	AF080255	AF080255 Homo sapi
40	85.5	37.3	171595	8	AL391476	AL391476 Human DNA
41	76.5	33.4	3782	5	BC072923	BC072923 Xenopus l
c 42	68.5	29.9	286181	14	AC156064	AC156064 Bos tauru
43	68	29.7	2311	8	BC002875	BC002875 Homo sapi
44	68	29.7	2715	6	CQ725289	CQ725289 Sequence
45	68	29.7	3186	8	AY882602	AY882602 Homo sapi
46	68	29.7	3216	8	BC010457	BC010457 Homo sapi
47	68	29.7	3456	8	BC069058	BC069058 Homo sapi
48	68	29.7	3910	8	HSU88153	U88153 Homo sapien
c 49	68	29.7	86945	14	AC002490	AC002490 Homo sapi
50	68	29.7	154353	8	AC027820	AC027820 Homo sapi
51	68	29.7	195494	8	AC091153	AC091153 Homo sapi
c 52	68	29.7	227406	14	AC118439	AC118439 Rattus no
53	68	29.7	235593	14	AC150165	AC150165 Gallus ga
c 54	68	29.7	235990	14	AC150044	AC150044 Gallus ga
55	68	29.7	251186	14	AC120750	AC120750 Rattus no
56	68	29.7	258285	14	AC095235	AC095235 Rattus no
57	68	29.7	301536	14	AC134158	AC134158 Rattus no
58	67.5	29.5	91152	14	BX255276_16	Continuation (17 o
59	67	29.3	2146	13	SFVDI6A	L27470 Semliki for
60	67	29.3	4093	13	SFVY17207	Y17207 Semliki For
61	67	29.3	7422	13	SFO251359	AJ251359 Semliki F
62	67	29.3	8000	6	BD138791	BD138791 Alphaviru
63	67	29.3	8100	6	BD226819	BD226819 Alphaviru
64	67	29.3	8100	6	AR243302	AR243302 Sequence
65	67	29.3	11000	6	BD138792	BD138792 Alphaviru
66	67	29.3	11406	13	AY112987	AY112987 Semliki f
67	67	29.3	11442	13	ALSFV42S	X04129 Semliki for
68	67	29.3	11517	6	A18788	A18788 complete nu
69	67	29.3	11517	6	AR001293	AR001293 Sequence
70	67	29.3	11517	6	AR130391	AR130391 Sequence
71	67	29.3	11517	6	AR569539	AR569539 Sequence
72	67	29.3	12464	6	BD263357	BD263357 Compositi
73	67	29.3	14529	13	SFVRNAIS	Z48163 Semliki for
74	67	29.3	15418	6	BD226816	BD226816 Alphaviru
75	67	29.3	15538	6	AR243299	AR243299 Sequence

ALIGNMENTS

```
RESULT 1
BD025398
LOCUS       BD025398                 274 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD025398
VERSION     BD025398.1  GI:22566621
KEYWORDS    JP 2001269182-A/1644.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 274)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 1644 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/1644
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers
            FT   CDS             108. .272.
FEATURES             Location/Qualifiers
     source          1. .274
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                3.42e-21        Length:          274
Score:                    229.00          Matches:         44
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BD025398 (1-274)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GGAGACCCGGGGTCGGARATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 170

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 230

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        231 GTTCTTGAATAC 242

RESULT 2
AX885788
LOCUS       AX885788                 274 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 1651 from Patent EP1033401.
ACCESSION   AX885788
VERSION     AX885788.1  GI:40042299
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 1651 06-SEP-2000;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .274
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             108. .>272
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE99926.1"
                     /db_xref="GI:40042300"
                     /translation="MGDPGSEIIESVPPAGPEASESTTDENEDDIQFVSEGPSRPVLE
                     YIDLVCGDDEN"
ORIGIN

Alignment Scores:
Pred. No.:                3.42e-21        Length:          274
Score:                    229.00          Matches:         44
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AX885788 (1-274)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GGAGACCCGGGGTCGGARATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 170

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 230

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        231 GTTCTTGAATAC 242

RESULT 3
BC012011
LOCUS       BC012011                 1377 bp    mRNA    linear   PRI 11-DEC-2003
DEFINITION  Homo sapiens KIAA1586, mRNA (cDNA clone IMAGE:4500277), partial
            cds.
ACCESSION   BC012011
VERSION     BC012011.1  GI:15080545
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1377)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
```

```
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1377)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 28 Row: e Column: 18.
FEATURES             Location/Qualifiers
     source          1. .1377
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4500277"
                     /tissue_type="Bladder, transitional cell papilloma"
                     /clone_lib="NIH_MGC_93"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

Alignment Scores:
Pred. No.:                1.83e-20        Length:         1377
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BC012011 (1-1377)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 105

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 165

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        166 GTTCTTGAATAC 177

RESULT 4
BC029372
LOCUS       BC029372               1383 bp    mRNA    linear   PRI 11-DEC-2003
DEFINITION  Homo sapiens KIAA1586, mRNA (cDNA clone IMAGE:4809575), partial
            cds.
ACCESSION   BC029372
VERSION     BC029372.1  GI:20809463
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1383)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1383)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 41 Row: p Column: 2
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Similarity but not
            identity to protein.
FEATURES             Location/Qualifiers
     source          1. .1383
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4809575"
                     /tissue_type="Lung, mucoepidermoid carcinoma"
                     /clone_lib="NIH_MGC_59"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     gene            <1. .1383
                     /gene="KIAA1586"
                     /db_xref="GeneID:57691"
     CDS             <1. .58
                     /gene="KIAA1586"
                     /codon_start=2
                     /product="KIAA1586 protein"
                     /protein_id="AAH29372.1"
                     /db_xref="GI:20809464"
```

/db_xref="GeneID:57691"
/translation="AVVLSARPSETWETRGRK"
ORIGIN

Alignment Scores:
Pred. No.: 1.83e-20 Length: 1383
Score: 229.00 Matches: 44
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-757-745-6_COPY_2_45 (1-44) x BC029372 (1-1383)

Qy 1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
Db 37 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 96

Qy 21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
Db 97 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 156

Qy 41 ValLeuGluTyr 44
Db 157 GTTCTTGAATAC 168

RESULT 5
BD205492
LOCUS BD205492 1536 bp DNA linear PAT 17-JUL-2003
DEFINITION CD40-Interacting and TRAF-interacting protein.
ACCESSION BD205492
VERSION BD205492.1 GI:33015262
KEYWORDS JP 2002512796-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1536)
AUTHORS Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
TITLE CD40-Interacting and TRAF-interacting protein
JOURNAL Patent: JP 2002512796-A 3 08-MAY-2002;
VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT OS Homo sapiens (human)
PN JP 2002512796-A/3
PD 08-MAY-2002
PF 28-APR-1999 JP 2000546003
PR 29-APR-1998 EP 98201392.2
PI STEFAN MARIA CHRISTIAAN PYPE,
PI JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI DANNY
FRANCOIS EVELINE HUYLEBROECK
PC C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
A61P37/02,
PC A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
PC G01N33/566,
PC C12N15/00,A61K37/02
CC CD40-Interacting and TRAF-interacting protein. FH Key
Location/Qualifiers
FT CDS (209). .(1534).
FEATURES Location/Qualifiers
source 1. .1536
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.: 2.05e-20 Length: 1536
Score: 229.00 Matches: 44
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-6_COPY_2_45 (1-44) x BD205492 (1-1536)

Qy 1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
Db 212 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 271

Qy 21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
Db 272 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 331

Qy 41 ValLeuGluTyr 44
Db 332 GTTCTTGAATAC 343

RESULT 6
AR594296
LOCUS AR594296 1536 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 5 from patent US 6812203.
ACCESSION AR594296
VERSION AR594296.1 GI:56643902
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1536)
AUTHORS Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
TITLE CD40-Interacting and TRAF-interacting proteins
JOURNAL Patent: US 6812203-A 5 02-NOV-2004;
Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
Zwijnaarde;
WOX;
FEATURES Location/Qualifiers
source 1. .1536
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.: 2.05e-20 Length: 1536
Score: 229.00 Matches: 44
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-6_COPY_2_45 (1-44) x AR594296 (1-1536)

Qy 1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
Db 212 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 271

Qy 21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
Db 272 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 331

Qy 41 ValLeuGluTyr 44
Db 332 GTTCTTGAATAC 343

RESULT 7
AX011603
LOCUS AX011603 1536 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 5 from Patent WO9955859.
ACCESSION AX011603
VERSION AX011603.1 GI:9998127
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1

```
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 5 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1536
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             209. .>1536
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07557.1"
                     /db_xref="GI:9998128"
                     /translation="MGDPGSEIIESVPPAGPEASESTTDENEDDIQFVSEGPSRPVLE
                     YIDLVCGDDENPSAYYSDILFPKMPKRQGDFLHFLNMKKVKTDTENNEVSKNHCRLSK
                     AKEPHFEYIEQPIIEEKPSLSSKKEIDNLVLPDCWNEKQAFMFTEQYKWLEIKEGKLG
                     CKDCSAVRHLGSKAEKHVHVSKEWIAYLVTPNGSNKTTRQASLRKKIREHDVSKAHGK
                     IQDLLKESTNDSICNLVHKQNNKNIDATVKVFNTVYSLVKHNRPLSDIEGARELQEKN
                     GEVNCLNTRYSATRIAEHIAKEMKMKIFKNIIEENAKICIIIDEASTVSKKTTLVIYL
                     QCTIQSAPAPVMLFVALKELVSTIAECIVNTLLTTLNDCGFTNEYLKANLIAFCSDGA
                     NTXLGRKSGVATKLLENFPEIIIWNCLNHRLQLSLDDSISEIKQINHLXY"
ORIGIN

Alignment Scores:
Pred. No.:              2.05e-20        Length:         1536
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AX011603 (1-1536)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 271

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 331

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        332 GTTCTTGAATAC 343

RESULT 8
BC041125
LOCUS       BC041125                2878 bp    mRNA    linear   PRI 08-MAR-2005
DEFINITION  Homo sapiens KIAA1586, mRNA (cDNA clone MGC:47798 IMAGE:5785266),
            complete cds.
ACCESSION   BC041125
VERSION     BC041125.1  GI:26996562
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2878)
  AUTHORS   Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
            Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg
            B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
            Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin
            GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,
            Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P,
            Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy
            SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH,
            Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW,
            Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
            Helton E, Ketteman M, Madan A, Rodrigues S, Sanchez A, Whiting M,
            Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW,
            Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J,
            Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U,
            Smailus DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED    12477932
REFERENCE   2  (bases 1 to 2878)
  AUTHORS   Director MGC Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-DEC-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 82 Row: m Column: 5
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1. .2878
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:47798 IMAGE:5785266"
                     /tissue_type="Uterus, leiomyosarcoma"
                     /clone_lib="NIH_MGC_71"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .2878
                     /gene="KIAA1586"
                     /db_xref="GeneID:57691"
     CDS             171. .2534
                     /gene="KIAA1586"
                     /codon_start=1
                     /product="KIAA1586"
                     /protein_id="AAH41125.1"
                     /db_xref="GI:26996563"
                     /db_xref="GeneID:57691"
                     /translation="MGDPGSEIIESVPPAGPEASESTTDENEDDIQFVSEGPSRPVLE
                     YIDLVCGDDENPSAYYSDILFPKMPKRQGDFLHFLNVKKVKTDTENNEVSKNHCRLSK
                     AKEPHFEYIEQPIIEEKPSLSSKKEIDNLVLPDCWNEKQAFMFTEQYKWLEIKEGKLG
                     CKDCSAVRHLGSKAEKHVHVSKEWIAYLVTPNGSNKTTRQASLRKKIREHDVSKAHGK
                     IQDLLKESTNDSICNLVHKQNNKNIDATVKVFNTVYSLVKHNRPLSDIEGARELQEKN
                     GEVNCLNTRYSATRIAEHIAKEMKMKIFKNIIEENAKICIIIDEASTVSKKTTLVIYL
                     QCTIQSAPAPVMLFVALKELVSTIAECIVNTLLTTLNDCGFTNEYLKANLIAFCSDGA
                     NTILGRKSGVATKLLENFPEIIIWNCLNHRLQLSLDDSISEIKQINHLKIFIDKIYSI
                     YHQPNKNQTKLLGTVAKELETEIIKIGRVMGPRWAACSLQAATAVWHAYPILYMHFSH
                     SYSGLAKRLANINFLQDLALMIDILEEFSVLSTALQSRSTNIKKAQKLIKRTIRALEN
                     LKIGTGKYESQIEDLIKSDKFKDIPFNKNNKFNALPRSILLDNIIQHMNLRLLSDRNH
```

EDIFNYFDLLEPSTWPYEEITSPWIAGEKTLFHLCKILKYEVDLNDFREFVNNNIKSN
NVSIPTTIYKAKKIVSTIAINSAEAERGFNLMNIICTRVRNSLTIDHVSDLMTINLLG
KELADWDATPFVKSWSNCNHRLATDTRVRQKSTKVFHENQLAIWNLK"

ORIGIN

Alignment Scores:

Pred. No.:	3.92e-20	Length:	2878
Score:	229.00	Matches:	44
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BC041125 (1-2878)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 233

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 293

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        294 GTTCTTGAATAC 305
```

RESULT 9
CQ730647

LOCUS	CQ730647	2022 bp DNA	linear PAT 03-FEB-2004
DEFINITION	Sequence 16581 from Patent WO02068579.		
ACCESSION	CQ730647		
VERSION	CQ730647.1 GI:42305057		
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.		
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof		
JOURNAL	Patent: WO 02068579-A 16581 06-SEP-2002; PE Corporation (NY) (US)		
FEATURES	Location/Qualifiers		
source	1. .2022 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"		

ORIGIN

Alignment Scores:

Pred. No.:	1.68e-19	Length:	2022
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x CQ730647 (1-2022)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1840 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 1899

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1900 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 1959

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db       1960 GTTCTTGAATAC 1971
```

RESULT 10
AX419140

LOCUS	AX419140	3042 bp DNA	linear PAT 18-JUN-2002
DEFINITION	Sequence 5 from Patent WO0216426.		
ACCESSION	AX419140		
VERSION	AX419140.1 GI:21523898		
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Dechering,K.J. and Mosselman,S.		
TITLE	Coactivation of nuclear receptors		
JOURNAL	Patent: WO 0216426-A 5 28-FEB-2002; Akzo Nobel N.V. (NL)		
FEATURES	Location/Qualifiers		
source	1. .3042 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"		

ORIGIN

Alignment Scores:

Pred. No.:	2.57e-19	Length:	3042
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x AX419140 (1-3042)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135
```

RESULT 11
AX419141

LOCUS	AX419141	3042 bp DNA	linear PAT 18-JUN-2002
DEFINITION	Sequence 6 from Patent WO0216426.		
ACCESSION	AX419141		
VERSION	AX419141.1 GI:21523899		
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Dechering,K.J. and Mosselman,S.		
TITLE	Coactivation of nuclear receptors		
JOURNAL	Patent: WO 0216426-A 6 28-FEB-2002; Akzo Nobel N.V. (NL)		
FEATURES	Location/Qualifiers		
source	1. .3042 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"		

ORIGIN

Alignment Scores:

```
Pred. No.:              2.57e-19        Length:         3042
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AX419141 (1-3042)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 12
AX419144
LOCUS       AX419144               3183 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 9 from Patent WO0216426.
ACCESSION   AX419144
VERSION     AX419144.1  GI:21523900
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dechering,K.J. and Mosselman,S.
  TITLE     Coactivation of nuclear receptors
  JOURNAL   Patent: WO 0216426-A 9 28-FEB-2002;
            Akzo Nobel N.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .3183
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.7e-19         Length:         3183
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AX419144 (1-3183)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 13
AX419145
LOCUS       AX419145               3183 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 10 from Patent WO0216426.
ACCESSION   AX419145
VERSION     AX419145.1  GI:21523901
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dechering,K.J. and Mosselman,S.
  TITLE     Coactivation of nuclear receptors
  JOURNAL   Patent: WO 0216426-A 10 28-FEB-2002;
            Akzo Nobel N.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .3183
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.7e-19         Length:         3183
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AX419145 (1-3183)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 14
BC098382
LOCUS       BC098382               3247 bp    mRNA    linear   PRI 03-AUG-2005
DEFINITION  Homo sapiens zinc finger protein 451, mRNA (cDNA clone MGC:104563
            IMAGE:5295845), complete cds.
ACCESSION   BC098382
VERSION     BC098382.1  GI:68534646
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3247)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
```

```
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3247)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) &  Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 198 Row: e Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 20127147.
FEATURES             Location/Qualifiers
     source          1. .3247
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:104563 IMAGE:5295845"
                     /tissue_type="Testis"
                     /clone_lib="NIH_MGC_97"
                     /lab_host="DH10B"
                     /note="Vector: pBluescriptR"
     gene            1. .3247
                     /gene="ZNF451"
                     /note="synonyms: COASTER, dJ417I1.1, KIAA0576"
                     /db_xref="GeneID:26036"
     CDS             68. .3193
                     /gene="ZNF451"
                     /codon_start=1
                     /product="ZNF451 protein"
                     /protein_id="AAH98382.1"
                     /db_xref="GI:68534647"
                     /db_xref="GeneID:26036"
                     /translation="MGDPGSEIIESVPPAGPEASESTTDENEDDIQFVSEGPLRPVLE
                     YIDLVSSDDEEPSTSYTDENIKRKDHIDYQKDKVALTLARLARHVEVEKQQKEEKNRA
                     FREKIDFQHAHGLQELEFIRGHSDTEAARLCVDQWLKMPGLKTGTINCGTKSSFRRGG
                     HTWVSGKPILCPIMHCNKEFDNGHLLLGHLKRFDHSPCDPTITLHGPFFSSFACVVCY
                     KKFVTQQQYRDHLFDKEATDDGHNNNLLPQIIQCFACPNCFLLFSRKEECSKHMSGKN
                     HFHQSFKLGDNKGIAHPISFPSFAKKLLISLCKDVPFQVKCVACHKTLRSHMELTAHF
                     RVHCRNAGPVAVAEKSITQVAEKFILRGYCPDCNQVFVDETSTQNHKQNSGHKVRVIN
                     SVEESVLLYCHSSEGNKDPSSDLHLLLDQSKFSSLKRTMSIKESSSLECIAIPKKKMN
                     LKDKSHEGVACVQKEKSVVKTWFCECNQRFPSEDAVEKHVFSANTMGYKCVVCGKVCD
                     DSGVIRLHMSRIHGGAHLNNFLFWCRTCKKELTRKDTIMAHVTEFHNGHRYFYEMDEV
                     EGETLPSSSTTLDNLTANKPSSAITVIDHSPANSSPRGKWQCRICEDMFDSQEYVKQH
                     CMSLASHKFHRYSCAHCRKPFHKIETLYRHCQDEHDNEIKIKYFCGLCDLIFNVEEAF
                     LSHYEEHHSIDYVFVSEKTETSIKTEDDFPVIETSNQLTCGCRESYICKVNRKEDYSR
                     CLQIMLDKGKLWFRCSLCSATAQNLTDMNTHIHQVHKEKSDEEEQQYVIKCGTCTKAF
                     HDPESAQQHFHRKHCFLQKPSVAHFGSEKSNLYKFTASASHTERRLKQAINYSKSLDM
                     EKGVENDLSYQNIEEEIVELPDLDYLRTMTHIVFVDFDNWSNFFGHLPGHLNQGTFIW
                     GFQGGNTNWKPPLNCKIYNYLNRIGCFFLHPRCSKRKDAADFAICMHAGRLDEQLPKQ
                     IPFTILSGDQGFLELENQFKKTQRPAHILNPHHLEGDMMCALLNSISDTTKGTQLQSQ
                     CKPPKRRRLHC"
ORIGIN

Alignment Scores:
Pred. No.:              2.75e-19        Length:         3247
```

```
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BC098382 (1-3247)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 130

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        131 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 190

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        191 GTTCTTGAATAC 202

RESULT 15
BC021712
LOCUS       BC021712               3420 bp    mRNA    linear   PRI 28-JUL-2005
DEFINITION  Homo sapiens zinc finger protein 451, mRNA (cDNA clone MGC:26701
            IMAGE:4821000), complete cds.
ACCESSION   BC021712
VERSION     BC021712.2  GI:34193451
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3420)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3420)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Aug 25, 2003 this sequence version replaced gi:20987289.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) &  Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
```

```
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 33 Row: a Column: 16.
FEATURES             Location/Qualifiers
     source          1. .3420
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:26701 IMAGE:4821000"
                     /tissue_type="Brain, hippocampus"
                     /clone_lib="NIH_MGC_95"
                     /lab_host="DH10B"
                     /note="Vector: pBluescriptR"
     gene            1. .3420
                     /gene="ZNF451"
                     /note="synonyms: COASTER, dJ417I1.1, KIAA0576"
                     /db_xref="GeneID:26036"
     CDS             142. .2940
                     /gene="ZNF451"
                     /codon_start=1
                     /product="ZNF451 protein"
                     /protein_id="AAH21712.2"
                     /db_xref="GI:34193452"
                     /db_xref="GeneID:26036"
                     /translation="MGDPGSEIIESVPPAGPEASESTTDENEDDIQFVSEGPLRPVLE
                     YIDLVSSDDEEPSTSYTDENIKRKDHIDYQKDKVALTLARLARHVEVEKQQKEEKNRA
                     FREKIDFQHAHGLQELEFIRGHSDTEAARLCVDQWLKMPGLKTGTINCGTKSSFRRGG
                     HTWVSGKPILCPIMHCNKEFDNGHLLLGHLKRFDHSPCDPTITLHGPFFSSFACVVCY
                     KKFVTQQQYRDHLFDKEATDDGHNNNLLPQIIQCFACPNCFLLFSRKEECSKHMSGKN
                     HFHQSFKLGDNKGIAHPISFPSFAKKLLISLCKDVPFQVKCVACHKTLRSHMELTAHF
                     RVHCRNAGPVAVAEKSITQVAEKFILRGYCPDCNQVFVDETSTQNHKQNSGHKVRVIN
                     SVEESVLLLYCHSSEGNKDPSSDLHLLLDQSKFSSLKRTMSIKESSSLECIAIPKKKMN
                     LKDKSHEGVACVQKEKSVVKTWFCECNQRFPSEDAVEKHVFSANTMGYKCVVCGKVCD
                     DSGVIRLHMSRIHGGAHLNNFLFWCRTCKKELTRKDTIMAHVTEFHNGHRYFYEMDEV
                     EGETLPSSSTTLDNLTANKPSSAITVIDHSPANSSPRGKWQCRICEDMFDSQEYVKQH
                     CMSLASHKFHRYSCAHCRKPFHKIETLYRHCQDEHDNEIKIKYFCGLCDLIFNVEEAF
                     LSHYEEHHSIDYVFVSEKTETSIKTEDDFPVIETSNQLTCGCRESYICKVNRKEDYSR
                     CLQIMLDKGKLWFRCSLCSATAQNLTDMNTHIHQVHKEKSDEEEQQYVIKCGTCTKAF
                     HDPESAQQHFHRKHCFLQKPSVAHFGSEKSNLYKFTASASHTERKLKQAINYSKSLDM
                     EKGVENDLSYQNIEEEIVELPDLDYLRTMTHIVFVDFDNWSNFFGHLPGHLNQGTFIW
                     GFQGTVNKKNKRKLFPTS"
ORIGIN

Alignment Scores:
Pred. No.:                2.91e-19        Length:         3420
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BC021712 (1-3420)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 204

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        205 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 264

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        265 GTTCTTGAATAC 276

RESULT 16
CQ861506
LOCUS       CQ861506                4961 bp    DNA     linear   PAT 10-SEP-2004
DEFINITION  Sequence 139 from Patent WO2004072265.
ACCESSION   CQ861506
VERSION     CQ861506.1  GI:51982495
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burczynski,M., Twine,N., Dorner,A.J. and Trepicchio,W.L.
  TITLE     METHODS FOR MONITORING DRUG ACTIVITIES IN VIVO /i
  JOURNAL   Patent: WO 2004072265-A 139 26-AUG-2004;
            Wyeth (US); Burczynski, Michael E. (US); Twine, Natalie C. (US);
            Dorner, Andrew J. (US); Trepicchio, William L. (US)
FEATURES             Location/Qualifiers
     source          1. .4961
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.27e-19        Length:         4961
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CQ861506 (1-4961)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 106

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        107 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 166

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        167 GTTCTTGAATAC 178

RESULT 17
AB011148
LOCUS       AB011148                4961 bp    mRNA    linear   PRI 10-APR-1998
DEFINITION  Homo sapiens mRNA for KIAA0576 protein, partial cds.
ACCESSION   AB011148
VERSION     AB011148.1  GI:3043675
KEYWORDS    KIAA0576 protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H.,
            Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes. IX.
            The complete sequences of 100 new cDNA clones from brain which can
            code for large proteins in vitro
  JOURNAL   DNA Res. 5 (1), 31-39 (1998)
   PUBMED   9628581
REFERENCE   2  (bases 1 to 4961)
  AUTHORS   Ohara,O., Nagase,T. and Ishikawa,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
            DNA Technology; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
FEATURES             Location/Qualifiers
     source          1. .4961
                     /organism="Homo sapiens"
```

```
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HJ0261"
                     /sex="male"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
     gene            1. .4961
                     /gene="KIAA0576"
     CDS             <1. .3229
                     /gene="KIAA0576"
                     /codon_start=2
                     /product="KIAA0576 protein"
                     /protein_id="BAA25502.1"
                     /db_xref="GI:3043676"
                     /translation="DSRSSGAVSAAVGDMGDPGSEIIESVPPAGPEASESTTDENEDD
                     IQFVSEGPLRPVLEYIDLVSSDDEEPSTSYTDENIKRKDHIDYQKDKVALTLARLARH
                     VEVEKQQKEEKNRAFREKIDFQHAHGLQELEFIRGHSDTEAARLCVDQWLKMPGLKTG
                     TINCGTKSSFRRGGHTWVSGKPILCPIMHCNKEFDNGHLLLGHLKRFDHSPCDPTITL
                     HGPFFSSFACVVCYKKFVTQQQYRDHLFDKEATDDGHNNNLLPQIIQCFACPNCFLLF
                     SRKEECSKHMSGKNHFHQSFKLGDNKGIAHPISFPSFAKKLLISLCKDVPFQVKCVAC
                     HKTLRSHMELTAHFRVHCRNAGPVAVAEKSITQVAEKFILRGYCPDCNQVFVDETSTQ
                     NHKQNSGHKVRVINSVEESVLLYCHSSEGNKDPSSDLHLLLDQSKFSSLKRTMSIKES
                     SSLECIAIPKKKMNLKDKSHEGVACVQKEKSVVKTWFCECNQRFPSEDAVEKHVFSAN
                     TMGYKCVVCGKVCDDSGVIRLHMSRIHGGAHLNNFLFWCRTCKKELTRKDTIMAHVTE
                     FHNGHRYFYEMDEVEGETLPSSSTTLDNLTANKPSSAITVIDHSPANSSPRGKWQCRI
                     CEDMFDSQEYVKQHCMSLASHKFHRYSCAHCRKPFHKIETLYRHCQDEHDNEIKIKYF
                     CGLCDLIFNVEEAFLSHYEEHHSIDYVFVSEKTETSIKTEDDFPVIETSNQLTCGCRE
                     SYICKVNRKEDYSRCLQIMLDKGKLWFRCSLCSATAQNLTDMNTHIHQVHKEKSDEEE
                     QQYVIKCGTCTKAFHDPESAQQHFHRKHCFLQKPSVAHFGSEKSNLYKFTASASHTER
                     KLKQAINYSKSLDMEKGVENDLSYQNIEEEIVELPDLDYLRTMTHIVFVDFDNWSNFF
                     GHLPGHLNQGTFIWGFQGGNTNWKPPLNCKIYNYLNRIGCFFLHPRCSKRKDAADFAI
                     CMHAGRLDEQLPKQIPFTILSGDQGFLELENQFKKTQRPAHILNPHHLEGDMMCALLN
                     SISDTTKECDSDDNMGAKNTSIGEEFISTEDVELEEAIRRSLEEM"
ORIGIN

Alignment Scores:
Pred. No.:                4.27e-19        Length:         4961
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AB011148 (1-4961)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 106

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        107 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 166

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        167 GTTCTTGAATAC 178

RESULT 18
AX419136
LOCUS       AX419136                4999 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 1 from Patent WO0216426.
ACCESSION   AX419136
VERSION     AX419136.1  GI:21523894
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dechering,K.J. and Mosselman,S.
  TITLE     Coactivation of nuclear receptors
  JOURNAL   Patent: WO 0216426-A 1 28-FEB-2002;
            Akzo Nobel N.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .4999
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.31e-19        Length:         4999
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AX419136 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 19
AX419137
LOCUS       AX419137                4999 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 2 from Patent WO0216426.
ACCESSION   AX419137
VERSION     AX419137.1  GI:21523895
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dechering,K.J. and Mosselman,S.
  TITLE     Coactivation of nuclear receptors
  JOURNAL   Patent: WO 0216426-A 2 28-FEB-2002;
            Akzo Nobel N.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .4999
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.31e-19        Length:         4999
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AX419137 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
```

```
                    ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 20
AY055204
LOCUS       AY055204               4999 bp    mRNA    linear   PRI 01-APR-2002
DEFINITION  Homo sapiens COASTER mRNA, complete cds.
ACCESSION   AY055204
VERSION     AY055204.1  GI:19879435
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 4999)
  AUTHORS   Heldens,I.M. and Dechering,K.J.
  TITLE     Isolation of a novel coactivator for steroid receptors that alters
            the intrinsic activity of the estrogen receptor alpha liganded with
            SERMs
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4999)
  AUTHORS   Heldens,I.M. and Dechering,K.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-SEP-2001) Target Discovery, N.V. Organon, Molenstraat
            110, Oss 5340 BH, The Netherlands
FEATURES             Location/Qualifiers
     source          1. .4999
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             226. .3267
                     /note="coactivator for steroid receptors"
                     /codon_start=1
                     /product="COASTER"
                     /protein_id="AAL17975.1"
                     /db_xref="GI:19879436"
                     /translation="MGDPGSEIIESVPPAGPEASESTTDENEDDIQFVSEGPLRPVLE
                     YIDLVSSDDEEPSTSYTDENIKRKDHIDYQKDKVALTLARLARHVEVEKQQKEEKNRA
                     FREKIDFQHAHGLQELEFIRGHSDTEAARLCVDQWLKMPGLKTGTINCGTKSSFRRGG
                     HTWVSGKPTLCPIMHCNKEFDNGHLLLGHLKRFDHSPCDPTITLHGPFFSSFACVVCY
                     KKFVTQQQYRDHLFDKEATDDGHNNNLLPQIIQCFACPNCFLLFSRKEECSKHMSGKN
                     HFHQSFKLGDNKGIAHPISFPSFAKKLLISLCKDVPFQVKCVACHKTLRSHMELTAHF
                     RVHCRNAGPVAVAEKSITQVAEKFILRGYCPDCNQVFVDETSTQNHKQNSGHKVRVIN
                     SVEESVLLYCHSSEGNKDPSSDLHLLLDQSKFSSLKRTMSIKESSSLECIAIPKKKMN
                     LKDKSHEGVACVQKEKSVVKTWFCECNQRFPSEDAVEKHVFSANTMGYKCVVCGKVCD
                     DSGVIRLHMSRIHGGAHLNNFLFWCRTCKKELTRKDTIMAHVTEFHNGHRYFYEMDEV
                     EGETLPSSSTTLDNLTANKPSSAITVIDHSPANSSPRGKWQCRICEDMFDSQEYVKQH
                     CMSLASHKFHRYSCAHCRKPFHKIETLYRHCQDEHDNEIKIKYFCGLCDLIFNVEEAF
                     LSHYEEHHSIDYVFVSEKTETSIKTEDDFPVIETSNQLTCGCRESYICKVNRKEDYSR
                     CLQIMLDKGKLWFRCSLCSATAQNLTDMNTHIHQVHKEKSDEEEQQYVIKCGTCTKAF
                     HDPESAQQHFHRKHCFLQKPSVAHFGSEKSNLYKFTASASHTERKLKQAINYSKSLDM
                     EKGVENDLSYQNIGGNTNWKPPLNCKIYNYLNRIGCFFLHPRCSKRKDAADFAICMHA
                     GRLDEQLPKQIPFTILSGDQGFLELENQFKKTQRPAHILNPHHLEGDMMCALLNSISD
                     TTKECDSDDNMGAKNTSIGEEFISTEDVELEEAIRRSLEEM"
ORIGIN

Alignment Scores:
Pred. No.:              4.31e-19        Length:         4999
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AY055204 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
```

```
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 21
HSM808227
LOCUS       HSM808227              5100 bp    mRNA    linear   PRI 30-AUG-2003
DEFINITION  Homo sapiens mRNA; cDNA DKFZp686M03226 (from clone DKFZp686M03226).
ACCESSION   BX648081
VERSION     BX648081.1  GI:34367240
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 5100)
  AUTHORS   Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  CONSRTM   The German Human cDNA Consortium
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            This clone (DKFZp686M03226) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .5100
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686M03226"
                     /tissue_type="human small intestine"
                     /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
                     DH10B; sites SfiIA + SfiIB"
                     /dev_stage="adult"
     polyA_site      5087
ORIGIN

Alignment Scores:
Pred. No.:              4.4e-19         Length:         5100
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x HSM808227 (1-5100)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 133

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        134 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 193

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        194 GTTCTTGAATAC 205

RESULT 22
AX419138
```

```
LOCUS       AX419138                5144 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 3 from Patent WO0216426.
ACCESSION   AX419138
VERSION     AX419138.1  GI:21523896
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dechering,K.J. and Mosselman,S.
  TITLE     Coactivation of nuclear receptors
  JOURNAL   Patent: WO 0216426-A 3 28-FEB-2002;
            Akzo Nobel N.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .5144
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.44e-19        Length:          5144
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AX419138 (1-5144)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 23
AX419139
LOCUS       AX419139                5144 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 4 from Patent WO0216426.
ACCESSION   AX419139
VERSION     AX419139.1  GI:21523897
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dechering,K.J. and Mosselman,S.
  TITLE     Coactivation of nuclear receptors
  JOURNAL   Patent: WO 0216426-A 4 28-FEB-2002;
            Akzo Nobel N.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .5144
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.44e-19        Length:          5144
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AX419139 (1-5144)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 24
BC019546
LOCUS       BC019546                1928 bp    mRNA    linear   ROD 13-FEB-2004
DEFINITION  Mus musculus zinc finger protein 451, mRNA (cDNA clone MGC:28438
            IMAGE:4038791), complete cds.
ACCESSION   BC019546
VERSION     BC019546.1  GI:18044542
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1928)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1928)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-DEC-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
```

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 37 Row: b Column: 8
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

```
FEATURES             Location/Qualifiers
     source          1. .1928
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="mix FVB/N, C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="MGC:28438 IMAGE:4038791"
                     /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
                     old, gross tissue."
                     /clone_lib="NCI_CGAP_Mam5"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1928
                     /gene="Zfp451"
                     /note="synonym: COASTER"
                     /db_xref="GeneID:98403"
                     /db_xref="MGI:2137896"
     CDS             74. .1753
                     /gene="Zfp451"
                     /codon_start=1
                     /product="Zfp451 protein"
                     /protein_id="AAH19546.1"
                     /db_xref="GI:18044543"
                     /db_xref="GeneID:98403"
                     /db_xref="MGI:2137896"
                     /translation="MGDPGPEIIESVPPAGPEASESTTDENEDDIQFVSEGPLRPVLE
                     YIDLVSSDDEEPSTSHSDRMPESKVPSSENHRPEMCSSCSVPLPIGDSSSSSGSCANS
                     PQRIVSQPSSVENPLENQKNDHNNSDVSVSEKETLKPQSNQTLPSASLLGSQESLAFS
                     EVKEDLPGTMSSASQSGQDAILYLQTQVAEMSRVIRDLQSRSCFRFHHSRPSESSSVP
                     WDISSSKEDSLSTVEEESDCKSPSADDKGQPPDPMQSSFTGLLKRMEQRGVIKRVTFQ
                     SEAETCEGKPDWMTSKKRLVPPLHPLLRIATTEVFKDPADCHPSSFMGHRVYPVAKDT
                     SPFQPNPPAEGPIIDVLEHSKRGNTASPLDSTSKEMEVMGCRFYHAASIAARAASYMA
                     YMTQYQRKLWEDMEDLVHDPEFDRGKARCIISDGMDAGLWQLCTTRDIMDSVVRVMAM
                     AVDYRRQAWLRLTSLTKKTQEKILHLPFDGTSLFGQDVNAIVAEDNSIKENEYKDHNK
                     YYTQHRYFYSHDQKAHYHNRGYSKGDWYKPRNHPYRYRKKGDSPERHGYKN"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                7.33e-19        Length:          1928
Score:                    218.00          Matches:         42
Percent Similarity:       95.5%           Conservative:    0
Best Local Similarity:    95.5%           Mismatches:      2
Query Match:              95.2%           Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BC019546 (1-1928)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         77 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 136

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 196

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        197 GTTTTGGAATAC 208

RESULT 25
BC062154
LOCUS       BC062154                3597 bp    mRNA    linear   ROD 12-FEB-2004
DEFINITION  Mus musculus zinc finger protein 451, mRNA (cDNA clone MGC:70050
            IMAGE:6331597), complete cds.
ACCESSION   BC062154
VERSION     BC062154.1  GI:38328433
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3597)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3597)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
            Ph.D.
            cDNA Library Preparation: ResGen, Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 131 Row: l Column: 7
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1. .3597
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="MGC:70050 IMAGE:6331597"
                     /tissue_type="Embryo, day 9 mouse (C57BL/6 background)
                     otocysts"
                     /clone_lib="NIH_MGC_130"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6.1"
```

```
     gene            1. .3597
                     /gene="Zfp451"
                     /note="synonym: COASTER"
                     /db_xref="GeneID:98403"
                     /db_xref="MGI:2137896"
     CDS             273. .2960
                     /gene="Zfp451"
                     /codon_start=1
                     /product="Zfp451 protein"
                     /protein_id="AAH62154.1"
                     /db_xref="GI:38328434"
                     /db_xref="GeneID:98403"
                     /db_xref="MGI:2137896"
                     /translation="MWRSEKPILCPIMHCNKEFDNGHLLLGHLKRFDHSPCDPTITLH
                     GPLANSFACAVCYEHFVTQQQYKDHLLSRTAAADGHSNSLLPQIIQCYACPQCFLLFS
                     TKDECLKHMSTKNHFHQSFKLSDNKGTARPISFPSFAKKRLVSLCKDVPFQVKSVACH
                     QTLRSHMELTAHFRVRCQNAGPVAIAEKSITQVAKEFIVRGYCSDCNQVFMDVASTQS
                     HKNSGHKITLANSVEESVLLYCHISEGSRPPCDLHLFSQPKISSLKRILSVKESSAED
                     CIVPTKKVNLGVESLGGATRVQRQSPAVTAWFCECRRQFPSEEAVEKHVFSANTMCYK
                     CVVCGKVCEDSGVMRLHMSRFHGGAHLNNFLFWCRTCKKELVKKDAIMAHITEFHSGH
                     RYFYEMDEVEEEEEEAMPSSSVESHLNTDKPPSPIAVVDHCPANSPPRGRWQCRICED
                     MFESQECVKQHCMSLTSHRFHRYSCAHCRKTFHKVETLYRHCQDEHDSEIMMKYFCGL
                     CDLIFNKEEEFLSHYKEHHSIDYVPVSEKTKTSIKTEGDFKIVETSSLLSCGCHESYM
                     CKINRKEDYDRCLPVLLEKGRLWFRCSSCSATAQNVTDINTHVCQVHRKEKSEEEQQY
                     VIKCGICTKAFQNTESAQQHFHRKHAALQKPTATPGGANRSSTCQLAASASHAEKNLK
                     QPSSQKHSDVEKGAEHDVRCQNIEEEVELPDVDYLRTMTHIVFVDFDNWSNFFGHLPG
                     HLNQGTFIWGFQGGNTNWKPPLSCKVYNYLSRIGCFFLHPRCSKRKDAADFAICMHAG
                     RLDEQLPKQIPFTILSGDQGFLELENQFKKTQRPAHILNPHHLEGDMMCALLNSISDT
                     TKECDSDDSSGMKGSPAEELRATEDVELEEAIRRSLEEM"
ORIGIN

Alignment Scores:
Pred. No.:                1.4e-18         Length:          3597
Score:                    218.00          Matches:         42
Percent Similarity:       95.5%           Conservative:    0
Best Local Similarity:    95.5%           Mismatches:      2
Query Match:              95.2%           Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BC062154 (1-3597)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         31 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 90

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         91 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 150

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        151 GTTTTGGAATAC 162

RESULT 26
AK122541
LOCUS       AK122541                5165 bp    mRNA    linear   ROD 15-MAR-2003
DEFINITION  Mus musculus mRNA for mKIAA1702 protein.
ACCESSION   AK122541
VERSION     AK122541.1  GI:28972813
KEYWORDS    FLI_CDNA.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S.,
            Nakajima,D., Nagase,T., Ohara,O. and Koga,H.
  TITLE     Prediction of the coding sequences of mouse homologues of KIAA
            gene: II. The complete nucleotide sequences of 400 mouse
            KIAA-homologous cDNAs identified by screening of terminal sequences
            of cDNA clones randomly sampled from size-fractionated libraries
  JOURNAL   DNA Res. 10, 35-48 (2003)
REFERENCE   2  (bases 1 to 5165)
  AUTHORS   Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research
            Institute, Laboratory for Genome Informatics; 2-6-7
            Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT     The CREATE program supported by Japan science and technology
            corporation; cDNA full insert sequencing: Kazusa DNA Research
            Institute; cDNA library construction, clone selection and 5'- &
            3'-end one pass sequencing.
FEATURES             Location/Qualifiers
     source          1. .5165
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="mbg07569"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /note="vector:modified pBC SK+"
     gene            1. .5165
                     /gene="mKIAA1702"
     CDS             <48. .1742
                     /gene="mKIAA1702"
                     /note="CDS is predicted by in silico analysis. Start codon
                     is not identified."
                     /codon_start=1
                     /evidence=not_experimental
                     /product="mKIAA1702 protein"
                     /protein_id="BAC65823.1"
                     /db_xref="GI:28972814"
                     /translation="QAVGEMGDPGPEIIESVPPAGPEASESTTDENEDDIQFVSEGPL
                     RPVLEYIDLVSSDDEEPSTSHSDRMPESKVPSSENHRPEMCSSCSVPLPIGDSSSSSG
                     SCANSPQRIVSQPSSVENPLENQKNDHNNSDVSVSEKETLKPQSNQTLPSASLLGSQE
                     SLAFSEVKEDLPGTMSSASQSGQDAILYLQTQVAEMSRVIRDLQSRSCFRFHHSRPSE
                     SSSVPWDISSSKEDSLSTVEEESDCKSPSADDKGQPPDPMQSSFTGLLKRMEQRGVIK
                     RVTFQSEAETCEGKPDWMTSKKRLVPPLHPLLRIATTEVFKDPADCHPSSFMGHRVYP
                     VAKDTSPFQPNPPAEGPIIDVLEHSKRGNTASPLDSTSKEMEVMGCRFYHAASIAARA
                     ASYMAYMTQYQRKLWEDMEDLVHDPEFDRGKARCIISDGMDAGLWQLCTTRDIMDSVV
                     RVMAMAVDYRRQAWLRLTSLTKKTQEKILHLPFDGTSLFGQDVNAIVAEDNSIKENEY
                     KDHNKYYTQHRYFYSHDQKAHYHNRGYSKGDWYKPRNHPYRYRKKGDSPERHGYKN"
ORIGIN

Alignment Scores:
Pred. No.:                2.04e-18        Length:          5165
Score:                    218.00          Matches:         42
Percent Similarity:       95.5%           Conservative:    0
Best Local Similarity:    95.5%           Mismatches:      2
Query Match:              95.2%           Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AK122541 (1-5165)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         66 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 125

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        126 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 185

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        186 GTTTTGGAATAC 197

RESULT 27
AC084103
LOCUS       AC084103                238881 bp    DNA     linear   HTG 13-OCT-2000
DEFINITION  Mus musculus clone RP23-194M15, WORKING DRAFT SEQUENCE, 16
            unordered pieces.
ACCESSION   AC084103
VERSION     AC084103.1  GI:10800221
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
```

```
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 238881)
  AUTHORS   McCombie,W.R., Baker,J.P., Bahret,A., Bal,H., Balija,V.,
            Dedhia,N.N., de la Bastide,M., Huang,E.N., King,L., Kirchoff,K.A.,
            Miller,B., Nascimento,L.U., O'Shaughnessy,A.L., Preston,R.R.,
            Rodriguez,S., Santos,L., Shah,R.S., Spiegel,L.A., Toth,K., Vil,M.D.
            and Zutavern,T.
  TITLE     Mouse Genomic Sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 238881)
  AUTHORS   McCombie,W.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-2000) Lita Annenberg Hazen Genome Sequencing
            Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
            Harbor, NY 11724, USA
COMMENT     ------------------- Genome Center
            Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
            Laboratory
            Center code: CSHL
            Web site: http://www.cshl.org/genseq
            Contact:  mccombie@cshl.org
            ------------------- Project Information
            Center project name: RP23-194M15
            Center clone name: RP23-194M15
            ------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    49777: contig of 49777 bp in length
            *    49778    50145: gap of unknown length
            *    50146    89723: contig of 39578 bp in length
            *    89724    90091: gap of unknown length
            *    90092   114797: contig of 24706 bp in length
            *   114798   115165: gap of unknown length
            *   115166   137619: contig of 22454 bp in length
            *   137620   137987: gap of unknown length
            *   137988   155279: contig of 17292 bp in length
            *   155280   155646: gap of unknown length
            *   155647   170917: contig of 15271 bp in length
            *   170918   171284: gap of unknown length
            *   171285   185125: contig of 13841 bp in length
            *   185126   185492: gap of unknown length
            *   185493   197574: contig of 12082 bp in length
            *   197575   197941: gap of unknown length
            *   197942   208656: contig of 10715 bp in length
            *   208657   209023: gap of unknown length
            *   209024   217380: contig of 8357 bp in length
            *   217381   217747: gap of unknown length
            *   217748   221877: contig of 4130 bp in length
            *   221878   222244: gap of unknown length
            *   222245   226091: contig of 3847 bp in length
            *   226092   226458: gap of unknown length
            *   226459   229820: contig of 3362 bp in length
            *   229821   230187: gap of unknown length
            *   230188   233215: contig of 3028 bp in length
            *   233216   233582: gap of unknown length
            *   233583   236341: contig of 2759 bp in length
            *   236342   236708: gap of unknown length
            *   236709   238881: contig of 2173 bp in length.
FEATURES             Location/Qualifiers
     source          1. .238881
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-194M15"
                     /clone_lib="RPCI-23"
     gap             49778. .50145
                     /estimated_length=unknown
     gap             89724. .90091
                     /estimated_length=unknown
     gap             114798. .115165
                     /estimated_length=unknown
     gap             137620. .137987
                     /estimated_length=unknown
     gap             155280. .155646
                     /estimated_length=unknown
     gap             170918. .171284
                     /estimated_length=unknown
     gap             185126. .185492
                     /estimated_length=unknown
     gap             197575. .197941
                     /estimated_length=unknown
     gap             208657. .209023
                     /estimated_length=unknown
     gap             217381. .217747
                     /estimated_length=unknown
     gap             221878. .222244
                     /estimated_length=unknown
     gap             226092. .226458
                     /estimated_length=unknown
     gap             229821. .230187
                     /estimated_length=unknown
     gap             233216. .233582
                     /estimated_length=unknown
     gap             236342. .236708
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:              1.17e-06      Length:          238881
Score:                  142.00        Matches:         29
Percent Similarity:     87.9%         Conservative:    0
Best Local Similarity:  87.9%         Mismatches:      4
Query Match:            62.0%         Indels:          0
DB:                     14            Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AC084103 (1-238881)

Qy          7 IleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn 26
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126410 ATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAGTCTACCACGGATGAAAAC 126469

Qy         27 GluAspAspIleGlnPheValSerGluGlyProSerArg 39
              |||||||||||||||||||||||  |||
Db     126470 GAAGATGACATTCAGTTTGTTAGTGTAAGTAGAGACAGA 126508

RESULT 28
AC163668/c
LOCUS       AC163668              247243 bp    DNA     linear   HTG 29-JUL-2005
DEFINITION  Mus musculus chromosome 1 clone RP23-32P22, WORKING DRAFT SEQUENCE,
            5 unordered pieces.
ACCESSION   AC163668
VERSION     AC163668.3  GI:71480400
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 247243)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 247243)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUN-2005) Genome Sequencing Center, 4444 Forest Park
```

```
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 247243)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUL-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On Jul 29, 2005 this sequence version replaced gi:71143410.

            ----------------- Genome Center ----------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information -------------------
            Center project name: M_BA0032P22
            ------------------------------------------------------
            ------------- Summary Statistics  --------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 244839 bases at least Q40
            Consensus quality: 245354 bases at least Q30
            Consensus quality: 245649 bases at least Q20
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1519: contig of 1519 bp in length
            *     1520     1619: gap of unknown length
            *     1620    12491: contig of 10872 bp in length
            *    12492    12591: gap of unknown length
            *    12592    50864: contig of 38273 bp in length
            *    50865    50964: gap of unknown length
            *    50965   100974: contig of 50010 bp in length
            *   100975   101074: gap of unknown length
            *   101075   247243: contig of 146169 bp in length.
FEATURES             Location/Qualifiers
     source          1. .247243
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /clone="RP23-32P22"
     misc_feature    1. .1519
                     /note="assembly_name:Contig23"
     gap             1520. .1619
                     /estimated_length=unknown
     misc_feature    1620. .12491
                     /note="assembly_name:Contig41"
     gap             12492. .12591
                     /estimated_length=unknown
     misc_feature    12592. .50864
                     /note="assembly_name:Contig42"
     gap             50865. .50964
                     /estimated_length=unknown
     misc_feature    50965. .100974
                     /note="assembly_name:Contig43"
     gap             100975. .101074
                     /estimated_length=unknown
     misc_feature    101075. .247243
                     /note="assembly_name:Contig44"
ORIGIN

Alignment Scores:
Pred. No.:          1.21e-06        Length:         247243
Score:              142.00          Matches:        29
Percent Similarity: 87.9%           Conservative:   0
Best Local Similarity: 87.9%        Mismatches:     4
Query Match:        62.0%           Indels:         0
DB:                 14              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AC163668 (1-247243)

Qy          7 IleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn 26
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     177037 ATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAGTCTACCACGGATGAAAAC 176978

Qy         27 GluAspAspIleGlnPheValSerGluGlyProSerArg 39
              |||||||||||||||||||||||          |||
Db     176977 GAAGATGACATTCAGTTTGTTAGTGTAAGTAGAGACAGA 176939

RESULT 29
AL450489
LOCUS       AL450489              101666 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-203B9 on chromosome 6 Contains
            the 3 'end of gene FLJ30162, a novel gene, the 5'end of gene
            KIAA0576, the 3' end of a novel gene and a CpG island, complete
            sequence.
ACCESSION   AL450489
VERSION     AL450489.12  GI:15022239
KEYWORDS    HTG; FLJ30162; KIAA0576.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 101666)
  AUTHORS   Leongamornlert,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 26, 2001 this sequence version replaced gi:14715364.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-203B9 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .101666
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-203B9"
                     /clone_lib="RPCI-11.1"
     misc_feature    1
                     /note="Clone_left_end: RP11-203B9"
```

```
gene            join(AL590037.7:24757. .24925,AL590037.7:51176. .51395,
                AL590037.7:61843. .62020,AL590037.7:84598. .84818,
                AL590037.7:86672. .86688,AL590037.7:86792. .86864,
                AL590037.7:87886. .88022,1494. .2754)
                /gene="C6orf65"
                /locus_tag="RP11-203B9.5-001"
gene            join(AL590037.7:24757. .24925,AL590037.7:51176. .51395,
                AL590037.7:61843. .62020,AL590037.7:84598. .84818,
                AL590037.7:86672. .86864,AL590037.7:87886. .88022,
                1494. .2754)
                /gene="C6orf65"
                /locus_tag="RP11-203B9.5-004"
mRNA            join(AL590037.7:24757. .24925,AL590037.7:51176. .51395,
                AL590037.7:61843. .62020,AL590037.7:84598. .84818,
                AL590037.7:86672. .86688,AL590037.7:86792. .86864,
                AL590037.7:87886. .88022,1494. .2754)
                /gene="C6orf65"
                /locus_tag="RP11-203B9.5-001"
                /product="chromosome 6 open reading frame 65"
                /note="match: ESTs: AW963721 BE145493 BE647602 BI913530
                BI914970 BI916461 N48613 R59280
                match: cDNAs: AK054724"
mRNA            join(AL590037.7:24757. .24925,AL590037.7:51176. .51395,
                AL590037.7:61843. .62020,AL590037.7:84598. .84818,
                AL590037.7:86672. .86864,AL590037.7:87886. .88022,
                1494. .2754)
                /gene="C6orf65"
                /locus_tag="RP11-203B9.5-001"
                /product="chromosome 6 open reading frame 65"
                /note="match: ESTs: BE647602.1 BG819883.1 BI913530.1
                BI914970.1 BI916461.1 H07984.1 R59280.1 Z44820.1"
gene            join(21961. .22188,22725. .22808,26186. .26266,28098. .30637)
                /locus_tag="RP11-203B9.2-001"
mRNA            join(21961. .22188,22725. .22808,26186. .26266,28098. .30637)
                /locus_tag="RP11-203B9.2-001"
                /note="match: ESTs: AA001244.1 AA779733.1 AA938639.1
                AA969559.1 AI474013.1 AW264000.1 AW977782.1 BE614550.1
                BF030066.1 BF218739.1 BF339845.1 BF687541.1 BG028983.1
                BG286553.1 BG777259.1 BG942992.1 BG944466.1 BI438747.1
                BI465045.1 BI560244.1 BM352924.1 BM835195.1 BQ230874.1
                match: cDNAs: BC012011.1 BC029372.1"
gene            join(22159. .22333,22725. .22808,26186. .26266,28098. .28458)
                /locus_tag="RP11-203B9.2-002"
mRNA            join(22159. .22333,22725. .22808,26186. .26266,28098. .28458)
                /locus_tag="RP11-203B9.2-002"
                /note="match: ESTs: BF699031.1"
CDS             join(22168. .22188,22725. .22808,26186. .26266,28098. .30275)
                /locus_tag="RP11-203B9.2-001"
                /standard_name="OTTHUMP00000016662"
                /note="match: proteins: Q8VCL4 Q9CUK0 Q9Y4E5"
                /codon_start=1
                /protein_id="CAH71219.1"
                /db_xref="GI:55665231"
                /db_xref="Genew:21360"
                /db_xref="UniProt/TrEMBL:Q8IW25"
                /translation="MGDPGSEIIESVPPAGPEASESTTDENEDDIQFVSEGPSRPVLE
                YIDLVCGDDENPSAYYSDILFPKMPKRQGDFLHFLNVKKVKTDTENNEVSKNHCRLSK
                AKEPHFEYIEQPIIEEKPSLSSKKEIDNLVLPDCWNEKQAFMFTEQYKWLEIKEGKLG
                CKDCSAVRHLGSKAEKHVHVSKEWIAYLVTPNGSNKTTRQASLRKKIREHDVSKAHGK
                IQDLLKESTNDSICNLVHKQNNKNIDATVKVFNTVYSLVKHNRPLSDIEGARELQEKN
                GEVNCLNTRYSATRIAEHIAKEMKMKIFKNIIEENAKICIIIDEASTVSKKTTLVIYL
                QCTIQSAPAPVMLFVALKELVSTIAECIVNTLLTTLNDCGFTNEYLKANLIAFCSDGA
                NTILGRKSGVATKLLENFPEIIIWNCLNHRLQLSLDDSISEIKQINHLKIFIDKIYSI
                YHQPNKNQTKLLGTVAKELETEIIKIGRVMGPRWAACSLQAATAVWHAYPILYMHFSH
                SYSGLAKRLANINFLQDLALMIDILEEFSVLSTALQSRSTNIKKAQKLIKRTIRALEN
                LKIGTGKYESQIEDLIKSDKFKDIPFNKNNKFNALPRSILLDNIIQHMNLRLLSDRNH
                EDIFNYFDLLEPSTWPYEEITSPWIAGEKTLFHLCKILKYEVDLNDFREFVNNNIKSN
                NVSIPTTIYKAKKIVSTIAINSAEAERGFNLMNIICTRVRNSLTIDHVSDLMTINLLG
                KELADWDATPFVKSWSNCNHRLATDTRVRQKSTKVFHENQLAIWNLK"
polyA_signal    29210. .29215
                /locus_tag="RP11-203B9.2-001"
polyA_site      29237
                /locus_tag="RP11-203B9.2-001"
polyA_signal    30604. .30609
                /locus_tag="RP11-203B9.2-001"
polyA_signal    30608. .30613
                /locus_tag="RP11-203B9.2-001"
polyA_signal    30624. .30629
                /locus_tag="RP11-203B9.2-001"
polyA_site      30637
                /locus_tag="RP11-203B9.2-001"
gene            join(65445. .65686,66223. .66306,74473. .74553,76015. .85128)
                /gene="ZNF451"
                /locus_tag="RP3-417I1.1-002"
mRNA            join(65445. .65686,66223. .66306,74473. .74553,76015. .85128)
                /gene="ZNF451"
                /locus_tag="RP3-417I1.1-002"
                /product="zinc finger protein 451"
                /note="match: ESTs: AA455600.1 AA486412.1 AL703808.1
                AW129544.1 AW179230.1 AW179233.1 BG259331.1 BG388119.1
                BG772992.1 BM459825.1 BM749349.1 BM980826.1 BQ362289.1
                H06401.1
                match: cDNAs: AK015797.1 BC019546.1"
gene            join(65450. .65686,66223. .66306,74473. .74553,
                100146. .100271,AL136311.7:4475. .4586,
                AL136311.7:8788. .8938,AL136311.7:10490. .10616,
                AL136311.7:17048. .17201,AL136311.7:17694. .17841,
                AL136311.7:22836. .24439,AL136311.7:26465. .26608,
                AL136311.7:27967. .28097,AL136311.7:29607. .29793,
                AL136311.7:36830. .36898,AL136311.7:44169. .46053)
                /gene="ZNF451"
                /locus_tag="RP3-417I1.1-001"
mRNA            join(65450. .65686,66223. .66306,74473. .74553,
                100146. .100271,AL136311.7:4475. .4586,
                AL136311.7:8788. .8938,AL136311.7:10490. .10616,
                AL136311.7:17048. .17201,AL136311.7:17694. .17841,
                AL136311.7:22836. .24439,AL136311.7:26465. .26608,
                AL136311.7:27967. .28097,AL136311.7:29607. .29793,
                AL136311.7:36830. .36898,AL136311.7:44169. .46053)
                /gene="ZNF451"
                /locus_tag="RP3-417I1.1-001"
                /product="zinc finger protein 451"
                /note="match: ESTs: AA127580.1 AA541929.1 AA811518.1
                AL044083 AW502925.1 BB659280.1 BE613457.1 BG701977.1
                BG705499.1 BG722501.1 BI559958.1 BI836404.1 L19145.1
                W03189.1
                match: cDNAs: AB011148.1 AY055204.1 BC021712.1"
gene            join(65489. .65686,66223. .66306,74473. .74553,99841. .99935)
                /gene="ZNF451"
                /locus_tag="RP3-417I1.1-006"
mRNA            join(65489. .65686,66223. .66306,74473. .74553,99841. .99935)
                /gene="ZNF451"
                /locus_tag="RP3-417I1.1-006"
                /product="zinc finger protein 451"
                /note="match: ESTs: AI017172.1 AW118103.1"
gene            join(65632. .65686,66223. .66306,74473. .74553,74881. .75022)
                /gene="ZNF451"
                /locus_tag="RP3-417I1.1-003"
mRNA            join(65632. .65686,66223. .66306,74473. .74553,74881. .75022)
                /gene="ZNF451"
                /locus_tag="RP3-417I1.1-003"
                /product="zinc finger protein 451"
                /note="match: ESTs: AW865717"
gene            join(65657. .65831,66223. .66306,74473. .74553)
                /gene="ZNF451"
                /locus_tag="RP3-417I1.1-004"
mRNA            join(65657. .65831,66223. .66306,74473. .74553)
                /gene="ZNF451"
                /locus_tag="RP3-417I1.1-004"
                /product="zinc finger protein 451"
                /note="match: ESTs: BF699031.1"
CDS             join(65666. .65686,66223. .66306,74473. .74553,
                100146. .100271,AL136311.7:4475. .4586,
                AL136311.7:8788. .8938,AL136311.7:10490. .10616,

Alignment Scores:
```

Pred. No.:	6.53e-07	Length:	101666
Score:	141.00	Matches:	28
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	61.6%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x AL450489 (1-101666)

```
Qy          7 IleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn 26
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      22725 ATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAAAT 22784

Qy         27 GluAspAspIleGlnPheValSer 34
              |||||||||||||||||||||||
Db      22785 GAAGACGACATTCAGTTTGTCAGT 22808
```

RESULT 30
AC097601/c

```
LOCUS       AC097601               235278 bp    DNA     linear   HTG 10-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-20C17, WORKING DRAFT SEQUENCE.
ACCESSION   AC097601
VERSION     AC097601.7  GI:30520446
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 235278)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 235278)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 235278)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:24819438.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GHSM
                Center clone name: CH230-20C17
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 228658 bases at least Q40
                Consensus quality: 230528 bases at least Q30
                Consensus quality: 231587 bases at least Q20
                Estimated insert size: 238639; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *         1   235278: contig of 235278 bp in length.
FEATURES              Location/Qualifiers
     source           1. .235278
                      /organism="Rattus norvegicus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:10116"
                      /clone="CH230-20C17"
     misc_feature     1. .1191
                      /note="wgs_end_extension
                      clone_end:Sp6"
     misc_feature     3327. .4193
                      /note="clone_boundary
                      clone_end:Sp6
```

```
                     site:EcoRI
                     end_sequence:BH318632"
     misc_feature    231272. .232079
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH318587"
     misc_feature    233552. .235278
                     /note="wgs_end_extension
                     clone_end:T7"
ORIGIN

Alignment Scores:
Pred. No.:                3.25e-05        Length:          235278
Score:                    131.00          Matches:         26
Percent Similarity:       92.9%           Conservative:    0
Best Local Similarity:    92.9%           Mismatches:      2
Query Match:              57.2%           Indels:          0
DB:                       14              Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AC097601 (1-235278)

Qy          7 IleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn 26
              ||||||||||||||||||||||||||||||||   ||||||||||||   |||||||||
Db     212867 ATAATAGAATCTGTCCCACCAGCTGGGCCCGAGCCATCAGAGTCAACGGCGGATGAAAAC 212808

Qy         27 GluAspAspIleGlnPheValSer 34
              |||||||||||||||||||||||
Db     212807 GAAGATGACATTCAGTTTGTTAGT 212784

RESULT 31
AC099359
LOCUS       AC099359               267870 bp    DNA     linear   HTG 26-SEP-2002
DEFINITION  Rattus norvegicus clone CH230-168C19, *** SEQUENCING IN PROGRESS
            ***, 19 unordered pieces.
ACCESSION   AC099359
VERSION     AC099359.5  GI:23321934
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 267870)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 267870)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 267870)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 26, 2002 this sequence version replaced gi:21729246.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequening reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/).  As a result, the
            sequence may extend beyond the ends of the clone and there may be
            contigs that consist entirely of whole genome shotgun sequence
            reads.  Both end sequences and whole genome shotgun sequence only
            contigs will be indicated in the feature table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GGIS
                Center clone name: CH230-168C19
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 215469 bases at least Q40
                Consensus quality: 221003 bases at least Q30
                Consensus quality: 224760 bases at least Q20
                Estimated insert size: 262408; sum-of-contigs estimation
                Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
            * NOTE: This sequence may represent more than one clone.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 19 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   100680: contig of 100680 bp in length
            *  100681   100780: gap of unknown length
            *  100781   104280: contig of 3500 bp in length
            *  104281   104380: gap of unknown length
```

```
          *   104381   115029: contig of 10649 bp in length
          *   115030   115129: gap of unknown length
          *   115130   118866: contig of 3737 bp in length
          *   118867   118966: gap of unknown length
          *   118967   132336: contig of 13370 bp in length
          *   132337   132436: gap of unknown length
          *   132437   244300: contig of 111864 bp in length
          *   244301   244400: gap of unknown length
          *   244401   249533: contig of 5133 bp in length
          *   249534   249633: gap of unknown length
          *   249634   250756: contig of 1123 bp in length
          *   250757   250856: gap of unknown length
          *   250857   251937: contig of 1081 bp in length
          *   251938   252037: gap of unknown length
          *   252038   253485: contig of 1448 bp in length
          *   253486   253585: gap of unknown length
          *   253586   254826: contig of 1241 bp in length
          *   254827   254926: gap of unknown length
          *   254927   256264: contig of 1338 bp in length
          *   256265   256364: gap of unknown length
          *   256365   257446: contig of 1082 bp in length
          *   257447   257546: gap of unknown length
          *   257547   258829: contig of 1283 bp in length
          *   258830   258929: gap of unknown length
          *   258930   260251: contig of 1322 bp in length
          *   260252   260351: gap of unknown length
          *   260352   262465: contig of 2114 bp in length
          *   262466   262565: gap of unknown length
          *   262566   263895: contig of 1330 bp in length
          *   263896   263995: gap of unknown length
          *   263996   266110: contig of 2115 bp in length
          *   266111   266210: gap of unknown length
          *   266211   267870: contig of 1660 bp in length.
FEATURES             Location/Qualifiers
     source          1. .267870
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-168C19"
     misc_feature    1. .1096
                     /note="wgs_contig"
     misc_feature    4169. .5400
                     /note="wgs_contig"
     gap             100681. .100780
                     /estimated_length=unknown
     misc_feature    100781. .101968
                     /note="wgs_contig"
     gap             104281. .104380
                     /estimated_length=unknown
     misc_feature    104381. .106219
                     /note="wgs_contig"
     gap             115030. .115129
                     /estimated_length=unknown
     misc_feature    115130. .116381
                     /note="wgs_contig"
     gap             118867. .118966
                     /estimated_length=unknown
     misc_feature    118967. .120659
                     /note="wgs_contig"
     gap             132337. .132436
                     /estimated_length=unknown
     gap             244301. .244400
                     /estimated_length=unknown
     misc_feature    244401. .246510
                     /note="wgs_contig"
     gap             249534. .249633
                     /estimated_length=unknown
     gap             250757. .250856
                     /estimated_length=unknown
     gap             251938. .252037
                     /estimated_length=unknown
     gap             253486. .253585
                     /estimated_length=unknown
     gap             254827. .254926
                     /estimated_length=unknown
     gap             256265. .256364
                     /estimated_length=unknown
     gap             257447. .257546
                     /estimated_length=unknown
     gap             258830. .258929
                     /estimated_length=unknown
     gap             260252. .260351
                     /estimated_length=unknown
     gap             262466. .262565
                     /estimated_length=unknown
     gap             263896. .263995
                     /estimated_length=unknown

Alignment Scores:
Pred. No.:                3.72e-05      Length:          267870
Score:                    131.00        Matches:         26
Percent Similarity:       92.9%         Conservative:    0
Best Local Similarity:    92.9%         Mismatches:      2
Query Match:              57.2%         Indels:          0
DB:                       14            Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AC099359 (1-267870)

Qy          7 IleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn 26
              ||||||||||||||||||||||||||||||||   ||||||||||||   |||||||||
Db     173414 ATAATAGAATCTGTCCCACCAGCTGGGCCCGAGCCATCAGAGTCAACGGCGGATGAAAAC 173473

Qy         27 GluAspAspIleGlnPheValSer 34
              |||||||||||||||||||||||
Db     173474 GAAGATGACATTCAGTTTGTTAGT 173497

RESULT 32
AC163142/c
LOCUS       AC163142               246801 bp    DNA     linear   HTG 01-JUL-2005
DEFINITION  Bos taurus clone CH240-79H13, *** SEQUENCING IN PROGRESS ***, 25
            unordered pieces.
ACCESSION   AC163142
VERSION     AC163142.2  GI:68300456
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 246801)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
```

```
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 246801)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 246801)
  AUTHORS   Cow Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jun 29, 2005 this sequence version replaced gi:66955947.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: FFBZ
                Center clone name: CH240-79H13
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 236819 bases at least Q40
                Consensus quality: 239101 bases at least Q30
                Consensus quality: 241053 bases at least Q20
                Estimated insert size: 239811; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    10508: contig of 10508 bp in length
            *    10509    10558: gap of 50 bp
            *    10559    77693: contig of 67135 bp in length
            *    77694    78724: gap of 1031 bp
            *    78725   100990: contig of 22266 bp in length
            *   100991   101040: gap of 50 bp
            *   101041   104443: contig of 3403 bp in length
            *   104444   104493: gap of 50 bp
            *   104494   125739: contig of 21246 bp in length
            *   125740   125789: gap of 50 bp
            *   125790   125790: contig of 1 bp in length
            *   125791   125917: gap of 127 bp
            *   125918   125957: contig of 40 bp in length
            *   125958   126007: gap of 50 bp
            *   126008   159696: contig of 33689 bp in length
            *   159697   159746: gap of 50 bp
            *   159747   172950: contig of 13204 bp in length
            *   172951   173000: gap of 50 bp
            *   173001   191200: contig of 18200 bp in length
            *   191201   191342: gap of 142 bp
            *   191343   211597: contig of 20255 bp in length
            *   211598   211647: gap of 50 bp
            *   211648   222857: contig of 11210 bp in length
            *   222858   222907: gap of 50 bp
            *   222908   224689: contig of 1782 bp in length
            *   224690   224739: gap of 50 bp
            *   224740   225906: contig of 1167 bp in length
            *   225907   226006: gap of unknown length
            *   226007   227761: contig of 1755 bp in length
            *   227762   227811: gap of 50 bp
            *   227812   233945: contig of 6134 bp in length
            *   233946   234045: gap of unknown length
            *   234046   235052: contig of 1007 bp in length
            *   235053   235152: gap of unknown length
            *   235153   236234: contig of 1082 bp in length
            *   236235   236334: gap of unknown length
            *   236335   237346: contig of 1012 bp in length
            *   237347   237446: gap of unknown length
            *   237447   238565: contig of 1119 bp in length
            *   238566   238665: gap of unknown length
            *   238666   240288: contig of 1623 bp in length
            *   240289   240388: gap of unknown length
            *   240389   241674: contig of 1286 bp in length
            *   241675   241774: gap of unknown length
            *   241775   243183: contig of 1409 bp in length
            *   243184   243283: gap of unknown length
            *   243284   244824: contig of 1541 bp in length
            *   244825   244924: gap of unknown length
            *   244925   246801: contig of 1877 bp in length.
FEATURES             Location/Qualifiers
     source          1. .246801
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9913"
                     /clone="CH240-79H13"
     gap             10509. .10558
                     /estimated_length=50
     gap             77694. .78724
                     /estimated_length=1031
     gap             100991. .101040
                     /estimated_length=50
     gap             104444. .104493
                     /estimated_length=50
     gap             125740. .125789
                     /estimated_length=50
     gap             125791. .125917
                     /estimated_length=127
     gap             125958. .126007
                     /estimated_length=50
     gap             159697. .159746
                     /estimated_length=50
```

```
     gap             172951. .173000
                     /estimated_length=50
     gap             191201. .191342
                     /estimated_length=142
     gap             211598. .211647
                     /estimated_length=50
     gap             222858. .222907
                     /estimated_length=50
     gap             224690. .224739
                     /estimated_length=50
     gap             225907. .226006
                     /estimated_length=unknown
     gap             227762. .227811
                     /estimated_length=50
     gap             233946. .234045

Alignment Scores:
Pred. No.:                0.000115        Length:         246801
Score:                    127.00          Matches:        25
Percent Similarity:       84.4%           Conservative:   2
Best Local Similarity:    78.1%           Mismatches:     5
Query Match:              55.5%           Indels:         0
DB:                       14              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AC163142 (1-246801)

Qy          7 IleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn 26
              |||||||||  :::|||||||||||||||||||||||||||||||||||   :::||||||
Db     122169 ATAATAGAATGTATCCCTCCAGCTGGGCCTGAGGCGTCCGAGTCAACCGTGGAGGAAAAC 122110

Qy         27 GluAspAspIleGlnPheValSerGluGlyProSer 38
              |||||||||||||||||||||||          |||
Db     122109 GAAGATGACATCCAGTTTGTGAGTGTAAGTAGCAGC 122074

RESULT 33
BX932889
LOCUS       BX932889                1300 bp    mRNA    linear   VRT 02-FEB-2004
DEFINITION  Gallus gallus finished cDNA, clone ChEST916m17.
ACCESSION   BX932889
VERSION     BX932889.1  GI:41633417
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1300)
  AUTHORS   Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
            Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
            Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
            Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
            Tickle,C. and Wilson,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
COMMENT     BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
            sequencing project.
            This sequence is from the
            BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
            from a library constructed by Elizabeth Bosch. cDNA was prepared
            from RNA extracted from whole embryo, normalised, and poly
            A-trimmed. EcoRI-NotI cut cDNA was then ligated into the vector.
            Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host:
            Escherichia coli DH10B.
FEATURES              Location/Qualifiers
     source           1. .1300
                      /organism="Gallus gallus"
                      /mol_type="mRNA"
                      /strain="White Leghorn, Hisex"
                      /db_xref="taxon:9031"
                      /clone="ChEST916m17"
                      /clone_lib="CSEQCHN03"
                      /dev_stage="stage 20-21"
ORIGIN

Alignment Scores:
Pred. No.:                0.0282          Length:         1300
Score:                    91.00           Matches:        20
Percent Similarity:       62.5%           Conservative:   5
Best Local Similarity:    50.0%           Mismatches:     15
Query Match:              39.7%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BX932889 (1-1300)

Qy          4 GlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThr 23
              |||       |||:::||||||                    |||:::|||   |||
Db        101 GGCCGTGGGATAATGGAGTCTTGCAGCTCTACCAGCATTTCTGAGAGCAAATCGCCGACT 160

Qy         24 AspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeuGlu 43
                |||:::||||||:::|||:::||||||   |||   |||   ||||||||||||||
Db        161 TCTGAAAGCGAAGATGAGATAGAGTTTGTTGGAGAAGAACCCTTAAGACCTGTACTTGAA 220

RESULT 34
BX932270
LOCUS       BX932270                824 bp    mRNA    linear   VRT 30-MAR-2004
DEFINITION  Gallus gallus finished cDNA, clone ChEST352c18.
ACCESSION   BX932270
VERSION     BX932270.2  GI:46017353
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 824)
  AUTHORS   Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
            Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
            Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
            Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
            Tickle,C. and Wilson,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAR-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
COMMENT     On Apr 1, 2004 this sequence version replaced gi:41632798.
            BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
            sequencing project.
            This sequence is from the
            BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
            from a library constructed by Elizabeth Bosch. cDNA was prepared
            from RNA extracted from heart, normalised, and poly A-trimmed.
            EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
            pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
            coli DH10B.
FEATURES              Location/Qualifiers
     source           1. .824
                      /organism="Gallus gallus"
                      /mol_type="mRNA"
                      /strain="White Leghorn, Hisex"
                      /db_xref="taxon:9031"
                      /clone="ChEST352c18"
                      /clone_lib="CSEQCHN70"
                      /dev_stage="stage 36"
ORIGIN

Alignment Scores:
Pred. No.:                0.0437          Length:         824
Score:                    88.00           Matches:        19
Percent Similarity:       64.9%           Conservative:   5
Best Local Similarity:    51.4%           Mismatches:     13
Query Match:              38.4%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BX932270 (1-824)

Qy          7 IleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn 26
```

```
Db        560 ATAATGGAGTCTTGCAGCTCTACCAGCATTTCTGAGAGCAAATCGCCGACTTCTGAAAGC 619

Qy         27 GluAspAspIleGlnPheValSerGluGlyProSerArgProValLeuGlu 43

Db        620 GAAGATGAGATAGAGTTTGTTGGAGAAGAACCCTTAAGACCTGTACTTGAA 670

RESULT 35
BC030058
LOCUS       BC030058               2628 bp    mRNA    linear   PRI 06-OCT-2003
DEFINITION  Homo sapiens transcription termination factor, RNA polymerase II,
            mRNA (cDNA clone MGC:39244 IMAGE:5457833), complete cds.
ACCESSION   BC030058
VERSION     BC030058.1  GI:20987695
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2628)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2628)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAY-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Pjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 42 Row: p Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 19923173.

FEATURES             Location/Qualifiers
     source          1. .2628
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:39244 IMAGE:5457833"
                     /tissue_type="Brain, neuroblastoma"
                     /clone_lib="NIH_MGC_47"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .2628
                     /gene="TTF2"
                     /note="synonym: HuF2"
                     /db_xref="GeneID:8458"
                     /db_xref="MIM:604718"
     CDS             5. .2035
                     /gene="TTF2"
                     /codon_start=1
                     /product="TTF2 protein"
                     /protein_id="AAH30058.1"
                     /db_xref="GI:20987696"
                     /db_xref="GeneID:8458"
                     /db_xref="MIM:604718"
                     /translation="MEEVRCPEHGTFCFLKTGVRDGPNKGKSFYVCRADTCSFVRATD
                     IPVSHCLLHEDFVVELQGLLLPQDKKEYRLFFRCIRSKAEGKRWCGSIPWQDPDSKEH
                     SVSNKSQHASETFHHSSNWLRNPFKVLDKNQEPALWKQLIKGEGEEKKADKKQREKGD
                     QLFDQKKEQKPEMMEKDLSSGLVPKKKQSVVQEKKQEEGAEIQCEAETGGTHKRDFSE
                     IKSQQCQGNELTRPSASSQEKSSGKSQDVQRESEPLREKVTQLLPQNVHSHNSISKPQ
                     KGGPLNKEYTNWEAKETKAKDGPSIQATQKSLPQGHFQERPETHSVPAPGGPAAQAAP
                     AAPGLSLGEGREAATSSDDEEEDDVVFVSSKPGSPLLFDSTLDLETKENLQFPDRSVQ
                     RKVSPASGVSKKVEPSDPVARRVYLTTQLKQKKSTLASVNIQALPDKGQKLIKQIQEL
                     EEVLSGLTLSPEQGTNEKSNSQVPQQSHFTKTTTGPPHLVPPQPLPRRGTQPVGSLEL
                     KSACQVTAGGSSQCYRGHTNQDHVHAVWKITSEAIGQLHRSLESCPGETVVAEDPAGL
                     KVPLLLHQKQALAWLLWRESQKPQGGILADDMGLGKTLTMIALILTQKNQEKKEEKEK
                     STALTWLSKDGRQKCLNSLPFPTSFEPPKRGTSSAKKGHLWSYLITLLENQY"
     misc_feature    1721. .1852
                     /gene="TTF2"
                     /note="SNF2_N; Region: SNF2 family N-terminal domain. This
                     domain is found in proteins involved in a variety of
                     processes including transcription regulation (e.g., SNF2,
                     STH1, brahma, MOT1) , DNA repair (e.g., ERCC6, RAD16,
                     RAD5), DNA recombination (e.g., RAD54), and chromatin
                     unwinding (e.g., ISWI) as well as a variety of other
                     proteins with little functional information (e.g.,
                     lodestar, ETL1)"
                     /db_xref="CDD:pfam00176"
ORIGIN

Alignment Scores:
Pred. No.:                0.311          Length:          2628
Score:                    85.50          Matches:         19
Percent Similarity:       53.2%          Conservative:    6
Best Local Similarity:    40.4%          Mismatches:      15
Query Match:              37.3%          Indels:          7
DB:                       8              Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x BC030058 (1-2628)

Qy          3 ProGlySerGluIleIleGluSerValProProAla--------------------Gly 15

Db        977 CCTGGAGGACCAGCGGCTCAGGCTGCACCAGCAGCACCAGGGCTTTCCCTGGGTGAGGGC 1036

Qy         16 ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35

Db       1037 CGTGAAGCTGCCACAAGCAGTGACGACGAGGAGGAAGATGATGTTGTTTTTGTTTCCTCT 1096

Qy         36 GlyProSerArgProValLeu 42

Db       1097 AAGCCTGGGAGCCCCCTACTC 1117

RESULT 36
CQ728000
LOCUS       CQ728000               3593 bp    DNA     linear   PAT 03-FEB-2004
```

```
DEFINITION  Sequence 13934 from Patent WO02068579.
ACCESSION   CQ728000
VERSION     CQ728000.1  GI:42295593
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 13934 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .3593
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                0.431          Length:        3593
Score:                    85.50          Matches:       19
Percent Similarity:       53.2%          Conservative:  6
Best Local Similarity:    40.4%          Mismatches:    15
Query Match:              37.3%          Indels:        7
DB:                       6              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x CQ728000 (1-3593)

Qy          3 ProGlySerGluIleIleGluSerValProProAla----------------------Gly 15
              ||||||           ::::::   |||   |||                      |||
Db        993 CCTGGAGGACCAGCGGCTCAGGCTGCACCAGCAGCACCAGGGCTTTCCCTGGGTGAGGGC 1052

Qy         16 ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                 ||||||:::   |||:::   ||||||   |||||||||:::   |||||||||
Db       1053 CGTGAAGCTGCCACAAGCAGTGACGACGAGGAGGAAGATGATGTTGTTTTTGTTTCCTCT 1112

Qy         36 GlyProSerArgProValLeu 42
                |||      |||:::|||
Db       1113 AAGCCTGGGAGCCCCCTACTC 1133

RESULT 37
AF073771
LOCUS       AF073771               3612 bp    mRNA    linear   PRI 15-OCT-1998
DEFINITION  Homo sapiens RNA polymerase II termination factor mRNA, complete
            cds.
ACCESSION   AF073771
VERSION     AF073771.1  GI:3702845
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 34 to 3612)
  AUTHORS   Liu,M., Xie,Z. and Price,D.H.
  TITLE     A human RNA polymerase II transcription termination factor is a
            SWI2/SNF2 family member
  JOURNAL   J. Biol. Chem. 273 (40), 25541-25544 (1998)
   PUBMED   9748214
REFERENCE   2  (bases 34 to 3612)
  AUTHORS   Liu,M., Xie,Z. and Price,D.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUN-1998) Biochemistry, Univ. of Iowa, 51 Newton
            Road, Iowa City, IA 52242, USA
REFERENCE   3  (bases 1 to 3612)
  AUTHORS   Liu,M., Xie,Z. and Price,D.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-SEP-1998) Biochemistry, Univ. of Iowa, 51 Newton
            Road, Iowa City, IA 52242, USA
  REMARK    Sequence updated by submitter
FEATURES             Location/Qualifiers
     source          1. .3612
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             21. .3509
                     /note="HuF2; factor 2; SWI2/SNF2 family member"
                     /codon_start=1
                     /product="RNA polymerase II termination factor"
                     /protein_id="AAC64044.1"
                     /db_xref="GI:3702846"
                     /translation="MEEVRCPEHGTFCFLKTGVRDGPNKGKSFYVCRADTCSFVRATD
                     IPVSHCLLHEDFVVELQGLLLPQDKKEYRLFFRCIRSKAEGKRWCGSIPWQDPDSKEH
                     SVSNKSQHASETFHHSSNWLRNPFKVLDKNQEPALWKQLIKGEGEEKKADKKQREKGD
                     QLFDQKKEQKPEMMEKDLSSGLVPKKKQSVVQEKKQEEGAEIQCEAETGGTHKRDFSE
                     IKSQQCQGNELTRPSASSQEKSSGKSQDVQRESEPLREKVTQLLPQNVHSHNSISKPQ
                     KGGPLNKEYTNWEAKETKAKDGPSIQATQKSLPQGHFQERPETHSVPAPGGPAAQAAP
                     AAPGLSLGEGREAATSSDDEEEDDVVFVSSKPGSPLLFDSTLDLETKENLQFPDRSVQ
                     RKVSPASGVSKKVEPSDPVARRVYLTTQLKQKKSTLASVNIQALPDKGQKLIKQIQEL
                     EEVLSGLTLSPEQGTNEKSNSQVPQQSHFTKTTTGPPHLVPPQPLPRRGTQPVGSLEL
                     KSACQVTAGGSSQCYRGHTNQDHVHAVWKITSEAIGQLHRSLESCPGETVVAEDPAGL
                     KVPLLLHQKQALAWLLWRESQKPQGGILADDMGLGKTLTMIALILTQKNQEKKEEKEK
                     STALTWLSKDDSCDFTSHGTLIICPASLIHHWKNEVEKRVNSNKLRVYLYRGPNRDSR
                     ARVLSTYDIVITTYSLVAKEIPTNKQEAEIPGANLNVEGTSTPLLRIAWARIILDEAH
                     NVKNPRVQTSIAVCKLQACARWAVTGTPIQNNLLDMYSLLKFLRCSPFDEFNLWRSQV
                     DNGSKKGGERLSILTKSLLLRRTKDQLDSTGRPLVILPQRKFQLHHLKLSEDEETVYN
                     VFFARSRSALQSYLKRHESRGNQSGRSPNNPFSRVALEFGSEEPRHSEAADSPRSSTV
                     HILSQLLRLRQCCCHLSLLKSALDPMELKGEGLVLSLEEQLSALTLSELRDSEPSSTV
                     SLNGTFFKMELFEGMRESTKISSLLAELEAIQRNSASQKSVIVSQWTNMLKVVALHLK
                     KHGLTYATIDGSVNPKQRMDLVEAFNHSRGPQVMLISLLAGGVGLNLTGGNHLFLLDM
                     HWNPSLEDQACDRIYRVGQQKDVVIHRFVCEGTVEEKILQLQEKKKDLAKQVLSGSGE
                     SVTKLTLADLRVLFGI"
ORIGIN

Alignment Scores:
Pred. No.:                0.433          Length:        3612
Score:                    85.50          Matches:       19
Percent Similarity:       53.2%          Conservative:  6
Best Local Similarity:    40.4%          Mismatches:    15
Query Match:              37.3%          Indels:        7
DB:                       8              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x AF073771 (1-3612)

Qy          3 ProGlySerGluIleIleGluSerValProProAla----------------------Gly 15
              ||||||           ::::::   |||   |||                      |||
Db        993 CCTGGAGGACCAGCGGCTCAGGCTGCACCAGCAGCACCAGGGCTTTCCCTGGGTGAGGGC 1052

Qy         16 ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                 ||||||:::   |||:::   ||||||   |||||||||:::   |||||||||
Db       1053 CGTGAAGCTGCCACAAGCAGTGACGACGAGGAGGAAGATGATGTTGTTTTTGTTTCCTCT 1112

Qy         36 GlyProSerArgProValLeu 42
                |||      |||:::|||
Db       1113 AAGCCTGGGAGCCCCCTACTC 1133

RESULT 38
CQ981168
LOCUS       CQ981168               4629 bp    DNA     linear   PAT 25-JAN-2005
DEFINITION  Sequence 23 from Patent EP1498424.
ACCESSION   CQ981168
VERSION     CQ981168.1  GI:58190427
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Rosenthal,A., Hermann,K., Heiden,E., Pilarsky,C., Bruemmendorf,T.,
            Staub,E., Roepcke,S., Mennerich,D., Kinnemann,H. and Li,X.
```

```
  TITLE     Human nucleic acid sequences from lung tumours
  JOURNAL   Patent: EP 1498424-A 23 19-JAN-2005;
            Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esmeralda (DE);
            Rosenthal, Andre (DE)
FEATURES             Location/Qualifiers
     source          1. .4629
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    4178
                     /note="n is a, c, g, or t"
     misc_feature    4211
                     /note="n is a, c, g, or t"
     misc_feature    4223
                     /note="n is a, c, g, or t"
     misc_feature    4229
                     /note="n is a, c, g, or t"
     misc_feature    4255
                     /note="n is a, c, g, or t"
     misc_feature    4261
                     /note="n is a, c, g, or t"
     misc_feature    4293
                     /note="n is a, c, g, or t"
     misc_feature    4377
                     /note="n is a, c, g, or t"
     misc_feature    4380
                     /note="n is a, c, g, or t"
     misc_feature    4498
                     /note="n is a, c, g, or t"
     misc_feature    4526
                     /note="n is a, c, g, or t"
ORIGIN

Alignment Scores:
Pred. No.:                0.56          Length:          4629
Score:                    85.50         Matches:         19
Percent Similarity:       53.2%         Conservative:    6
Best Local Similarity:    40.4%         Mismatches:      15
Query Match:              37.3%         Indels:          7
DB:                       6             Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x CQ981168 (1-4629)

Qy          3 ProGlySerGluIleIleGluSerValProProAla----------------------Gly 15
              ||||||              ::::::   |||   |||                    |||
Db       1003 CCTGGAGGACCAGCGGCTCAGGCTGCACCAGCAGCACCAGGGCTTTCCCTGGGTGAGGGC 1062

Qy         16 ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
              ||||||:::   |||:::   ||||||   |||||||||:::   |||||||||
Db       1063 CGTGAAGCTGCCACAAGCAGTGACGACGAGGAGGAAGATGATGTTGTTTTTGTTTCCTCT 1122

Qy         36 GlyProSerArgProValLeu 42
                |||      |||:::|||
Db       1123 AAGCCTGGGAGCCCCCTACTC 1143

RESULT 39
AF080255
LOCUS       AF080255               4629 bp    mRNA    linear   PRI 16-AUG-1999
DEFINITION  Homo sapiens lodestar protein mRNA, complete cds.
ACCESSION   AF080255
VERSION     AF080255.1  GI:5733121
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 4629)
  AUTHORS   Leonard,D.M. and Legerski,R.J.
  TITLE     Sequence of hLODESTAR mRNA
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4629)
  AUTHORS   Leonard,D.M. and Legerski,R.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-1998) Molecular Genetics, M. D. Anderson Cancer
            Center, 1515 Holcombe Blvd., Houston, TX 77030, USA
FEATURES             Location/Qualifiers
     source          1. .4629
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1p22"
     CDS             31. .3519
                     /note="similar to Drosophila lodestar; related to SNF2
                     gene family"
                     /codon_start=1
                     /product="lodestar protein"
                     /protein_id="AAD49435.1"
                     /db_xref="GI:5733122"
                     /translation="MEEVRCPEHGTFCFLKTGVRDGPNKGKSFYVCRADTCSFVRATD
                     IPVSHCLLHEDFVVELQGLPLPQDKKEYRLFFRCIGSKAEGKRWCGSIPWQDPDSKEH
                     SVSNKSQHASETFHHSSNWLRNPFKVLDKNQEPALWKQLIKGEGEEKKADKKQREKGD
                     QLFDQKKEQKPEMMEKDLSSGLVPKKKQSVVQEKKQEEGAEIQCEAETGGTHKRDFSE
                     IKSQQCQGNELTRPSASSQEKSSGKSQDVQRESEPLREKVTQLLPQNVHSHNSISKPQ
                     KGGPLNKEYTNWEAKETKAKDGPSIQATQKSLPQGHFQERPETHSVPAPGGPAAQAAP
                     AAPGLSLGEGREAATSSDDEEEDDVVFVSSKPGSPLLFDSTLDLETKENLQFPDRSVQ
                     RKVSPASGVSKKVEPSDPVARRVYLTTQLKQKKSTLASVNIQALPDKGQKLIKQIQEL
                     EEVLSGLTLSPEQGTNEKSNSQVPQQSHFTKTTTGPPHLVPPQPLPRRGTQPVGSLEL
                     KSACQVTAGGSSQCYRGHTNQDHVHAVWKITSEAIGQLHRSLESCPGETVVAEDPAGL
                     KVPLLLHQKQALAWLLWRESQKPQGGILADDMGLGKTLTMIALILTQKNQEKKEEKEK
                     STALTWLSKDDSCDFTSHGTLIICPASLIHHWKNEVEKRVNSNKLRVYLYHGPNRDSR
                     ARVLSTYDIVITTYSLVAKEIPTNKQEAEIPGANLNVEGTSTPLLRIAWARIILDEAH
                     NVKNPRVQTSIAVCKLQACARWAVTGTPIQNNLLDMYSLLKFLRCSPFDEFNLWRSQV
                     DNGSKKGGERLSILTKSLLLRRTKDQLDSTGRPLVILPQRKFQLHHLKLSEDEETVYN
                     VFFARSRSALQSYLKRHESRGNQSGRSPNNPFSRVALEFGSEEPRHSEAADSPRSSTV
                     HILSQLLRLRQCCCHLSLLKSALDPMELKGEGLVLSLEEQLSALTLSELRDSEPSSTV
                     SLNGTFFKMELFEGMRESTKISSLLAELEAIQRNSASQKSVIVSQWTNMLKVVALHLK
                     KHGLTYATIDGSVNPKQRMDLVEAFNHSRGPQVMLISLLAGGVGLNLTGGNHLFLLDM
                     HWNPSLEDQACDRIYRVGQQKDVVIHRFVCEGTVEEKILQLQEKKKDLAKQVLSGSGE
                     SVTKLTLAHLRVLFGI"
ORIGIN

Alignment Scores:
Pred. No.:                0.56          Length:          4629
Score:                    85.50         Matches:         19
Percent Similarity:       53.2%         Conservative:    6
Best Local Similarity:    40.4%         Mismatches:      15
Query Match:              37.3%         Indels:          7
DB:                       8             Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x AF080255 (1-4629)

Qy          3 ProGlySerGluIleIleGluSerValProProAla----------------------Gly 15
              ||||||              ::::::   |||   |||                    |||
Db       1003 CCTGGAGGACCAGCGGCTCAGGCTGCACCAGCAGCACCAGGGCTTTCCCTGGGTGAGGGC 1062

Qy         16 ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
              ||||||:::   |||:::   ||||||   |||||||||:::   |||||||||
Db       1063 CGTGAAGCTGCCACAAGCAGTGACGACGAGGAGGAAGATGATGTTGTTTTTGTTTCCTCT 1122

Qy         36 GlyProSerArgProValLeu 42
                |||      |||:::|||
Db       1123 AAGCCTGGGAGCCCCCTACTC 1143

RESULT 40
AL391476
LOCUS       AL391476               171595 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-229A19 on chromosome 1 Contains
            the 3' end of the TTF2 gene for RNA polymerase II transcription
            termination factor, the TRIM45 gene for tripartite motif-containing
            45, a ribosomal protein S15a (RPS15A) pseudogene, the gene for
            immune costimulatory protein B7-H4 and a CpG island, complete
            sequence.
ACCESSION   AL391476
VERSION     AL391476.20  GI:15131484
```

```
KEYWORDS    HTG; B7-H4; CpG island; RPS15A; transcription termination factor;
            TRIM45; tripartite motif; TTF2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 171595)
  AUTHORS   Wallis,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 9, 2001 this sequence version replaced gi:14970375.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-229A19 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .171595
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-229A19"
                     /clone_lib="RPCI-11.1"
     misc_feature    1
                     /note="Clone_left_end: RP11-229A19"
     gene            join(AL445231.8:124209. .124280,AL445231.8:124361. .124463,
                     AL445231.8:126293. .126379,10723. .10789,11932. .12921,
                     13242. .13356,13703. .13885,14978. .15105,16630. .16711,
                     18890. .19009,21080. .21230,23479. .23584,25863. .26044,
                     27117. .27270,27594. .27685,28390. .28472,28879. .29012,
                     29793. .29963,32371. .32429,33211. .33361,34413. .34495,
                     35895. .35969,38442. .39932)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-001"
     mRNA            join(AL445231.8:124209. .124280,AL445231.8:124361. .124463,
                     AL445231.8:126293. .126379,10723. .10789,11932. .12921,
                     13242. .13356,13703. .13885,14978. .15105,16630. .16711,
                     18890. .19009,21080. .21230,23479. .23584,25863. .26044,
                     27117. .27270,27594. .27685,28390. .28472,28879. .29012,
                     29793. .29963,32371. .32429,33211. .33361,34413. .34495,
                     35895. .35969,38442. .39932)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-001"
                     /product="transcription termination factor, RNA polymerase
                     II"
                     /note="match: ESTs: Em:AA227424.1 Em:AI991355.1
                     Em:AL039234.1 Em:AL596888.1 Em:AW198215.1 Em:BE273891.1
                     Em:BG337302.1 Em:BG826554.1 Em:BG828016.1 Em:BM011302.1
                     Em:BM835083.1 Em:BQ220407.1 Em:BQ423665.1 Em:BQ921458.1
                     Em:BU076149.1 Em:BX281340.1 Em:BX642926.1 Em:CA430266.1
                     Em:CA867234.1 Em:H81916.1 Em:N40873.1
                     match: cDNAs: Em:AF073771.1 Em:AF080255.1"
     CDS             join(AL445231.8:124253. .124280,AL445231.8:124361. .124463,
                     AL445231.8:126293. .126379,10723. .10789,11932. .12921,
                     13242. .13356,13703. .13885,14978. .15105,16630. .16711,
                     18890. .19009,21080. .21230,23479. .23584,25863. .26044,
                     27117. .27270,27594. .27685,28390. .28472,28879. .29012,
                     29793. .29963,32371. .32429,33211. .33361,34413. .34495,
                     35895. .35969,38442. .38586)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-001"
                     /standard_name="OTTHUMP00000013818"
                     /note="match: proteins: Tr:O75921 Tr:Q9UNY4"
                     /codon_start=1
                     /product="transcription termination factor, RNA polymerase
                     II"
                     /protein_id="CAI12738.1"
                     /db_xref="GI:55959604"
                     /db_xref="GOA:Q5VVU8"
                     /db_xref="InterPro:IPR000330"
                     /db_xref="InterPro:IPR001650"
                     /db_xref="InterPro:IPR002464"
                     /db_xref="InterPro:IPR011545"
                     /db_xref="UniProt/TrEMBL:Q5VVU8"
                     /translation="MEEVRCPEHGTFCFLKTGVRDGPNKGKSFYVCRADTCSFVRATD
                     IPVSHCLLHEDFVVELQGLLLPQDKKEYRLFFRCIRSKAEGKRWCGSIPWQDPDSKEH
                     SVSNKSQHASETFHHSSNWLRNPFKVLDKNQEPALWKQLIKGEGEEKKADKKQREKGD
                     QLFDQKKEQKPEMMEKDLSSGLVPKKKQSVVQEKKQEEGAEIQCEAETGGTHKRDFSE
                     IKSQQCQGNELTRPSASSQEKSSGKSQDVQRESEPLREKVTQLLPQNVHSHNSISKPQ
                     KGGPLNKEYTNWEAKETKAKDGPSIQATQKSLPQGHFQERPETHSVPAPGGPAAQAAP
                     AAPGLSLGEGREAATSSDDEEEDDVVFVSSKPGSPLLFDSTLDLETKENLQFPDRSVQ
                     RKVSPASGVSKKVEPSDPVARRVYLTTQLKQKKSTLASVNIQALPDKGQKLIKQIQEL
                     EEVLSGLTLSPEQGTNEKSNSQVPQQSHFTKTTTGPPHLVPPQPLPRRGTQPVGSLEL
                     KSACQVTAGGSSQCYRGHTNQDHVHAVWKITSEAIGQLHRSLESCPGETVVAEDPAGL
                     KVPLLLHQKQALAWLLWRESQKPQGGILADDMGLGKTLTMIALILTQKNQEKKEEKEK
                     STALTWLSKDDSCDFTSHGTLIICPASLIHHWKNEVEKRVNSNKLRVYLYHGPNRDSR
                     ARVLSTYDIVITTYSLVAKEIPTNKQEAEIPGANLNVEGTSTPLLRIAWARIILDEAH
                     NVKNPRVQTSIAVCKLQACARWAVTGTPIQNNLLDMYSLLKFLRCSPFDEFNLWRSQV
                     DNGSKKGGERLSILTKSLLLRRTKDQLDSTGRPLVILPQRKFQLHHLKLSEDEETVYN
                     VFFARSRSALQSYLKRHESRGNQSGRSPNNPFSRVALEFGSEEPRHSEAADSPRSSTV
                     HILSQLLRLRQCCCHLSLLKSALDPMELKGEGLVLSLEEQLSALTLSELRDSEPSSTV
                     SLNGTFFKMELFEGMRESTKISSLLAELEAIQRNSASQKSVIVSQWTNMLKVVALHLK
                     KHGLTYATIDGSVNPKQRMDLVEAFNHSRGPQVMLISLLAGGVGLNLTGGNHLFLLDM
                     HWNPSLEDQACDRIYRVGQQKDVVIHRFVCEGTVEEKILQLQEKKKDLAKQVLSGSGE
                     SVTKLTLADLRVLFGI"
     gene            join(AL445231.8:124249. .124280,AL445231.8:124361. .124463,
                     AL445231.8:126293. .126379,10723. .10789,11932. .12921,
                     13242. .13356,13703. .13885,14978. .15105,16630. .16711,
                     18890. .19437,20057. .20281)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-008"
     mRNA            join(AL445231.8:124249. .124280,AL445231.8:124361. .124463,
                     AL445231.8:126293. .126379,10723. .10789,11932. .12921,
                     13242. .13356,13703. .13885,14978. .15105,16630. .16711,
                     18890. .19437,20057. .20281)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-008"
                     /product="transcription termination factor, RNA polymerase
                     II"
                     /note="match: cDNAs: Em:BC030058.1"
     gene            join(AL445231.8:124268. .124280,AL445231.8:124361. .124467,
                     AL445231.8:126293. .126379,10723. .10789,11932. .12506)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-002"
     mRNA            join(AL445231.8:124268. .124280,AL445231.8:124361. .124467,
                     AL445231.8:126293. .126379,10723. .10789,11932. .12506)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-002"
                     /product="transcription termination factor, RNA polymerase
                     II"
                     /note="match: ESTs: Em:BE781382.1"
```

```
     gene            join(12357. .12921,13242. .13451)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-007"
     mRNA            join(12357. .12921,13242. .13451)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-007"
                     /product="transcription termination factor, RNA polymerase
                     II"
                     /note="match: ESTs: Em:AI652738.1 Em:AI991587.1
                     Em:BQ009100.1"
     polyA_signal    13427. .13432
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-007"
     polyA_site      13451
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-007"
     gene            join(29849. .29963,32371. .32429,33211. .33361,35268. .35300,
                     35895. .35969,38442. .38811)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-003"
     mRNA            join(29849. .29963,32371. .32429,33211. .33361,35268. .35300,
                     35895. .35969,38442. .38811)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-003"
                     /product="transcription termination factor, RNA polymerase
                     II"
                     /note="match: ESTs: Em:BF038049.1"
     gene            join(32387. .32429,33211. .33361,34413. .34495,35268. .35300,
                     35895. .35969,38442. .38718)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-004"
     mRNA            join(32387. .32429,33211. .33361,34413. .34495,35268. .35300,
                     35895. .35969,38442. .38718)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-004"
                     /product="transcription termination factor, RNA polymerase
                     II"
                     /note="match: ESTs: Em:AV713684.1"
     gene            join(<33233. .33361,34413. .34495,35895. .35969,38442. .38624,
                     39256. .39424)
                     /gene="TTF2"
                     /locus_tag="RP11-229A19.1-005"
     mRNA            join(<33233. .33361,34413. .34495,35895. .35969,38442. .38624,
                     39256. .39424)
                     /gene="TTF2"

Alignment Scores:
Pred. No.:                23.8          Length:          171595
Score:                    85.50         Matches:         19
Percent Similarity:       53.2%         Conservative:    6
Best Local Similarity:    40.4%         Mismatches:      15
Query Match:              37.3%         Indels:          7
DB:                       8             Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x AL391476 (1-171595)

Qy          3 ProGlySerGluIleIleGluSerValProProAla--------------------Gly 15
              ||||||             :::::::   |||   |||                    |||
Db      12619 CCTGGAGGACCAGCGGCTCAGGCTGCACCAGCAGCACCAGGGCTTTCCCTGGGTGAGGGC 12678

Qy         16 ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                ||||||:::   |||:::   ||||||   |||||||||:::   |||||||||
Db      12679 CGTGAAGCTGCCACAAGCAGTGACGACGAGGAGGAAGATGATGTTGTTTTTGTTTCCTCT 12738

Qy         36 GlyProSerArgProValLeu 42
                |||       |||:::|||
Db      12739 AAGCCTGGGAGCCCCCTACTC 12759

RESULT 41
BC072923
LOCUS       BC072923               3782 bp    mRNA    linear   VRT 03-AUG-2004
DEFINITION  Xenopus laevis MGC80407 protein, mRNA (cDNA clone MGC:80407
            IMAGE:5155019), complete cds.

ACCESSION   BC072923
VERSION     BC072923.1  GI:49116882
KEYWORDS    MGC.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 3782)
  AUTHORS   Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
            and Richardson,P.
  TITLE     Genetic and genomic tools for Xenopus research: The NIH Xenopus
            initiative
  JOURNAL   Dev. Dyn. 225 (4), 384-391 (2002)
   PUBMED   12454917
REFERENCE   2  (bases 1 to 3782)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   3  (bases 1 to 3782)
  AUTHORS   Klein,S. and Gerhard,D.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2004) National Institutes of Health, Xenopus Gene
            Collection (XGC), National Institute of Child Health and Human
            Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
            20892-7510, USA
  REMARK    NIH-MGC Project
COMMENT     Contact: XGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 153 Row: g Column: 18
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1. .3782
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
```

```
                     /db_xref="taxon:8355"
                     /clone="MGC:80407 IMAGE:5155019"
                     /tissue_type="Embryo, stage 10"
                     /clone_lib="NICHD_XGC_Emb1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .3782
                     /gene="MGC80407"
                     /db_xref="GeneID:443981"
     CDS             117. .3296
                     /gene="MGC80407"
                     /codon_start=1
                     /product="MGC80407 protein"
                     /protein_id="AAH72923.1"
                     /db_xref="GI:49116883"
                     /db_xref="GeneID:443981"
                     /translation="MDSSVSVDDKHVQNTESLCNEEEIEFVSEGPIRPVLECIDLNSS
                     DEDSKSSFVIRKVKDHVDRQKDRVASTLDRLALHVEVEKKNKAEKNKAFQEKLHGQHA
                     HGLQELECIKGLPGTDAARICVNQWLKMPGPKPGVVCSSRRNMNRSHDQSSVINTPIM
                     CPIVHCYRKYDNGQLLLGHLKRFDHSPCDPTIHLHGAPANSYACIVCLQRFNTRDKYD
                     DHVMAKARLADGHENNIRPQVINSFACPLCFLLFNLRDECLQHMSDSNHFLQAIQKDD
                     KGIPCPVPFPLYGKNVLIALCKDIPFQVKCSSCNQELRSHTELSAHFRTRCRNAHPVA
                     YSEQSIAEVVAVFLMEAYCLSCRKMLKSDMHVAKHAQRTGHKAKLITSLEESIMAFCY
                     INEGIKTPSDLCLSACNARLKPNMLKRTINNNVQVRTDFINVGKNTFKDSPTKRHKQT
                     RDLGISGTAWFCECLQKCSTEKDAEKHIMMSNGICHKCLVCSKLANDLSVIRLHMSRF
                     HGGAHLNSFIFWCKGCRAELPRTENMMAHVCECHKGHSYYFEQETWDNEPSTSTQICS
                     PPQQVVSETSSTTPETTAGHWQCHICEEMFDSEDSVRQHCNNISIHQFHKFSCDICKN
                     RFHKVETLFRHSQHQHDGEINMKYVCGLCDLYFEDEKDFHGHYNGLHSSEYGFVPEQM
                     QSLVKNEDQSFSILKPNEDCLTCGCLENYSSKAQREEANRLCLARLLQKGSLWYSCSN
                     CSATGQTYDSLKIHDCTKSPQPLAIDCVVKCSCCSTSFKDPTSALQHYHTTHCFLQKP
                     NIKQFKSLDSQKEVFEFMASGACAKAKHGKSVTLPAPKLHGETPNRCPIGMELSKPES
                     CDAENKKTGEDTQSGKIDIEEESDELPDLDFLCTMTHIVFVDLDNWAQFFSHLPGYLN
                     QGIFVWGFQGGKSNWKRPVNCKVFNYLSNTGSFFLHPHCSDSKDAADFAICMHAGRLD
                     EQLPKQIPFTILSGDKGFLELESQFKKTQRSTHILNPHHLEGDMMCALLNSISDTTQD
                     ANDDTEDELAELDEDANLNEAIKRSMTEM"
ORIGIN

Alignment Scores:
Pred. No.:                7              Length:          3782
Score:                    76.50          Matches:         17
Percent Similarity:       63.9%          Conservative:    6
Best Local Similarity:    47.2%          Mismatches:      4
Query Match:              33.4%          Indels:          9
DB:                       5              Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x BC072923 (1-3782)

Qy          17 GluAlaSerGluSerThrThrAspGlu----------------------------AsnGlu 27
               :::::|||   |||      |||:::                            ||||||
Db         120 GATTCGTCTGTATCTGTTGATGATAAACATGTTCAGAATACGGAGAGCCTTTGCAATGAG 179

Qy          28 AspAspIleGlnPheValSerGluGlyProSerArgProValLeuGlu 43
               :::::|||:::||||||||||||||||||   ||||||||||||||
Db         180 GAGGAGATTGAATTTGTCAGTGAGGGCCCAATACGACCTGTCCTGGAA 227

RESULT 42
AC156064/c
LOCUS       AC156064              286181 bp    DNA     linear   HTG 01-JUL-2005
DEFINITION  Bos taurus clone CH240-47J1, *** SEQUENCING IN PROGRESS ***, 47
            unordered pieces.
ACCESSION   AC156064
VERSION     AC156064.3  GI:68265450
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 286181)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 286181)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JAN-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 286181)
  AUTHORS   Cow Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jun 28, 2005 this sequence version replaced gi:58082191.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
```

```
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
    Center project name: FCOW
    Center clone name: CH240-47J1
-------------- Summary Statistics
    Assembly program: Atlas 3.0;
    Consensus quality: 263799 bases at least Q40
    Consensus quality: 267653 bases at least Q30
    Consensus quality: 270844 bases at least Q20
    Estimated insert size: 259829; sum-of-contigs estimation
    Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
--------------
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 47 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*       1     2106: contig of 2106 bp in length
*    2107     2176: gap of 70 bp
*    2177     4591: contig of 2415 bp in length
*    4592     4641: gap of 50 bp
*    4642     8967: contig of 4326 bp in length
*    8968     9017: gap of 50 bp
*    9018    13869: contig of 4852 bp in length
*   13870    13969: gap of unknown length
*   13970    28003: contig of 14034 bp in length
*   28004    28053: gap of 50 bp
*   28054    44714: contig of 16661 bp in length
*   44715    44764: gap of 50 bp
*   44765    46092: contig of 1328 bp in length
*   46093    46192: gap of unknown length
*   46193    47990: contig of 1798 bp in length
*   47991    48501: gap of 511 bp
*   48502    53685: contig of 5184 bp in length
*   53686    53735: gap of 50 bp
*   53736    74499: contig of 20764 bp in length
*   74500    74549: gap of 50 bp
*   74550    98298: contig of 23749 bp in length
*   98299    98348: gap of 50 bp
*   98349   117823: contig of 19475 bp in length
*  117824   117923: gap of unknown length
*  117924   119188: contig of 1265 bp in length
*  119189   119238: gap of 50 bp
*  119239   123341: contig of 4103 bp in length
*  123342   123391: gap of 50 bp
*  123392   126992: contig of 3601 bp in length
*  126993   127042: gap of 50 bp
*  127043   137392: contig of 10350 bp in length
*  137393   137442: gap of 50 bp
*  137443   138563: contig of 1121 bp in length
*  138564   145741: gap of 7178 bp
*  145742   158989: contig of 13248 bp in length
*  158990   159039: gap of 50 bp
*  159040   167491: contig of 8452 bp in length
*  167492   167708: gap of 217 bp
*  167709   168975: contig of 1267 bp in length
*  168976   169243: gap of 268 bp
*  169244   174555: contig of 5312 bp in length
*  174556   174605: gap of 50 bp
*  174606   179913: contig of 5308 bp in length
*  179914   179963: gap of 50 bp
*  179964   182608: contig of 2645 bp in length
*  182609   182871: gap of 263 bp
*  182872   184353: contig of 1482 bp in length
*  184354   184517: gap of 164 bp
*  184518   191534: contig of 7017 bp in length
*  191535   191634: gap of unknown length
*  191635   192895: contig of 1261 bp in length
*  192896   192945: gap of 50 bp
*  192946   201248: contig of 8303 bp in length
*  201249   201298: gap of 50 bp
*  201299   202873: contig of 1575 bp in length
*  202874   202923: gap of 50 bp
*  202924   213946: contig of 11023 bp in length
*  213947   214479: gap of 533 bp
*  214480   216708: contig of 2229 bp in length
*  216709   216808: gap of unknown length
*  216809   225505: contig of 8697 bp in length
*  225506   225555: gap of 50 bp
*  225556   227913: contig of 2358 bp in length
*  227914   227963: gap of 50 bp
*  227964   236492: contig of 8529 bp in length
*  236493   236592: gap of unknown length
*  236593   254614: contig of 18022 bp in length
*  254615   254664: gap of 50 bp
*  254665   256942: contig of 2278 bp in length
*  256943   257042: gap of unknown length
*  257043   258120: contig of 1078 bp in length
*  258121   258220: gap of unknown length
*  258221   259680: contig of 1460 bp in length
*  259681   259780: gap of unknown length
*  259781   261065: contig of 1285 bp in length
*  261066   261165: gap of unknown length
*  261166   262370: contig of 1205 bp in length
*  262371   262470: gap of unknown length
*  262471   263589: contig of 1119 bp in length
*  263590   263689: gap of unknown length
*  263690   265026: contig of 1337 bp in length
*  265027   265126: gap of unknown length
*  265127   267109: contig of 1983 bp in length
*  267110   267209: gap of unknown length
*  267210   269042: contig of 1833 bp in length

Alignment Scores:
Pred. No.:              7.08e+03        Length:         286181
Score:                  68.50           Matches:        14
Percent Similarity:     67.9%           Conservative:   5
Best Local Similarity:  50.0%           Mismatches:     8
Query Match:            29.9%           Indels:         1
DB:                     14              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x AC156064 (1-286181)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
              |||||||||||||||   ||||||               |||   |||::::::   :::
Db      61884 CCGCCGGCCGGCCCGCACGCCTCACGCCTCGCGGCAGGCGAGACAGAGGAAGAAAGGGAG 61825

Qy         32 PheValSerGluGlyProSerArg 39
              |||:::   :::||||||||||||
Db      61824 TTCTTA---AAGGGACCGTCACGC 61804

RESULT 43
BC002875
LOCUS       BC002875                2311 bp    mRNA    linear   PRI 12-NOV-2003
DEFINITION  Homo sapiens proline-, glutamic acid-, leucine-rich protein 1, mRNA
            (cDNA clone IMAGE:3940843), partial cds.
ACCESSION   BC002875
VERSION     BC002875.2  GI:38197220
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2311)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
```

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 2311)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT On Nov 6, 2003 this sequence version replaced gi:12804044.
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 15 Row: e Column: 5.

FEATURES Location/Qualifiers

source 1. .2311
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3940843"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH_MGC_7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

gene <1. .2311
/gene="PELP1"
/note="synonyms: P160, MNAR"
/db_xref="GeneID:27043"

CDS <1. .2234
/gene="PELP1"
/codon_start=3
/product="PELP1 protein"
/protein_id="AAH02875.2"
/db_xref="GI:38197221"
/db_xref="GeneID:27043"
/translation="GSRLLRFGILIGRLLPQVLNSWSIGRDSLSPGQERPYSTVRTKV
YAILELWVQVCGASAGMLQGGASGEALLTHLLSDISPPADALKLRSPRGSPDGSLQTG
KPSAPKKLKLDVGEAMAPPSHRKGDSNANSDVCAAALRGLSRTILMCGPLIKEETHRR
LHDLVLPLVMGVQQGEVLGSSPYTSSRCRRELYCLLLALLLAPSPRCPPPLACALQAF
SLGQREDSLEVSSFCSEALVTCAALTHPRVPPLQPMGPTCPTPAPVPPPEAPSPFRAP
PFHPPGPMPSVGSMPSAGPMPSAGPMPSAGPVPSARPGPPTTANHLGLSVPGLVSVPP
RLLPGPENHRAGSNEDPILAPSGTPPPTIPPDETFGGRVPRPAFVHYDKEEASDVEIS
LESDSDDSVVIVPEGLPPLPPPPPSGATPPPIAPTGPPTASPPVPAKEEPEELPAAPG
PLPPPPPPPPPVPGPVTLPPPQLVPEGTPGGGGPPALEEDLTVININSSDEEEEEEEE
EEEEEEEEEEEEDFEEEEDEEEYFEEEEEEEEFEEEFEEEEGELEEEEEEEDEEE
EEELEEVEDLEFGTAGGEVEEGAPPPPTLPPALPPPESPPKVQPEPEPEPGLLLEVEE
PGTEEERGADTAPTLAPEALPSQGEVEREGESPAAGPPPQELVEEEPSAPPTLLEEET
EDGSDKVQPPPETPAEEEMETETEAEALQEKEQDDTAAMLADFIDCPPDDEKPPPPTE
PDS"

ORIGIN

Alignment Scores:

Pred. No.:	55.6	Length:	2311
Score:	68.00	Matches:	14
Percent Similarity:	48.8%	Conservative:	7
Best Local Similarity:	32.6%	Mismatches:	22
Query Match:	29.7%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BC002875 (1-2311)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||::::::|||      |||   |||   |||          |||
Db       1980 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 2039

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                |||::::::      |||   :::|||      |||    |||:::
Db       2040 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 2099

Qy         41 ValLeuGlu 43
                :::|||
Db       2100 GAGATGGAG 2108
```

RESULT 44
CQ725289
LOCUS CQ725289 2715 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 11223 from Patent WO02068579.
ACCESSION CQ725289
VERSION CQ725289.1 GI:42286146
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 11223 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES Location/Qualifiers

source 1. .2715
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	65.7	Length:	2715
Score:	68.00	Matches:	14
Percent Similarity:	48.8%	Conservative:	7
Best Local Similarity:	32.6%	Mismatches:	22
Query Match:	29.7%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x CQ725289 (1-2715)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||  |||        |||
Db       2461 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 2520

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||:::::      |||   :::|||         |||   |||:::
Db       2521 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 2580

Qy         41 ValLeuGlu 43
                :::|||
Db       2581 GAGATGGAG 2589

RESULT 45
AY882602
LOCUS       AY882602               3186 bp    mRNA    linear   PRI 09-FEB-2005
DEFINITION  Homo sapiens transcription factor HMX3 mRNA, complete cds.
ACCESSION   AY882602
VERSION     AY882602.1  GI:58618183
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3186)
  AUTHORS   Lei,W. and Harrod,K.S.
  TITLE     Human transcription factor HMX3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3186)
  AUTHORS   Lei,W. and Harrod,K.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JAN-2005) Infectious Disease Program, Lovelace
            Respiratory Research Institute, 2425 Ridgecrest Dr. SE,
            Albuquerque, NM 87108, USA
FEATURES             Location/Qualifiers
     source          1. .3186
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             1. .3186
                     /codon_start=1
                     /product="transcription factor HMX3"
                     /protein_id="AAW80659.1"
                     /db_xref="GI:58618184"
                     /translation="MCLLRLHGSVGGAQNLSALGALVSLSNARLSSIKTRPEGLCLLS
                     LLVGESPTELFQQHCVSWLRSIQQVLQTQDPPATMELAVAVLRDLLRYAAQLPALFRD
                     ISMNHLPGLLTSLLGLRPECEQSALEGMKACMTYFPRACGSLKGKLASFFLSRVDALS
                     PQLQQLACECYSRLPSLGAGFSQGLKHTESWEQELHSLLASLHTLLGALYEGAETAPV
                     QNEGPGVEMLLSSEDGDAHVLLQLRQRFSGLARCLGLMLSSEFGAPVSVPVQEILDFI
                     CRTLSVSSKNISLHGDGPLRLLLLPSIHLEALDLLSALILACGSRLLRFGILIGRLLP
                     QVLNSWSIGRDSLSPGQERPYSTVRTKVYAILELWVQVCGASAGMLQGGASGEALLTH
                     LLSDISPPADALKLRSPRGSPDGSLQTGKPSAPKKLKLDVGEAMAPPSHRKGDSNANS
                     DVCAAALRGLSRTILMCGPLIKEETHRRLHDLVLPLVMGVQQGEVLGSSPYTSSRCRR
                     ELYCLLLALLLAPSPRCPPPLACALQAFSLGQREDSLEVSSFCSEALVTCAALTHPRV
                     PPLQPMGPTCPTPAPVPPPEAPSPFRAPPFHPPGPMPSVGSMPSAGPMPSAGPMPSAG
                     PVPSARPGPPTTANHLGLSVPGLVSVPPRLLPGPENHRAGSNEDPILAPSGTPPPTIP
                     PDETFGGRVPRPAFVHYDKEEASDVEISLESDSDDSVVIVPEGLPPLPPPPPSGATPP
                     PIAPTGPPTASPPVPAKEEPEELPAAPGPLPPPPPPPPPVPGPVTLPPPQLVPEGTPG
                     GGGPPALEEDLTVININSSDEEEEEEEEEEEEEEEEEEEDFEEEEDEEEYFEEEE
                     EEEEEFEEEFEEEEGELEEEEEEEDEEEEEELEEVEDLEFGTAGGEVEEGAPPPPTLP
                     PALPPPESPPKVQPEPEPEPGLLLEVEEPGTEEERGADTAPTLAPEALPSQGEVEREG
                     ESPAAGPPPQELVEEEPSAPPTLLEEETEDGSDKVQPPPETPAEEEMETETEAEALQE
                     KEQDDTAAMLADFIDCPPDDEKPPPPTEPDS"
ORIGIN

Alignment Scores:
Pred. No.:                77.6          Length:          3186
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
DB:                       8             Gaps:            0
```

```
US-10-757-745-6_COPY_2_45 (1-44) x AY882602 (1-3186)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||  |||        |||
Db       2932 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 2991

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||:::::      |||   :::|||         |||   |||:::
Db       2992 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3051

Qy         41 ValLeuGlu 43
                :::|||
Db       3052 GAGATGGAG 3060

RESULT 46
BC010457
LOCUS       BC010457               3216 bp    mRNA    linear   PRI 12-NOV-2003
DEFINITION  Homo sapiens proline-, glutamic acid-, leucine-rich protein 1, mRNA
            (cDNA clone IMAGE:4156419), partial cds.
ACCESSION   BC010457
VERSION     BC010457.2  GI:34783133
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3216)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3216)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Sep 16, 2003 this sequence version replaced gi:14714635.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.
```

```
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 12 Row: d Column: 1
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Hexamer frequency ORF
          analysis, Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1. .3216
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4156419"
                     /tissue_type="Brain, anaplastic oligodendroglioma with
                     1p/19q loss"
                     /clone_lib="NCI_CGAP_Brn67"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            <1. .3216
                     /gene="PELP1"
                     /note="synonyms: P160, MNAR"
                     /db_xref="GeneID:27043"
     CDS             <1. .3147
                     /gene="PELP1"
                     /codon_start=1
                     /product="PELP1 protein"
                     /protein_id="AAH10457.2"
                     /db_xref="GI:34783134"
                     /db_xref="GeneID:27043"
                     /translation="QNLSALGALVSLSNARLSSIKTRFEGLCLLSLLVGESPTELFQQ
                     HCVSWLRSIQQVLQTQDPPATMELAVAVLRDLLRYAAQLPALFRDISMNHLPGLLTSL
                     LGLRPECEQSALEGMKACMTYFPRACGSLKGKLASFFLSRVDALSPQLQQLACECYSR
                     LPSLGAGFSQGLKHTESWEQELHSLLASLHTLLGALYEGAETAPVQNEGPGVEMLLSS
                     EDGDAHVLLQLRQRFSGLARCLGLMLSSEFGAPVSVPVQEILDFICRTLSVSSKNISL
                     HGDGPLRLLLLPSIHLEALDLLSALILACGSRLLRFGILIGRLLPQVLNSWSIGRDSL
                     SPGQERPYSTVRTKVYAILELWVQVCGASAGMLQGGASGEALLTHLLSDISPPADALK
                     LRSPRGSPDGSLQTGKPSAPKKLKLDVGEAMAPPSHRKGDSNANSDVCAAALRGLSRT
                     ILMCGPLIKEETHRRLHDLVLPLVMGVQQGEVLGSSPYTSSRCRRELYCLLLALLLAP
                     SPRCPPPLACALQAFSLGQREDSLEVSSFCSEALVTCAALTHPRVPPLQPMGPTCPTP
                     APVPPPEAPSPFRAPPFHPPGPMPSVGSMPSAGPMPSAGPMPSAGPVPSARPGPPTTA
                     NHLGLSVPGLVSVPPRLLPGPENHRAGSNEDPILAPSGTPPPTIPPDETFGGRVPRPA
                     FVHYDKEEASDVEISLESDSDDSVVIVPEGLPPLPPPPPSGATPPPIAPTGPPTASPP
                     VPAKEEPEELPAAPGPLPPPPPPPPPVPGPVTLPPPQLVPEGTPGGGGPPALEEDLTV
                     ININSSDEEEEEEEEEEEEEEEEEEEEDFEEEEDEEEYFEEEEEEEEFEEEFEEE
                     EGELEEEEEEEDEEEEEELEEVEDLEFGTAGGEVEEGAPPPPTLPPALPPPESPPKVQ
                     PEPEPEPGLLLEVEEPGTEEERGADTAPTLAPEALPSQGEVEREGESPAAGPPPQELV
                     EEEPSAPPTLLEEETEDGSDKVQPPPETPAEEEMETETEAEALQEKEQDDTAAMLADF
                     IDCPPDDEKPPPPTEPDS"
ORIGIN

Alignment Scores:
Pred. No.:                78.3          Length:          3216
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
DB:                       8             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BC010457 (1-3216)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||:::::::|||      |||   |||   |||         |||
Db       2893 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 2952

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::     |||   :::|||          |||   |||:::
Db       2953 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3012

Qy         41 ValLeuGlu 43
                :::|||
Db       3013 GAGATGGAG 3021

RESULT 47
BC069058
```

```
LOCUS       BC069058                3456 bp    mRNA    linear   PRI 30-JUN-2004
DEFINITION  Homo sapiens proline-, glutamic acid-, leucine-rich protein 1, mRNA
            (cDNA clone MGC:70420 IMAGE:6068755), complete cds.
ACCESSION   BC069058
VERSION     BC069058.1  GI:46399216
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3456)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3456)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 140 Row: g Column: 6
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 24415382.
FEATURES             Location/Qualifiers
     source          1. .3456
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:70420 IMAGE:6068755"
                     /tissue_type="Testis, embryonal carcinoma"
```

```
                     /clone_lib="NIH_MGC_92"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .3456
                     /gene="PELP1"
                     /note="synonyms: P160, MNAR"
                     /db_xref="GeneID:27043"
     CDS             1. .3393
                     /gene="PELP1"
                     /codon_start=1
                     /product="proline-, glutamic acid-, leucine-rich protein
                     1"
                     /protein_id="AAH69058.1"
                     /db_xref="GI:46399217"
                     /db_xref="GeneID:27043"
                     /translation="MAAAVLSGPSAGSAAGVPGGTGGLSAVSSGPRLRLLLLESVSGL
                     LQPRTGSAVAPVHPPNRSAPHLPGLMCLLRLHGSVGGAQNLSALGALVSLSNARLSSI
                     KTRFEGLCLLSLLVGESPTELFQQHCVSWLRSIQQVLQTQDPPATMELAVAVLRDLLR
                     YAAQLPALFRDISMNHLPGLLTSLLGLRPECEQSALEGMKACMTYFPRACGSLKGKLA
                     SFFLSRVDALSPQLQQLACECYSRLPSLGAGFSQGLKHTESWEQELHSLLASLHTLLG
                     ALYEGAETAPVQNEGPGVEMLLSSEDGDAHVLLQLRQRFSGLARCLGLMLSSEFGAPV
                     SVPVQEILDFICRTLSVSSKNISLHGDGPLRLLLLPSIHLEALDLLSALILACGSRLL
                     RFGILIGRLLPQVLNSWSIGRDSLSPGQERPYSTVRTKVYAILELWVQVCGASAGMLQ
                     GGASGEALLTHLLSDISPPADALKLRSPRGSPDGSLQTGKPSAPKKLKLDVGEAMAPP
                     SHRKGDSNANSDVCAAALRGLSRTILMCGPLIKEETHRRLHDLVLPLVMGVQQGEVLG
                     SSPYTSSRCRRELYCLLLALLLAPSPRCPPPLACALQAFSLGQREDSLEVSSFCSEAL
                     VTCAALTHPRVPPLQPMGPTCPTPAPVPPPEAPSPFRAPPFHPPGPMPSVGSMPSAGP
                     MPSAGPMPSAGPVPSARPGPPTTANHLGLSVPGLVSVPPRLLPGPENHRAGSNEDPIL
                     APSGTPPPTIPPDETFGGRVPRPAFVHYDKEEASDVEISLESDSDDSVVIVPEGLPPL
                     PPPPPSGATPPPIAPTGPPTASPPVPAKEEPEELPAAPGPLPPPPPPPPPVPGPVTLP
                     PPQLVPEGTPGGGGPPALEEDLTVININSSDEEEEEEEEEEEEEEEEEEEEDFEEEE
                     EDEEEYFEEEEEEEEEFEEEFEEEEGELEEEEEEEDEEEEEELEEVEDLEFGTAGGEV
                     EEGAPPPPTLPPALPPPESPPKVQPEPEPEPGLLLEVEEPGTEEERGADTAPTLAPEA
                     LPSQGEVEREGESPAAGPPPQELVEEEPSAPPTLLEEETEDGSDKVQPPPETPAEEEM
                     ETETEAEALQEKEQDDTAAMLADFIDCPPDDEKPPPPTEPDS"
ORIGIN

Alignment Scores:
Pred. No.:                84.4          Length:         3456
Score:                    68.00         Matches:        14
Percent Similarity:       48.8%         Conservative:   7
Best Local Similarity:    32.6%         Mismatches:     22
Query Match:              29.7%         Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BC069058 (1-3456)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::::|||      |||   |||   |||          |||
Db       3139 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 3198

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::      |||   :::|||         |||   |||:::
Db       3199 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3258

Qy         41 ValLeuGlu 43
                :::|||
Db       3259 GAGATGGAG 3267

RESULT 48
HSU88153
LOCUS       HSU88153                3910 bp    mRNA    linear   PRI 14-JUN-2002
DEFINITION  Homo sapiens PELP1 mRNA, complete cds.
ACCESSION   U88153
VERSION     U88153.2  GI:21426921
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 3910)
  AUTHORS   Vadlamudi,R.K., Wang,R.A., Mazumdar,A., Kim,Y., Shin,J., Sahin,A.
            and Kumar,R.
  TITLE     Molecular cloning and characterization of PELP1, a novel human
            coregulator of estrogen receptor alpha
  JOURNAL   J. Biol. Chem. 276 (41), 38272-38279 (2001)
  PUBMED    11481323
REFERENCE   2  (bases 1 to 3910)
  AUTHORS   Vadlamudi,R.K., Chung,J.K. and Shin,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-1997) Tumor Virology, Dana Farber Cancer
            Institute, 44 Binney St., Norwood, MA 02115, USA
REFERENCE   3  (bases 1 to 3910)
  AUTHORS   Vadlamudi,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-1998) MD Anderson Cancer Center, 1515 Holcomb
            Blvd., Houston, TX 77030, USA
  REMARK    CDS extended by submitter
REFERENCE   4  (bases 1 to 3910)
  AUTHORS   Vadlamudi,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-2002) MD Anderson Cancer Center, 1515 Holcomb
            Blvd., Houston, TX 77030, USA
  REMARK    Sequence update by submitter
COMMENT     On Jun 14, 2002 this sequence version replaced gi:2444068.
FEATURES             Location/Qualifiers
     source          1. .3910
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_line="Hela"
     CDS             1. .3855
                     /codon_start=1
                     /product="PELP1"
                     /protein_id="AAC17708.2"
                     /db_xref="GI:21426922"
                     /translation="MVAAVLSGPSAGSAAGVPGGTGGLSAVSSGPRLRLLLLESVSGL
                     LQPRTGSAVAPVHPPNRSAPHLPGLMCLLRLHGSVGGAQNLSALGALVSLSNARLSSI
                     KTRFEGLCLLSLLVGESPTELFQQHCVSWLRSIQQVLQTQDPPATMELAVAVLRDLLR
                     YAAQLPALFRDISMNHLPGLLTSLLGLRPECEQSALEGMKACMTYFPRACGSLKGKLA
                     SFFLSRVDALSPQLQQLACECYSRLPSLGAGFSQGLKHTESWEQELHSLLASLHTLLG
                     ALYEGAETAPVQNEGPGVEMLLSSEDGDAHVLLQLRQRFSGLARCLGLMLSSEFGAPV
                     SVPVQEILDFICRTLSVSSKNIVSGICHLFRALAQDTRQPGKYWGPESPQTVSSWSPS
                     QRASTFVQITSLPMCRDTGAQCQSVANASLGEGEFGDSAESLLRGPAILLTFHPGSIL
                     EDRGLILLGEMRSGVGFLTYVYICKWSFPVSVSLWLSLSSSTLYLCPFFLQSLHGDGP
                     LRLLLLPSIHLEALDLLSALILACGSRLLRFGILIGRLLPQVLNSWSIGRDSLSPGQE
                     RPYSTVRTKVYAILELWVQVCGASAGMLQGGASGEALLTHLLSDISPPADALKLRSPR
                     GSPDGSLQTGKPSAPKKLKLDVGEAMAPPSHRKGDSNANSDVCAAALRGLSRTILMCG
                     PLIKEETHRRLHDLVLPLVMGVQQGEVLGSSPYTSSRCRRELYCLLLALLLAPSPRCP
                     PPLACALQAFSLGQREDSLEVSSFCSEALVTCAALTHPRVPPLQPMGPTCPTPAPVPP
                     PEPHRPSGPHRSILRAPCPSVGSMPSAGPMPSAGPMPSAGPVPSDSPGPPTTANLLGL
                     SVPGLVSVPPRLLPGPENHRAGSNEDPILAPSGTPPPTIPPDETFGGRVPRPAFVHYD
                     KEEASDVEISLESDSDDSVVIVPEGLPPLPPPPPSGATPPPIAPTGPPTASPPVPAKE
                     EPEELPAAPGPLPPPPPPPPVPGPVTLPPPQLVPEGTPGGGGPPALEEDLTVININS
                     SDEEEEEEGEEEEEEEEEEEEEEEEEEEEEEEDFEEEEDEEEYFEEEEEEEEFEEE
                     FEEEEGELEEEEEEDEEEEELEEVEDLEFGTAGGEVEEGAPPPPTLPPALPPPESP
                     PKVQPEPEPEPGLLLEVEEPGTEEERGADTAPTLAPEALPSQGEVEREGESPAAGPPP
                     QELVEEEPSAPPTLLEEETEDGSDKVQPPPETPAEEEMETETEAEALQEKEQDDTAAM
                     LADFIDCPPDDEKPPPPTEPDS"
ORIGIN

Alignment Scores:
Pred. No.:                95.9          Length:         3910
Score:                    68.00         Matches:        14
Percent Similarity:       48.8%         Conservative:   7
Best Local Similarity:    32.6%         Mismatches:     22
Query Match:              29.7%         Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x HSU88153 (1-3910)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::::|||      |||   |||   |||          |||
Db       3601 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 3660
```

```
Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||::::::      |||   :::|||          |||   |||:::
Db       3661 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3720

Qy         41 ValLeuGlu 43
                :::|||
Db       3721 GAGATGGAG 3729

RESULT 49
AC002490/c
LOCUS       AC002490               86945 bp    DNA     linear   HTG 08-SEP-1997
DEFINITION  Homo sapiens chromosome 17 clone CITI180A21, *** SEQUENCING IN
            PROGRESS ***, 11 unordered pieces.
ACCESSION   AC002490
VERSION     AC002490.1  GI:2341009
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 86945)
  AUTHORS   Hawkins,T.L., Birren,B.W., Fasman,K.H., Nussbaum,C. and Lander,E.S.
  TITLE     Genomic sequence from Human 17
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 86945)
  AUTHORS   Hawkins,T.L., Birren,B.W., Fasman,K.H., Nussbaum,C., Lander,E.S.,
            McKernan,K., Munro,C., Richardson,P., Barna,N., Chang,A., Cooke,P.,
            Daly,M.J., Devon,K., Dewar,K., Forrest,C., Gage,D., Geraigery,K.,
            Hagos,B., Huang,J., Hui,L., Jacotot,L., Kirby,A., Lane,M.,
            MacKenzie,J., Marquis,N., McDermott,J., Molla,M., Morrow,J.,
            Nachman,A., Naylor,J., Nusbaum,C., O'Connor,T., Olotu,A.,
            Peterson,K., Reeve,M.P., Roberts,D., Rollins,G., Stilwell,J.,
            Stone,C., Strickland,C., Sydney,K., Tang,L., Wilmer,F., Zemtseva,I.
            and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1997) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     The Staden databases, finishing information, and all
            chromatographic files used in the assembly of this clone are
            available from our anonymous ftp site.

            All repeats were identified using RepeatMasker: Smit, A.F.A. &
            Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3993: contig of 3993 bp in length
            *                   gap of unknown length
            *     3994     6023: contig of 2030 bp in length
            *                   gap of unknown length
            *     6024    21986: contig of 15963 bp in length
            *                   gap of unknown length
            *    21987    26296: contig of 4310 bp in length
            *                   gap of unknown length
            *    26297    30564: contig of 4268 bp in length
            *                   gap of unknown length
            *    30565    42164: contig of 11600 bp in length
            *                   gap of unknown length
            *    42165    51729: contig of 9565 bp in length
            *                   gap of unknown length
            *    51730    78333: contig of 26604 bp in length
            *                   gap of unknown length
            *    78334    81497: contig of 3164 bp in length
            *                   gap of unknown length
            *    81498    83990: contig of 2493 bp in length
            *                   gap of unknown length
            *    83991    86945: contig of 2955 bp in length.
FEATURES            Location/Qualifiers
     source         1. .86945
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="17"
                    /clone="CITI180A21"
                    /clone_lib="unknown"
ORIGIN

Alignment Scores:
Pred. No.:                2.4e+03        Length:          86945
Score:                    68.00          Matches:         14
Percent Similarity:       48.8%          Conservative:    7
Best Local Similarity:    32.6%          Mismatches:      22
Query Match:              29.7%          Indels:          0
DB:                       14             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AC002490 (1-86945)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||         |||
Db      78798 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 78739

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||::::::      |||   :::|||          |||   |||:::
Db      78738 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 78679

Qy         41 ValLeuGlu 43
                :::|||
Db      78678 GAGATGGAG 78670

RESULT 50
AC027820
LOCUS       AC027820               154353 bp    DNA     linear   PRI 07-SEP-2002
DEFINITION  Homo sapiens chromosome 17, clone RP11-198F11, complete sequence.
ACCESSION   AC027820
VERSION     AC027820.12  GI:22758431
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 154353)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-198F11
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 154353)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
```

```
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 154353)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 154353)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 7, 2002 this sequence version replaced gi:22417407.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L9155
                Center clone name: 198_F_11
            --------------.
FEATURES                Location/Qualifiers
     source             1. .154353
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="17"
                        /map="17"
                        /clone="RP11-198F11"
                        /clone_lib="RPCI-11 Human Male BAC"
     repeat_region      1. .238
                        /rpt_family="AluJb"
     repeat_region      321. .349
                        /rpt_family="AT_rich"
     repeat_region      420. .725
                        /rpt_family="AluY"
     repeat_region      734. .745
                        /rpt_family="(T)n"
     repeat_region      746. .1049
                        /rpt_family="AluY"
     repeat_region      1050. .1064
                        /rpt_family="(T)n"
     repeat_region      1065. .1365
                        /rpt_family="AluSq"
     repeat_region      1366. .1377
                        /rpt_family="(T)n"
     repeat_region      1378. .1669
                        /rpt_family="AluSp"
     repeat_region      1670. .1691
                        /rpt_family="(T)n"
     repeat_region      complement(1729. .1905)
                        /rpt_family="AluSq"
     repeat_region      complement(3649. .3944)
                        /rpt_family="AluY"
     repeat_region      complement(3945. .4240)
                        /rpt_family="AluSc"
     repeat_region      complement(5223. .5521)
                        /rpt_family="AluSx"
     repeat_region      6322. .6557
                        /rpt_family="(TCCA)n"
     repeat_region      7437. .7756
                        /rpt_family="L1MC/D"
     repeat_region      complement(7757. .8043)
                        /rpt_family="AluSx"
     repeat_region      8044. .8200
                        /rpt_family="L1MC/D"
     repeat_region      8312. .8595
                        /rpt_family="AluSx"
     repeat_region      8596. .8624
                        /rpt_family="(CAAAC)n"
     repeat_region      8625. .8636
                        /rpt_family="AluSx"
     repeat_region      complement(8830. .9028)
                        /rpt_family="L1ME2"
     repeat_region      9029. .9337
                        /rpt_family="AluSc"
     repeat_region      9338. .9499
                        /rpt_family="AluSq"
     repeat_region      9507. .9806
                        /rpt_family="AluJb"
     repeat_region      9851. .10043
                        /rpt_family="L2"
     repeat_region      10044. .10332
                        /rpt_family="AluSq"
     repeat_region      10333. .10351
                        /rpt_family="L2"
     repeat_region      complement(10956. .11258)
                        /rpt_family="AluSq"
     repeat_region      12628. .12921
                        /rpt_family="AluJo"
     repeat_region      14060. .14082
                        /rpt_family="GC_rich"
     repeat_region      14469. .14546
                        /rpt_family="GC_rich"
     repeat_region      14577. .14611
                        /rpt_family="GC_rich"
     repeat_region      14794. .14846
                        /rpt_family="GC_rich"
     repeat_region      complement(15012. .15131)
                        /rpt_family="L2"
     repeat_region      complement(15184. .15615)
                        /rpt_family="L1ME2"
     repeat_region      complement(15616. .15649)
                        /rpt_family="AluSq"
     repeat_region      complement(15650. .15951)
                        /rpt_family="AluY"
     repeat_region      complement(15952. .16178)
                        /rpt_family="AluSq"
     repeat_region      complement(16188. .16472)
```

```
                     /rpt_family="AluSp"
     repeat_region   16474. .16714
                     /rpt_family="L1ME2"
     repeat_region   16738. .17025
                     /rpt_family="AluSc"
     repeat_region   17118. .17428
                     /rpt_family="AluY"
     repeat_region   17600. .17895
                     /rpt_family="AluSx"
     repeat_region   17896. .18204
                     /rpt_family="AluY"
     repeat_region   18247. .18402
                     /rpt_family="MIR"
     repeat_region   18403. .18697
                     /rpt_family="AluSx"
     repeat_region   18698. .18774

Alignment Scores:
Pred. No.:                4.34e+03        Length:          154353
Score:                    68.00           Matches:         14
Percent Similarity:       48.8%           Conservative:    7
Best Local Similarity:    32.6%           Mismatches:      22
Query Match:              29.7%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AC027820 (1-154353)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               |||   |||       |||::::::|||       |||   |||   |||         |||
Db      149774 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 149833

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::      |||   :::|||           |||   |||:::
Db      149834 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 149893

Qy          41 ValLeuGlu 43
                 :::|||
Db      149894 GAGATGGAG 149902

RESULT 51
AC091153
LOCUS       AC091153               195494 bp    DNA     linear   PRI 01-OCT-2002
DEFINITION  Homo sapiens chromosome 17, clone RP11-314A20, complete sequence.
ACCESSION   AC091153
VERSION     AC091153.10  GI:23396331
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 195494)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-314A20
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 195494)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
            Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K.,
            Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
            Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-APR-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 195494)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 195494)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Oct 1, 2002 this sequence version replaced gi:22779581.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L13227
                Center clone name: 314_A_20
            --------------.
FEATURES                Location/Qualifiers
     source             1. .195494
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="17"
                        /map="17"
                        /clone="RP11-314A20"
                        /clone_lib="RPCI-11 Human Male BAC"
```

```
repeat_region   220. .532
                /rpt_family="AluSx"
repeat_region   complement(543. .1136)
                /rpt_family="AluSq"
repeat_region   1555. .1644
                /rpt_family="MIR"
repeat_region   1680. .1980
                /rpt_family="AluSp"
repeat_region   complement(2582. .2878)
                /rpt_family="AluJo"
repeat_region   complement(2960. .3258)
                /rpt_family="AluSc"
repeat_region   complement(3418. .3701)
                /rpt_family="AluJb"
repeat_region   complement(3974. .4136)
                /rpt_family="AluJo"
repeat_region   complement(4137. .4431)
                /rpt_family="AluSp"
repeat_region   complement(4432. .4577)
                /rpt_family="AluJo"
repeat_region   4711. .4811
                /rpt_family="L2"
repeat_region   complement(4812. .5116)
                /rpt_family="AluSg"
repeat_region   5117. .5192
                /rpt_family="L2"
repeat_region   5206. .5432
                /rpt_family="L1ME3A"
repeat_region   5433. .5746
                /rpt_family="AluSx"
repeat_region   5747. .5885
                /rpt_family="L1ME3A"
repeat_region   5886. .6172
                /rpt_family="AluSx"
repeat_region   6173. .6426
                /rpt_family="L1ME3A"
repeat_region   6476. .6771
                /rpt_family="AluY"
repeat_region   6774. .7063
                /rpt_family="L2"
repeat_region   complement(7099. .7408)
                /rpt_family="AluSp"
repeat_region   complement(7467. .7574)
                /rpt_family="FRAM"
repeat_region   complement(7597. .7703)
                /rpt_family="L2"
repeat_region   complement(7799. .8096)
                /rpt_family="AluSx"
repeat_region   8702. .8830
                /rpt_family="AluJb"
repeat_region   8831. .9141
                /rpt_family="AluSx"
repeat_region   9142. .9326
                /rpt_family="AluJb"
repeat_region   10478. .10784
                /rpt_family="AluSq"
repeat_region   10812. .10853
                /rpt_family="(TAAA)n"
repeat_region   complement(11279. .11561)
                /rpt_family="AluSx"
repeat_region   complement(11577. .11808)
                /rpt_family="AluJo"
repeat_region   complement(11809. .11855)
                /rpt_family="AluSg/x"
repeat_region   complement(12298. .12606)
                /rpt_family="AluSp"
repeat_region   complement(12641. .12806)
                /rpt_family="AluJb"
repeat_region   12821. .13131
                /rpt_family="AluY"
repeat_region   complement(13851. .14137)
                /rpt_family="AluSq"
unsure          14448. .14470
                /note="<30 qual SNGL region"
repeat_region   14459. .14760
                /rpt_family="AluSx"
repeat_region   14764. .15062
                /rpt_family="AluSx"
repeat_region   complement(15631. .15929)
                /rpt_family="AluSx"
repeat_region   complement(15952. .16126)
                /rpt_family="AluJb"
repeat_region   16127. .16153
                /rpt_family="(TA)n"
repeat_region   complement(16154. .16442)
                /rpt_family="AluSp"
repeat_region   complement(16443. .16577)
                /rpt_family="AluJb"
repeat_region   16600. .16902
                /rpt_family="AluSp"
repeat_region   complement(17010. .17419)
                /rpt_family="MLT1K"
repeat_region   complement(17420. .17585)
                /rpt_family="THE1C"
repeat_region   17886. .18020

Alignment Scores:
Pred. No.:              5.55e+03        Length:         195494
Score:                  68.00           Matches:        14
Percent Similarity:     48.8%           Conservative:   7
Best Local Similarity:  32.6%           Mismatches:     22
Query Match:            29.7%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AC091153 (1-195494)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||  |||      |||::::::|||      |||  |||  |||         |||
Db     110182 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 110241

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                |||::::::      |||   :::|||         |||   |||:::
Db     110242 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 110301

Qy         41 ValLeuGlu 43
                :::|||
Db     110302 GAGATGGAG 110310

RESULT 52
AC118439/c
LOCUS       AC118439              227406 bp    DNA     linear   HTG 15-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-292B6, WORKING DRAFT SEQUENCE.
ACCESSION   AC118439
VERSION     AC118439.5  GI:25009636
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 227406)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
```

```
          Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
          Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
          Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
          Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
          Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
          Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
          Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
          Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
          Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
          Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
          Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
          Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
          Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
          Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
          Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
          Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
          Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
          Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
          Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
          Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
          Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
          Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
          Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
          Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
          Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
          Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
          Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
          Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
          Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
          Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
          Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
          Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
          Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
          Weinstock,G. and Gibbs,R.A.
  TITLE   Direct Submission
  JOURNAL Unpublished
REFERENCE 2  (bases 1 to 227406)
  AUTHORS Worley,K.C.
  TITLE   Direct Submission
  JOURNAL Submitted (17-APR-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3  (bases 1 to 227406)
  AUTHORS Rat Genome Sequencing Consortium.
  TITLE   Direct Submission
  JOURNAL Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Nov 15, 2002 this sequence version replaced gi:23810271.
          The sequence in this assembly is a combination of BAC based reads
          and whole genome shotgun sequencing reads assembled using Atlas
          (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
          in the feature table below represents a scaffold in the Atlas
          assembly (a 'contig-scaffold'). Within each contig-scaffold,
          individual sequence contigs are ordered and oriented, and separated
          by sized gaps filled with Ns to the estimated size. The sequence
          may extend beyond the ends of the clone and there may be sequence
          contigs within a contig-scaffold that consist entirely of whole
          genome shotgun sequence reads. Both end sequences and whole genome
          shotgun sequence only contigs will be indicated in the feature
          table.
          -------------- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
          -------------- Project Information
              Center project name: GUVV
              Center clone name: CH230-292B6
          -------------- Summary Statistics
              Assembly program: Phrap; version 0.990329
              Consensus quality: 220256 bases at least Q40
              Consensus quality: 221805 bases at least Q30
              Consensus quality: 222900 bases at least Q20
              Estimated insert size: 228346; sum-of-contigs estimation
              Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
          --------------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 1 contigs. Gaps between the contigs
          * are represented as runs of N. The order of the pieces
          * is believed to be correct as given, however the sizes
          * of the gaps between them are based on estimates that have
          * provided by the submittor.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          *       1   227406: contig of 227406 bp in length.
FEATURES             Location/Qualifiers
     source          1. .227406
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-292B6"
     misc_feature    3. .897
                     /note="clone_boundary
                     clone_end:T7
                     site:
                     end_sequence:BZ190731"
     misc_feature    complement(225981. .226797)
                     /note="clone_boundary
                     clone_end:Sp6
                     site:
                     end_sequence:BZ190742"
ORIGIN

Alignment Scores:
Pred. No.:                6.49e+03        Length:          227406
Score:                    68.00           Matches:         18
Percent Similarity:       50.0%           Conservative:    3
Best Local Similarity:    42.9%           Mismatches:      11
Query Match:              29.7%           Indels:          10
DB:                       14              Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x AC118439 (1-227406)

Qy          2 AspProGlySerGluIleIleGluSerValProPro---------AlaGlyProGluAla 18
              ||||||||||||:::::|||      :::|||||||          ||||||||||||
Db     103008 GACCCTGGAAGCCAATTGATTGTGCCCTTACCCCCCACTCCCCAAGGTGGACCGGAGGCT 102949

Qy         19 SerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSer 38
                                   |||                |||   |||      ||||||
Db     102948 --------------------AACCACTCCCAGCCCTTCTTCCATTCGCCGGATCCATCA 102910

Qy         39 ArgPro 40
              ||||||
Db     102909 CGACCG 102904

RESULT 53
AC150165
LOCUS       AC150165               235593 bp    DNA     linear   HTG 02-JUL-2004
DEFINITION  Gallus gallus clone CH261-149E23, WORKING DRAFT SEQUENCE, 21
            unordered pieces.
ACCESSION   AC150165
VERSION     AC150165.1  GI:49615147
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 235593)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Unpublished
  JOURNAL   Unpublished
```

```
REFERENCE   2  (bases 1 to 235593)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2004) Production Genomics Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
            94598-1698, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 2988112
            Center clone name: JF2-149E23
            -------
            Summary Statistics
            Consensus quality: 222914 bases at least Q40
            Consensus quality: 227963 bases at least Q30
            Consensus quality: 230980 bases at least Q20
            Estimated insert size: 283000; agarose-fp estimation
            Estimated insert size: 232851; sum-of-contigs estimation
            Quality coverage:  8.98 in Q20 bases; agarose-fp estimation
            Quality coverage:  10.92 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 21 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    37010: contig of 37010 bp in length
            *    37011    37110: gap of unknown length
            *    37111    59947: contig of 22837 bp in length
            *    59948    60047: gap of unknown length
            *    60048    66563: contig of 6516 bp in length
            *    66564    66663: gap of unknown length
            *    66664    68188: contig of 1525 bp in length
            *    68189    68288: gap of unknown length
            *    68289    69017: contig of 729 bp in length
            *    69018    69117: gap of unknown length
            *    69118    86157: contig of 17040 bp in length
            *    86158    86257: gap of unknown length
            *    86258   121365: contig of 35108 bp in length
            *   121366   121465: gap of unknown length
            *   121466   123318: contig of 1853 bp in length
            *   123319   123418: gap of unknown length
            *   123419   137562: contig of 14144 bp in length
            *   137563   137662: gap of unknown length
            *   137663   146776: contig of 9114 bp in length
            *   146777   146876: gap of unknown length
            *   146877   148268: contig of 1392 bp in length
            *   148269   148368: gap of unknown length
            *   148369   150052: contig of 1684 bp in length
            *   150053   150152: gap of unknown length
            *   150153   156661: contig of 6509 bp in length
            *   156662   156761: gap of unknown length
            *   156762   159783: contig of 3022 bp in length
            *   159784   159883: gap of unknown length
            *   159884   161692: contig of 1809 bp in length
            *   161693   161792: gap of unknown length
            *   161793   182912: contig of 21120 bp in length
            *   182913   183012: gap of unknown length
            *   183013   197519: contig of 14507 bp in length
            *   197520   197619: gap of unknown length
            *   197620   201302: contig of 3683 bp in length
            *   201303   201402: gap of unknown length
            *   201403   209553: contig of 8151 bp in length
            *   209554   209653: gap of unknown length
            *   209654   229970: contig of 20317 bp in length
            *   229971   230070: gap of unknown length
            *   230071   235593: contig of 5523 bp in length.
FEATURES             Location/Qualifiers
     source          1. .235593
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9031"
                     /clone="CH261-149E23"
                     /clone_lib="CHORI-261 Chicken BAC Library"
     gap             37011. .37110
                     /estimated_length=unknown
     gap             59948. .60047
                     /estimated_length=unknown
     gap             66564. .66663
                     /estimated_length=unknown
     gap             68189. .68288
                     /estimated_length=unknown
     gap             69018. .69117
                     /estimated_length=unknown
     gap             86158. .86257
                     /estimated_length=unknown
     gap             121366. .121465
                     /estimated_length=unknown
     gap             123319. .123418
                     /estimated_length=unknown
     gap             137563. .137662
                     /estimated_length=unknown
     gap             146777. .146876
                     /estimated_length=unknown
     gap             148269. .148368
                     /estimated_length=unknown
     gap             150053. .150152
                     /estimated_length=unknown
     gap             156662. .156761
                     /estimated_length=unknown
     gap             159784. .159883
                     /estimated_length=unknown
     gap             161693. .161792
                     /estimated_length=unknown
     gap             182913. .183012
                     /estimated_length=unknown
     gap             197520. .197619
                     /estimated_length=unknown
     gap             201303. .201402
                     /estimated_length=unknown
     gap             209554. .209653
                     /estimated_length=unknown
     gap             229971. .230070
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:              6.74e+03        Length:         235593
Score:                  68.00           Matches:        17
Percent Similarity:     48.8%           Conservative:   4
Best Local Similarity:  39.5%           Mismatches:     20
Query Match:            29.7%           Indels:         2
DB:                     14              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x AC150165 (1-235593)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu------ 20
              |||       |||:::         ||||||||||||   |||      |||:::
Db     154558 CCACCCCCCGAGGTGGGCCCCGAAGTGCCCCCCGCTGCCCCCGGCAAGTCCCAGGTTGGG 154617

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||| ||||||  |||  :::          |||||||||      |||
Db     154618 ATGACGCTGGACGAGGAGGAGGGAGAGTCGGACACGGAGCCAGAGGGGCCCGAGAGCCCT 154677

Qy         41 ValLeuGlu 43
              :::  |||
Db     154678 TTGGAGGAG 154686

RESULT 54
AC150044/c
LOCUS       AC150044               235990 bp    DNA     linear   HTG 01-JUL-2004
```

```
DEFINITION  Gallus gallus clone CH261-127L24, WORKING DRAFT SEQUENCE, 12
            ordered pieces.
ACCESSION   AC150044
VERSION     AC150044.1  GI:49533639
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 235990)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Unpublished
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 235990)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2004) Production Genomics Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
            94598-1698, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 2979832
            Center clone name: JF2-127L24
            -------
            Summary Statistics
            Consensus quality: 232263 bases at least Q40
            Consensus quality: 233529 bases at least Q30
            Consensus quality: 234333 bases at least Q20
            Estimated insert size: 254000; agarose-fp estimation
            Estimated insert size: 234890; sum-of-contigs estimation
            Quality coverage:  10.53 in Q20 bases; agarose-fp estimation
            Quality coverage:  11.39 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1     4675: contig of 4675 bp in length
            *    4676     4775: gap of unknown length
            *    4776    53511: contig of 48736 bp in length
            *   53512    53611: gap of unknown length
            *   53612   153067: contig of 99456 bp in length
            *  153068   153167: gap of unknown length
            *  153168   173870: contig of 20703 bp in length
            *  173871   173970: gap of unknown length
            *  173971   212196: contig of 38226 bp in length
            *  212197   212296: gap of unknown length
            *  212297   214687: contig of 2391 bp in length
            *  214688   214787: gap of unknown length
            *  214788   222610: contig of 7823 bp in length
            *  222611   222710: gap of unknown length
            *  222711   223940: contig of 1230 bp in length
            *  223941   224040: gap of unknown length
            *  224041   227626: contig of 3586 bp in length
            *  227627   227726: gap of unknown length
            *  227727   228734: contig of 1008 bp in length
            *  228735   228834: gap of unknown length
            *  228835   230316: contig of 1482 bp in length
            *  230317   230416: gap of unknown length
            *  230417   235990: contig of 5574 bp in length.
FEATURES             Location/Qualifiers
     source          1. .235990
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9031"
                     /clone="CH261-127L24"
                     /clone_lib="CHORI-261 Chicken BAC Library"
     gap             4676. .4775
                     /estimated_length=unknown
     gap             53512. .53611
                     /estimated_length=unknown
     gap             153068. .153167
                     /estimated_length=unknown
     gap             173871. .173970
                     /estimated_length=unknown
     gap             212197. .212296
                     /estimated_length=unknown
     gap             214688. .214787
                     /estimated_length=unknown
     gap             222611. .222710
                     /estimated_length=unknown
     gap             223941. .224040
                     /estimated_length=unknown
     gap             227627. .227726
                     /estimated_length=unknown
     gap             228735. .228834
                     /estimated_length=unknown
     gap             230317. .230416
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:                6.75e+03        Length:          235990
Score:                    68.00           Matches:         17
Percent Similarity:       48.8%           Conservative:    4
Best Local Similarity:    39.5%           Mismatches:      20
Query Match:              29.7%           Indels:          2
DB:                       14              Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x AC150044 (1-235990)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu------ 20
              |||       |||:::           ||||||||||||   |||     |||:::
Db     216486 CCACCCCCCGAGGTGGGCCCCGAAGTGCCCCCCGCTGCCCCCGGCAAGTCCCAGGTTGGG 216427

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                |||   ||||||   |||   :::                |||||||||       |||
Db     216426 ATGACGCTGGACGAGGAGGAGGGAGAGTCGGACACGGAGCCAGAGGGGCCCGAGAGCCCT 216367

Qy         41 ValLeuGlu 43
                :::   |||
Db     216366 TTGGAGGAG 216358

RESULT 55
AC120750
LOCUS       AC120750               251186 bp    DNA     linear   HTG 21-SEP-2002
DEFINITION  Rattus norvegicus clone CH230-54I13, *** SEQUENCING IN PROGRESS
            ***, 3 unordered pieces.
ACCESSION   AC120750
VERSION     AC120750.4  GI:23265324
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 251186)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
```

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 251186)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 251186)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Sep 21, 2002 this sequence version replaced gi:21908122.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequening reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GXZE
Center clone name: CH230-54I13
---------- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 217571 bases at least Q40
Consensus quality: 219475 bases at least Q30
Consensus quality: 220633 bases at least Q20
Estimated insert size: 237562; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
---------- * NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 247947: contig of 247947 bp in length
* 247948 248047: gap of unknown length
* 248048 249438: contig of 1391 bp in length
* 249439 249538: gap of unknown length
* 249539 251186: contig of 1648 bp in length.

FEATURES Location/Qualifiers
source 1. .251186
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-54I13"
misc_feature 1. .1555
/note="wgs_end_extension
clone_end:Sp6"
misc_feature 3004. .3858
/note="clone_boundary
clone_end:Sp6
site:EcoRI
end_sequence:BH349020"
misc_feature complement(246067. .246774)
/note="clone_boundary
clone_end:T7
site:EcoRI
end_sequence:BH349018"
gap 247948. .248047
/estimated_length=unknown
gap 249439. .249538
/estimated_length=unknown
ORIGIN

Alignment Scores:

Pred. No.:	7.2e+03	Length:	251186
Score:	68.00	Matches:	17
Percent Similarity:	52.5%	Conservative:	4
Best Local Similarity:	42.5%	Mismatches:	15
Query Match:	29.7%	Indels:	4
DB:	14	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x AC120750 (1-251186)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   ||||||   |||:::|||   :::   |||   |||           :::
Db     200334 GGTGACACTGGCTCATCAATACTGGAGAAAATACGGCCTGGAGGAGGACCGATTAGGGAC 200393

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||||||||   |||:::|||                  |||      ||||||   |||
Db     200394 AGCACCACAGCTGAGGATGAACATCAG-----------AGTGTCTACCCATCACTGCCC 200441

RESULT 56
AC095235
LOCUS       AC095235              258285 bp    DNA     linear   HTG 21-SEP-2002
DEFINITION  Rattus norvegicus clone CH230-10F21, *** SEQUENCING IN PROGRESS
            ***, 3 unordered pieces.
ACCESSION   AC095235
VERSION     AC095235.4  GI:22772987
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 258285)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 258285)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 258285)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Sep 10, 2002 this sequence version replaced gi:21722866.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequening reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```
            ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
                Center project name: GDMA
                Center clone name: CH230-10F21
            ---------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 214955 bases at least Q40
                Consensus quality: 218426 bases at least Q30
                Consensus quality: 221055 bases at least Q20
                Estimated insert size: 237775; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            ----------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   253660: contig of 253660 bp in length
            *   253661   253760: gap of unknown length
            *   253761   254890: contig of 1130 bp in length
            *   254891   254990: gap of unknown length
            *   254991   258285: contig of 3295 bp in length.
FEATURES             Location/Qualifiers
     source          1. .258285
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-10F21"
     misc_feature    1. .1361
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    6290. .7208
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH305475"
     misc_feature    77877. .79084
                     /note="wgs_contig"
     misc_feature    complement(252613. .253531)
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH305476"
     gap             253661. .253760
                     /estimated_length=unknown
     gap             254891. .254990
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:                7.41e+03        Length:          258285
Score:                    68.00           Matches:         17
Percent Similarity:       52.5%           Conservative:    4
Best Local Similarity:    42.5%           Mismatches:      15
Query Match:              29.7%           Indels:          4
DB:                       14              Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x AC095235 (1-258285)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   ||||||   |||:::|||  :::   |||  |||           :::
Db     151408 GGTGACACTGGCTCATCAATACTGGAGAAAATACGGCCTGGAGGAGGACCGATTAGGGAC 151467
```

```
Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               |||||||||   |||:::|||                |||      ||||||   |||
Db      151468 AGCACCACAGCTGAGGATGAACATCAG-----------AGTGTCTACCCATCACTGCCC 151515

RESULT 57
AC134158
LOCUS       AC134158               301536 bp    DNA     linear   HTG 13-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-172M3, *** SEQUENCING IN PROGRESS
            ***, 13 unordered pieces.
ACCESSION   AC134158
VERSION     AC134158.2  GI:24941826
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 301536)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,P., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 301536)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 301536)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 13, 2002 this sequence version replaced gi:23269965.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GSKI
                Center clone name: CH230-172M3
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 239973 bases at least Q40
                Consensus quality: 244429 bases at least Q30
                Consensus quality: 247205 bases at least Q20
                Estimated insert size: 238981; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 13 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *         1   243755: contig of 243755 bp in length
            *    243756   243855: gap of unknown length
            *    243856   253246: contig of 9391 bp in length
            *    253247   253346: gap of unknown length
            *    253347   264612: contig of 11266 bp in length
            *    264613   264712: gap of unknown length
            *    264713   270919: contig of 6207 bp in length
            *    270920   271019: gap of unknown length
            *    271020   285377: contig of 14358 bp in length
            *    285378   285477: gap of unknown length
            *    285478   290593: contig of 5116 bp in length
            *    290594   290693: gap of unknown length
            *    290694   292250: contig of 1557 bp in length
            *    292251   292350: gap of unknown length
            *    292351   293352: contig of 1002 bp in length
            *    293353   293452: gap of unknown length
            *    293453   294453: contig of 1001 bp in length
            *    294454   294553: gap of unknown length
            *    294554   295736: contig of 1183 bp in length
            *    295737   295836: gap of unknown length
            *    295837   297160: contig of 1324 bp in length
            *    297161   297260: gap of unknown length
            *    297261   299442: contig of 2182 bp in length
            *    299443   299542: gap of unknown length
            *    299543   301536: contig of 1994 bp in length.
FEATURES             Location/Qualifiers
```

```
     source          1. .301536
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-172M3"
     misc_feature    1. .1160
                     /note="wgs_contig"
     misc_feature    241830. .243755
                     /note="wgs_contig"
     gap             243756. .243855
                     /estimated_length=unknown
     gap             253247. .253346
                     /estimated_length=unknown
     misc_feature    253347. .254480
                     /note="wgs_contig"
     misc_feature    254531. .255708
                     /note="wgs_contig"
     gap             264613. .264712
                     /estimated_length=unknown
     gap             270920. .271019
                     /estimated_length=unknown
     misc_feature    271020. .272426
                     /note="wgs_contig"
     gap             285378. .285477
                     /estimated_length=unknown
     gap             290594. .290693
                     /estimated_length=unknown
     gap             292251. .292350
                     /estimated_length=unknown
     gap             293353. .293452
                     /estimated_length=unknown
     gap             294454. .294553
                     /estimated_length=unknown
     gap             295737. .295836
                     /estimated_length=unknown
     gap             297161. .297260
                     /estimated_length=unknown
     gap             299443. .299542
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:                8.7e+03        Length:          301536
Score:                    68.00          Matches:         18
Percent Similarity:       50.0%          Conservative:    3
Best Local Similarity:    42.9%          Mismatches:      11
Query Match:              29.7%          Indels:          10
DB:                       14             Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x AC134158 (1-301536)

Qy          2 AspProGlySerGluIleIleGluSerValProPro---------AlaGlyProGluAla 18
              ||||||||||||:::::|||      :::||||||           ||||||||||
Db      20657 GACCCTGGAAGCCAATTGATTGTGCCCTTACCCCCCACTCCCCAAGGTGGACCGGAGGCT 20716

Qy         19 SerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSer 38
                                 |||                |||   |||      ||||||
Db      20717 --------------------AACCACTCCCAGCCCTTCTTCCATTCGCCGGATCCATCA 20755

Qy         39 ArgPro 40
              ||||||
Db      20756 CGACCG 20761

RESULT 58
BX255276_16
WPCOMMENT
  Sequence split into 17 fragments  LOCUS BX255276 Accession BX255276
     Fragment Name      Begin     End
     BX255276_00            1     110000
     BX255276_01       100001     210000
     BX255276_02       200001     310000
     BX255276_03       300001     410000
     BX255276_04       400001     510000
     BX255276_05       500001     610000
     BX255276_06       600001     710000
     BX255276_07       700001     810000
     BX255276_08       800001     910000
     BX255276_09       900001    1010000
     BX255276_10      1000001    1110000
     BX255276_11      1100001    1210000
     BX255276_12      1200001    1310000
     BX255276_13      1300001    1410000
     BX255276_14      1400001    1510000
     BX255276_15      1500001    1610000
     BX255276_16      1600001    1691152
Continuation (17 of 17) of BX255276 from base 1600001 (BX255276 Homo sapiens chromosome

Alignment Scores:
Pred. No.:                2.93e+03       Length:          91152
Score:                    67.50          Matches:         17
Percent Similarity:       47.2%          Conservative:    8
Best Local Similarity:    32.1%          Mismatches:      17
Query Match:              29.5%          Indels:          11
DB:                       14             Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x BX255276_16 (1-91152)

Qy          1 GlyAspProGlySerGluIleIle---GluSerValProProAlaGlyProGluAlaSer 19
              |||   ||||||:::::::::::::   :::   |||          ||||||   :::
Db      55747 GGAACGCCGGGTGCGCAACTGGTCAACCAGGGCGTGATCGGGGTTCCACCCGAACTCAAC 55806

Qy         20 GluSerThrThrAspGluAsnGluAspAsp----------------------------- 29
              |||              |||  |||
Db      55807 GAGCCCAACAAGACCAACAACACCGATAGCGCGGCGGTGAGCTTCAGCAACGATGCGCGG 55866

Qy         30 IleGlnPheValSerGluGlyProSerArgProValLeu 42
              ::::::   |||       ||||||||||||||||||||
Db      55867 CTGCGCCTCGTCAAGGCCGGCCCCAGCCGCCCGGTACTG 55905

RESULT 59
SFVDI6A
LOCUS       SFVDI6A                 2146 bp ss-RNA     linear   VRL 27-JUN-1994
DEFINITION  Semliki forest virus gene, DI-6.
ACCESSION   L27470
VERSION     L27470.1  GI:440843
KEYWORDS    .
SOURCE      Semliki forest virus
  ORGANISM  Semliki forest virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
            Alphavirus; SFV complex.
REFERENCE   1  (sites)
  AUTHORS   Thomson,M. and Dimmock,N.J.
  TITLE     Common sequence elements in structurally unrelated genomes of
            defective interfering Semliki Forest virus
  JOURNAL   Virology 199 (2), 354-365 (1994)
   PUBMED   7510084
COMMENT     Original source text: Semliki forest virus (individual_isolate
            DI-6) cDNA to genomic RNA.
FEATURES             Location/Qualifiers
     source          1. .2146
                     /organism="Semliki forest virus"
                     /mol_type="genomic RNA"
                     /isolate="DI-6"
                     /db_xref="taxon:11033"
ORIGIN

Alignment Scores:
Pred. No.:                69.8           Length:          2146
Score:                    67.00          Matches:         14
Percent Similarity:       51.2%          Conservative:    7
Best Local Similarity:    34.1%          Mismatches:      20
Query Match:              29.3%          Indels:          0
DB:                       13             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x SFVDI6A (1-2146)
```

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||      :::::::::   ::::::|||                |||:::
Db       1566 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 1625

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                     |||       |||   |||      |||   ||||||||||:::   |||
Db       1626 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 1685

Qy         41 Val 41
              |||
Db       1686 GTG 1688

RESULT 60
SFVY17207
LOCUS       SFVY17207               4093 bp    RNA     linear   VRL 18-APR-2005
DEFINITION  Semliki Forest virus genomic RNA for nsP1 and nsP2 proteins.
ACCESSION   Y17207
VERSION     Y17207.1  GI:3114969
KEYWORDS    NSP1; NSP2.
SOURCE      Semliki forest virus
  ORGANISM  Semliki forest virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
            Alphavirus; SFV complex.
REFERENCE   1
  AUTHORS   Tuittila,M.T., Santagati,M.G., Roytta,M., Maatta,J.A. and
            Hinkkanen,A.E.
  TITLE     Replicase complex genes of Semliki Forest virus confer lethal
            neurovirulence
  JOURNAL   J. Virol. 74 (10), 4579-4589 (2000)
   PUBMED   10775594
REFERENCE   2  (bases 1 to 4093)
  AUTHORS   Hinkkanen,A.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-1998) A.E. Hinkkanen, Abo Akademi University,
            Dept Biochemistry and Pharmacy, Biocity, Tykistoekatu 6 A, 20520
            Turku, FINLAND
FEATURES             Location/Qualifiers
     source          1. .4093
                     /organism="Semliki forest virus"
                     /mol_type="genomic RNA"
                     /strain="SFV A7 (74)"
                     /isolate="MA774"
                     /db_xref="taxon:11033"
     CDS             86. .>4093
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="CAA76683.1"
                     /db_xref="GI:3114970"
                     /db_xref="GOA:O73550"
                     /db_xref="InterPro:IPR000606"
                     /db_xref="InterPro:IPR002620"
                     /db_xref="InterPro:IPR010916"
                     /db_xref="UniProt/TrEMBL:O73550"
                     /translation="MAAKVHVDIEADSPFIKSLQKAFPSFEVESLQVTPNDHANARAF
                     SHLATKLIEQETDKDTLILDIGSAPSRRMMSTHKYHCVCPMRSAEDPERLVCYAKKLA
                     AASGKVLDREIAGKITDLQTVMATPDAESPTFCLHTDVTCRTAAEVAVYQDVYAVHAP
                     TSLYHQAMKGVRTAYWIGFDTTPFMFDALAGAYPTYATNWADEQVLQARNIGLCAASL
                     TEGRLGKLSILRKKQLKPSDTVMFSVGSTLYTESRKLLRSWHLPSVFHLKGKQSFTCR
                     CDTIVSCEGYVVKKITMCPGLYGKTVGYAVTHHAEGFLVCKTTDTVKGERVSFPVCTY
                     VPSTICDQMTGILATDVTPEDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPVVAVAF
                     SKWAREYKADLDDEKPLGVRERSLTCCCLWAFKTKKMHTMYKKPDTQTIVKVPSEFNS
                     FVIPSLWSTGLAIPVRSRIKMLLAKKTKRELIPALDASSARDAEQEEKERLEAELTRE
                     ALPPLVPIAPAETGVVDVDVEELEYRAGAGVVETPRSALKVTAQPNDVLLGNYVVLSP
                     QTVLKSSKLAPVHPLAEQVKIITHNGRAGRYQVDGYDGRVLLPCGSAIPVPEFQALSE
                     SATMVYNEREFVNRKLYHIAVHGPSLNTDEENYEKVRAERTDAEYVFDVDKKCCIKRE
                     EASGLVLVGELTNPPFHEFAYEGLKIRPSAPYKTTVVGVFGVPGSGKSAIIKSLVTKH
                     DLVTSGKKENCQEIVNDVKKHRGLDIQAKTVDSILLNGCRRAVDILYVDEAFACHSGT
                     LLALIALVKPRSKVVLCGDPKQCGFFNMMQLKVNFNHNICTEVCHKSISRRCTRPVTA
                     IVSTLHYGGKMRTTNPCNKPIIIDTTGQTKPKPGDIVLTCFRGWVKQLQLDYRGHEVM
                     TAAASQGLTRKGVYAVRQKVNENPLYAPASEHVNVLLTRTEDRLVWKTLAGDPWIKVL
                     SNIPQGNFTATLEEWQEEHDKIMKVIEGPAAPVDAFQNKANVCWAKSLVPVLDTAGIR
                     LTAEEWSTIITAFKEDRAYSPEVALNEICTKYYGVDLDSGLFSAPKVSLYYENNHWDN
                     RPGGRMYGFNAATAARLEARHTFLKGQWHTGKQAVIAERKIQPLSVLDNVIPINRRLP
                     HALVAEYKTVKGSRVEWLVNKVRGYHVLLVSEYNLALPRRRVTWLSPLNVTGADRCYD
                     LSLGLPADAGRYDLVFVNIHTEFRIHHYQQCVDHAMKLQMLGGDALRLLKPGGNLLMR
                     AYGYADKISEAVVSSLSRKFSSARVLRPDCVTSNTEVFLLFSNFDNGKRPSTLHQMNT
                     KLSAVYAGEAMHTAGC"
     mat_peptide     86. .1696
                     /product="nsP1"
     mat_peptide     1697. .>4093
                     /product="unnamed"
     variation       223
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="c"
     variation       369. .371
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="ata"
     variation       660
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="g"
     variation       826
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="a"
     variation       1365
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="g"
     variation       1650. .1651
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="tg"
     variation       1808
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       2021
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="a"
     variation       2193
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       2284
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       2350
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       2372. .2398
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="gggaactggacatccagggaaaacagt"
     variation       2472
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="a"
     variation       2879. .2881
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="aaa"
     variation       3280
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       3355
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       3359
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="a"
     variation       3428
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="a"
     variation       3466
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       3520
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="g"
     variation       3616
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       3643
                     /note="compared to Z48163, strain A7 of SFV"
```

```
                     /replace="g"
     variation       3656
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       3660
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       3673
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       3675
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="a"
     variation       3679
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="c"
     variation       3693
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       3712
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="c"
     variation       3727
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="g"
     variation       3730
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="c"
     variation       3732. .3733
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="ta"
     variation       3748
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="a"
     variation       3820
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
     variation       3883
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="c"
     variation       3946
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="g"
     variation       4093
                     /note="compared to Z48163, strain A7 of SFV"
                     /replace="t"
ORIGIN

Alignment Scores:
Pred. No.:                136             Length:         4093
Score:                    67.00           Matches:        14
Percent Similarity:       51.2%           Conservative:   7
Best Local Similarity:    34.1%           Mismatches:     20
Query Match:              29.3%           Indels:         0
DB:                       13              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x SFVY17207 (1-4093)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||                 |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||      |||   |||      |||   |||||||||:::   |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100

RESULT 61
SFO251359
LOCUS       SFO251359               7422 bp    RNA     linear   VRL 15-APR-2005
DEFINITION  Semliki Forest virus ORF1, genomic RNA, clone SFV4.
ACCESSION   AJ251359
VERSION     AJ251359.1  GI:6523490
KEYWORDS    nonstructural protein; NSP gene; NSP1; NSP2; NSP3; NSP4;
            polyprotein nsP1234.
SOURCE      Semliki forest virus
  ORGANISM  Semliki forest virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
            Alphavirus; SFV complex.
REFERENCE   1
  AUTHORS   Tuittila,M.T., Santagati,M.G., Roytta,M., Maatta,J.A. and
            Hinkkanen,A.E.
  TITLE     Replicase complex genes of Semliki Forest virus confer lethal
            neurovirulence
  JOURNAL   J. Virol. 74 (10), 4579-4589 (2000)
   PUBMED   10775594
REFERENCE   2  (bases 1 to 7422)
  AUTHORS   Tuittila,M.T.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-NOV-1999) Tuittila M.T., Department of Biochemistry
            and Pharmacy, Abo Akademi University, Tykistokatu 6A, 20520 Turku,
            FINLAND
FEATURES             Location/Qualifiers
     source          1. .7422
                     /organism="Semliki forest virus"
                     /virion
                     /mol_type="genomic RNA"
                     /isolate="L10"
                     /specific_host="rodents, mosquitoes"
                     /db_xref="taxon:11033"
                     /map="1-7422"
                     /clone="SFV4"
                     /lab_host="mouse"
                     /country="Uganda:Bwamba"
     CDS             86. .7384
                     /codon_start=1
                     /product="polyprotein nsP1234"
                     /protein_id="CAB62256.1"
                     /db_xref="GI:6523491"
                     /db_xref="GOA:Q9QBM1"
                     /db_xref="InterPro:IPR000606"
                     /db_xref="InterPro:IPR001788"
                     /db_xref="InterPro:IPR002589"
                     /db_xref="InterPro:IPR002620"
                     /db_xref="InterPro:IPR007094"
                     /db_xref="InterPro:IPR007095"
                     /db_xref="InterPro:IPR010916"
                     /db_xref="UniProt/TrEMBL:Q9QBM1"
                     /translation="MAAKVHVDIEADSPFIKSLQKAFPSFEVESLQVTPNDHANARAF
                     SHLATKLIEQETDKDTLILDIGSAPSRRMMSTHKYHCVCPMRSAEDPERLVCYAKKLA
                     AASGKVLDREIAGKITDLQTVMATPDAESPTFCLHTDVTCRTAAEVAVYQDVYAVHAP
                     TSLYHQAMKGVRTAYWIGFDTTPFMFDALAGAYPTYATNWADEQVLQARNIGLCAASL
                     TEGRLGKLSILRKKQLKPCDTVMFSVGSTLYTESRKLLRSWHLPSVFHLKGKQSFTCR
                     CDTIVSCEGYVVKKITMCPGLYGKTVGYAVTYHAEGFLVCKTTDTVKGERVSFPVCTY
                     VPSTICDQMTGILATDVTPEDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPIVAVAF
                     SKWAREYKADLDDEKPLGVRERSLTCCCLWAFKTRKMHTMYKKPDTQTIVKVPSEFNS
                     FVIPSLWSTGLAIPVRSRIKMLLAKKTKRELIPVLDASSARDAEQEEKERLEAELTRE
                     ALPPLVPIAPAETGVVDVDVEELEYHAGAGVVETPRSALKVTAQPNDVLLGNYVVLSP
                     QTVLKSSKLAPVHPLAEQVKIITHNGRAGRYQVDGYDGRVLLPCGSAIPVPEFQALSE
                     SATMVYNEREFVNRKLYHIAVHGPSLNTDEENYEKVRAERTDAEYVFDVDKKCCVKRE
                     EASGLVLVGELTNPPFHEFAYEGLKIRPSAPYKTTVVGVFGVPGSGKSAIIKSLVTKH
                     DLVTSGKKENCQEIVNDVKKHRGLDIQAKTVDSILLNGCRRAVDILYVDEAFACHSGT
                     LLALIALVKPRSKVVLCGDPKQCGFFNMMQLKVNFNHNICTEVCHKSISRRCTRPVTA
                     IVSTLHYGGKMRTTNPCNKPIIIDTTGQTKPKPGDIVLTCFRGWVKQLQLDYRGHEVM
                     TAAASQGLTRKGVYAVRQKVNENPLYAPASEHVNVLLTRTEDRLVWKTLAGDPWIKVL
                     SNIPQGNFTATLEEWQEEHDKIMKVIEGPAAPVDAFQNKANVCWAKSLVPVLDTAGIR
                     LTAEEWSTIITAFKEDRAYSPVVALNEICTKYYGVDLDSGLFSAPKVSLYYENNHWDN
                     RPGGRMYGFNAATAARLEARHTFLKGQWHTGKQAVIAERKIQPLSVLDNVIPINRRLP
                     HALVAEYKTVKGSRVEWLVNKVRGYHVLLVSEYNLALPRRRVTWLSPLNVTGADRCYD
                     LSLGLPADAGRFDLVFVNIHTEFRIHHYQQCVDHAMKLQMLGGDALRLLKPGGSLLMR
                     AYGYADKISEAVVSSLSRKFSSARVLRPDCVTSNTEVFLLFSNFDNGKRPSTLHQMNT
                     KLSAVYAGEAMHTAGCAPSYRVKRADIATCTEAAVVNAANARGTVGDGVCRAVAKKWP
                     SAFKGEATPVGTIKTVMCGSYPVIHAVAPNFSATTEAEGDRELAAVYRAVAAEVNRLS
```

```
                     LSSVAIPLLSTGVFSGGRDRLQQSLNHLFTAMDATDADVTIYCRDKSWEKKIQEAIDM
                     RTAVELLNDDVELTTDLVRVHPDSSLVGRKGYSTTDGSLYSYFEGTKFNQAAIDMAEI
                     LTLWPRLQEANEQICLYALGETMDNIRSKCPVNDSDSSTPPRTVPCLCRYAMTAERIA
                     RLRSHQVKSMVVCSSFPLPKYHVDGVQKVKCEKVLLFDPTVPSVVSPRKYAASTTDHS
                     DRSLRGFDLDWTTDSSSTASDTMSLPSLQSCDIDSIYEPMAPIVVTADVHPEPAGIAD
                     LAADVHPEPADHVDLENPIPPPRPKRAAYLASRAAERPVPAPRKPTPAPRTAFRNKLP
                     LTFGDFDEHEVDALASGITFGDFDDVLRLGRAGAYIFSSDTGSGHLQQKSVRQHNLQC
                     AQLDAVEEEKMYPPKLDTEREKLLLLKMQMHPSEANKSRYQSRKVENMKATVVDRLTS
                     GARLYTGADVGRIPTYAVRYPRPVYSPTVIERFSSPDVAIAACNEYLSRNYPTVASYQ
                     ITDEYDAYLDMVDGSDSCLDRATFCPAKLRCYPKHHAYHQPTVRSAVPSPFQNTLQNV
                     LAAATKRNCNVTQMRELPTMDSAVFNVECFKRYACSGEYWEEYAKQPIRITTENITTY
                     VTKLKGPKAAALFAKTHNLVPLQEVPMDRFTVDMKRDVKVTPGTKHTEERPKVQVIQA
                     AEPLATAYLCGIHRELVRRLNAVLRPNVHTLFDMSAEDFDAIIASHFHPGDPVLETDI
                     ASFDKSQDDSLALTGLMILEDLGVDQYLLDLIEAAFGEISSCHLPTGTRFKFGAMMKS
                     GMFLTLFINTVLNITIASRVLEQRLTDSACAAFIGDDNIVHGVISDKLMAERCASWVN
                     MEVKIIDAVMGEKPPYFCGGFIVFDSVTQTACRVSDPLKRLFKLGKPLTAEDKQDEDR
                     RRALSDEVSKWFRTGLGAELEVALTSRYEVEGCKSILIAMATLARDIKAFKKLRGPVI
                     HLYGGPRLVR"
     mat_peptide     89. .1696
                     /gene="nsp1"
                     /product="nonstructural protein nsP1"
     mat_peptide     1697. .4093
                     /gene="nsp2"
                     /product="nonstructural protein nsP2"
     mat_peptide     4094. .5539
                     /gene="nsp3"
                     /product="nonstructural protein nsP3"
     mat_peptide     5540. .7381
                     /gene="nsp4"
                     /product="nonstructural protein nsP4"
     gene            89. .1696
                     /gene="nsp1"
     gene            1697. .4093
                     /gene="nsp2"
     gene            4094. .5539
                     /gene="nsp3"
     gene            5540. .7381
                     /gene="nsp4"
ORIGIN

Alignment Scores:
Pred. No.:              253          Length:        7422
Score:                  67.00        Matches:       14
Percent Similarity:     51.2%        Conservative:  7
Best Local Similarity:  34.1%        Mismatches:    20
Query Match:            29.3%        Indels:        0
DB:                     13           Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x SFO251359 (1-7422)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||               |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||      |||  |||      |||  |||||||||:::  |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100

RESULT 62
BD138791
LOCUS       BD138791               8000 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Alphavirus vectors.
ACCESSION   BD138791
VERSION     BD138791.1  GI:23233736
KEYWORDS    JP 2002508925-A/3.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 8000)
  AUTHORS   Parrington,M.
  TITLE     Alphavirus vectors
  JOURNAL   Patent: JP 2002508925-A 3 26-MAR-2002;
            CONNAUGHT LABORATORIES LTD
COMMENT     OS   Alpha virus
            PN   JP 2002508925-A/3
            PD   26-MAR-2002
            PF   13-NOV-1998 JP 2000521222
            PR   14-NOV-1997 US    60/065793
            PI   MARK PARRINGTON
            PC   C12N15/09//(C12N15/09,C12R1:94),C12N15/00,(C12N15/00,C12R1:94)
            CC   Alphavirus vectors
            FH   Key             Location/Qualifiers
            FT   source          1. .8010
            FT                   /organism='Alpha virus'.
FEATURES             Location/Qualifiers
     source          1. .8000
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:              273          Length:        8000
Score:                  67.00        Matches:       14
Percent Similarity:     51.2%        Conservative:  7
Best Local Similarity:  34.1%        Mismatches:    20
Query Match:            29.3%        Indels:        0
DB:                     6            Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x BD138791 (1-8000)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||               |||:::
Db       2701 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 2760

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||      |||  |||      |||  |||||||||:::  |||
Db       2761 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 2820

Qy         41 Val 41
              |||
Db       2821 GTG 2823

RESULT 63
BD226819
LOCUS       BD226819               8100 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Alphavirus vectors for paramyxovirus vaccines.
ACCESSION   BD226819
VERSION     BD226819.1  GI:33036589
KEYWORDS    JP 2002512003-A/4.
SOURCE      Semliki forest virus
  ORGANISM  Semliki forest virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
            Alphavirus; SFV complex.
REFERENCE   1  (bases 1 to 8100)
  AUTHORS   Parrington,M. and Li,X.
  TITLE     Alphavirus vectors for paramyxovirus vaccines
  JOURNAL   Patent: JP 2002512003-A 4 23-APR-2002;
            CONNAUGHT LABORATORIES LTD
COMMENT     OS   Semliki forest virus (SFV)
            PN   JP 2002512003-A/4
            PD   23-APR-2002
            PF   13-NOV-1998 JP 2000521221
            PR   14-NOV-1997 US    60/065791
            PI   MARK PARRINGTON,XIAOMAO LI
            PC   C12N15/09,A61K31/711,A61K39/155,A61P31/12,C07K14/115,C07K14/
            PC   135,C12N15/00
            CC   Alphavirus vectors for paramyxovirus vaccines FH   Key
                 Location/Qualifiers
            FT   source          1. .8100
            FT                   /organism='Semliki forest virus (SFV)'.
FEATURES             Location/Qualifiers
```

```
     source          1. .8100
                     /organism="Semliki forest virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11033"
ORIGIN

Alignment Scores:
Pred. No.:                277             Length:         8100
Score:                    67.00           Matches:        14
Percent Similarity:       51.2%           Conservative:   7
Best Local Similarity:    34.1%           Mismatches:     20
Query Match:              29.3%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BD226819 (1-8100)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||      :::::::::   :::::::|||                ||||::
Db       2701 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 2760

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                       |||       |||   |||       |||   ||||||||||:::   |||
Db       2761 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 2820

Qy         41 Val 41
              |||
Db       2821 GTG 2823

RESULT 64
AR243302
LOCUS       AR243302                8100 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 4 from patent US 6475780.
ACCESSION   AR243302
VERSION     AR243302.1  GI:27290492
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 8100)
  AUTHORS   Parrington,M., Li,X. and Klein,M.H.
  TITLE     Alphavirus vectors for paramyxovirus vaccines
  JOURNAL   Patent: US 6475780-A 4 05-NOV-2002;
            Aventis Pasteur Limited; Toronto;
            CAX;
FEATURES             Location/Qualifiers
     source          1. .8100
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                277             Length:         8100
Score:                    67.00           Matches:        14
Percent Similarity:       51.2%           Conservative:   7
Best Local Similarity:    34.1%           Mismatches:     20
Query Match:              29.3%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AR243302 (1-8100)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||      :::::::::   :::::::|||                ||||::
Db       2701 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 2760

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                       |||       |||   |||       |||   ||||||||||:::   |||
Db       2761 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 2820

Qy         41 Val 41
              |||
Db       2821 GTG 2823

RESULT 65
BD138792
LOCUS       BD138792                11000 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Alphavirus vectors.
ACCESSION   BD138792
VERSION     BD138792.1  GI:23233737
KEYWORDS    JP 2002508925-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 11000)
  AUTHORS   Parrington,M.
  TITLE     Alphavirus vectors
  JOURNAL   Patent: JP 2002508925-A 4 26-MAR-2002;
            CONNAUGHT LABORATORIES LTD
COMMENT     OS   Alpha virus
            PN   JP 2002508925-A/4
            PD   26-MAR-2002
            PF   13-NOV-1998 JP 2000521222
            PR   14-NOV-1997 US    60/065793
            PI   MARK PARRINGTON
            PC   C12N15/09//(C12N15/09,C12R1:94),C12N15/00,(C12N15/00,C12R1:94)
            CC   Alphavirus vectors
            FH   Key             Location/Qualifiers
            FT   source          1. .12474
            FT                   /organism='Alpha virus'.
FEATURES             Location/Qualifiers
     source          1. .11000
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                380             Length:         11000
Score:                    67.00           Matches:        14
Percent Similarity:       51.2%           Conservative:   7
Best Local Similarity:    34.1%           Mismatches:     20
Query Match:              29.3%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BD138792 (1-11000)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||      :::::::::   :::::::|||                ||||::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                       |||       |||   |||       |||   ||||||||||:::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101

RESULT 66
AY112987
LOCUS       AY112987                11406 bp    RNA     linear   VRL 01-JUL-2002
DEFINITION  Semliki forest virus strain L10, complete genome.
ACCESSION   AY112987
VERSION     AY112987.1  GI:21655311
KEYWORDS    .
SOURCE      Semliki forest virus
  ORGANISM  Semliki forest virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
            Alphavirus; SFV complex.
REFERENCE   1  (bases 1 to 11406)
  AUTHORS   Logue,C., Mooney,D., Shanley,R. and Atkins,G.J.
  TITLE     Semliki Forest Virus - L10 Strain Complete Genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 11406)
  AUTHORS   Logue,C., Mooney,D., Shanley,R. and Atkins,G.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2002) Microbiology, Trinity College Dublin, 2
```

College Green, Dublin 2, Ireland

```
FEATURES             Location/Qualifiers
     source          1. .11406
                     /organism="Semliki forest virus"
                     /mol_type="genomic RNA"
                     /strain="L10"
                     /db_xref="taxon:11033"
     5'UTR           1. .86
     CDS             87. .7382
                     /codon_start=1
                     /product="non-structural polyprotein"
                     /protein_id="AAM64226.1"
                     /db_xref="GI:21655312"
                     /translation="MAAKVYVDIEADSPFIKSLQKAFPSFEVESLQVTPNDHANARAF
                     SHLATKLIEQETDKDTLILDIGSAPSRRMMSTHKYHCVCPMRSAEDPERLVCYAKKLA
                     AASGKVLDREIAGKITDLQTVMATPDAESPTFCLHTDVTCRTAAEVAVYQDVYAVHAP
                     TSLYHQAMKGVRTAYWIGFDTTPFMFDALAGAYPTYATNWADEQVLQARNIGLCAASL
                     TEGRLGKLSILRKKQLKPCDTVMFSVGSTLYTESRKLLRSWHLPSVFHLKGKQSFTCR
                     CDTIVSCEGYVVKKITMCPGLYGKTVGYAVTYHAKGFLVCKTTDTVKGERVSFPVCTY
                     VPSTICDQMTGILATDVTPEDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPIVAVAF
                     SKWAREYKADLDDEKPLGVRERSLTCCCLWAFKTRKMHTMYKKPDTQTIVKVPSEFNS
                     FVIPSLWSTGLAIPVRSRIKMLLAKKTKRELIPVLDASSARDAEQEEKERLEAELTRE
                     ALPPLVPIAPAETGVVDVDVEELEYHAGAGVVETPRSALKVTAQPNDVLLGNYVVLSP
                     QTVLKSSKLAPVHPLAEQVKIITHNGRAGRYQVDGYDGRVLLPCGSAIPVPEFQALSE
                     SATMVYNEREFVNRKLYHIAVHGPSLNTDEENYEKVRAERTDAEYVFDVDKKCCVKRE
                     EASGLVLVGELTNPPFHEFAYEGLKIRPSAPYKTTVVGVFGVPGSGKSAIIKSLVTKH
                     DLVTSGKKENCQEIVNDVKKHRGNWTSRKNSDSILLNGCRRAVDILYVDEAFACHSGT
                     LLALIALVKPRSKVVLCGNPKQCGFFNTMQLKVNFNHNICTEVCNKSISRRCTRPVTA
                     IVCTLHYGGKMRTTNPCNKPIIIDTTGQTKPKPGDIVLTCFRGWVKQLQLDYRGHEVM
                     TAAASQGLTRKGVYAVRQKVNENPLYAPASEHVNVLLTRTEDRLVWKTLAGDPWIKVL
                     SNIPQGNFTATLEEWQEEHDKIMKVIEGPAAPVDAFQNKANVCWAKSLVPVLDTAGIR
                     LTAEEWSTIITAFKEDRAYSPVVALNEICTKYYGVDLDSGLFSAPKVSLYYENNHWDN
                     RPGGRMYGFNAATAARLEARHTFLKGQWHTGKQAVIAERKIQPLSVLDNVIPINRRLP
                     HALVAEYKTVKGSRVEWLVNKVRGYHVLLVSEYNLALPRRRVTWLSPLNVTGADRCYD
                     LSLGLPADAGRFDLVFVNIHTEFRIHHYQQCVDHAMKLQMLGGDALRLLKPGGILMRA
                     YGYADKISEAVVSSLSRKFSSARVLRPDCVTSNTEVFLLFSNFDNGKRPSTLHQMNTK
                     LSAVYAGEAMHTAGCAPSYRVKRADIATCTEAAVVNAANARGTVGDGVCRAVAKKWPS
                     AFKGAATPVGTIKTVMCGSYPVIHAVAPNFSATTEAEGDRELAAAVYRAVAAEVNRLSL
                     SSVAIPLLSTGVFSGGRDRLQQSLNHLFTAMDATDADVTIYCRDKSWEKKIQEAIDMR
                     TAVELLNDDVELTTDLVRVHPDSSLVGRKGYSTTDGSLYSYFEGTKFNQAAIDMAEIL
                     TLWPRLQEANEQICLYALGETMDNIRSKCPVNDSDSSTPPRTVPCLCRYAMTAERIAR
                     LRSHQVKSMVVCSSFPLPKYHVDGVQKVKCEKVLLFDPTVPSVVSPRKYAASTTDHSD
                     RSLRGFDLDWTTDSSSTASDTMSLPSLQSCDIDSIYEPMAPIVVTADVHPEPAGIADL
                     AADVHPEPADHVDLENPIPPPRPKRAAYLASRAAERPVPAPRKPTPAPRTAFRNKLPL
                     TFGDFDEHEVDALASGITFGDFDDVLRLGRAGAYIFSSDTGSGHLQQKSVRQHNLQCA
                     QLDAVEEEKMYPPKLDTEREKLLLLKMQMHPSEANKSRYQSRKVENMKATVVDRLTSG
                     ARLYTGADVGRIPTYAVRYRRPVYSPTVIERFSSPDAAIAACNEYLSRNYPTVASYQI
                     TDEYDAYLDMVDGSDSCLDRATFCPAKLRCYPKHHAYHQPTVRSAVPSPFQNTLQNVL
                     AAATKRNCNVTQMRELPTMDSAVFNVECFKRYACSGEYWEEYAKQPIRITTENITTYV
                     TKLKGPKAAALFDKTHNLVPLQEVPMDRFTVDMKRDVKVTPGTKHTEERPKVQVIQAA
                     EPLATAYLCGIHRELVRRLNAVLRPNVHTLFDMSAEDFDAIIASHFHPGDPVLETDIA
                     SFDKSQDDSLALTGLMILEDLGVDQYLLDLIEAAFGEISSCHLPTGTRFKFGAMMKSG
                     MFLTLFINTVLNITIASRVLEQRLTDSACAAFIGDDNIVHGVISDKLMAERCASWVNM
                     EVKIIDAVMGEKPPYFCGGFIVFDSVTQTACRVSDPLKRLFKLGKPLTAEDKQDEDRR
                     RALSDEVSKWFRTGLGAELEVALTSRYEVEGCKSILIAMATLARDIKAFKKLRGPVIH
                     LYGGPRLVR"
     mat_peptide     87. .1697
                     /product="NSP1"
     mat_peptide     1698. .4091
                     /product="NSP2"
     mat_peptide     4092. .5537
                     /product="NSP3"
     mat_peptide     5538. .7370
                     /product="NSP4"
     rRNA            7371. .7420
                     /product="26S ribosomal RNA"
     CDS             7421. .11182
                     /codon_start=1
                     /product="structural polyprotein"
                     /protein_id="AAM64227.1"
                     /db_xref="GI:21655313"
                     /translation="MNYIPTQTFYGRRWRPRPAARPWPLQATPVAPVVPDFQAQQMQQ
                     LISAVNALTMRQNAIAPAGPPKPKKKKTTKPKPKTQPKKIKGKTQQQKKKDKQADKKK
                     KKPGKRERMCMKIENDCIFEVKHEGKVTGYACLVGDKVMKPAHVKGVIDNADLAKLAF
                     KKSSKYDLECAQIPVHMRSDASKYTHEKPEGHYNWHHGAVQYSGGRFTIPTGAGKPGD
                     SGRPIFDNKGRVVAIVLGGANEGSRTALSVVTWNKDMVTRVTPEGSEEWSAPLITAMC
                     VLANATFPCFQPPCVPCCYENNAEATLRMLEDNVDRPGYYDLLQAALTCRNGTRHRRS
                     VSQHFNVYKATRPYIAYCADCGAGHSCHSPVAIEAIRSEATDGMLKIQFSAQIGIDKS
                     DNHDYTKIRYADGHAIENAVRSSLKVATSGDCFVHGTMGHFILATCPPGEFLQVSIQD
                     TRNAVRACRIQYHHDPQPVGREKFTIRPHYGKEIPCTTYQQTTAETVEEIDMHMPPDT
                     PDRTLLSQQSGNVKITVGGKKVKYNCTCGTGNVGTTNSDMTINTCLIEQCHVSVTDHK
                     KWQFNSPFVPRADEPARKGKVHIPFPLDNITCRVPMAREPTVIHGKREVTLHLHPDHP
                     TLFSYRTLGEDPQYHEEWVTAAVERTIPVPVDGMEYHWGNNDPVRLWSQLTTEGKPHG
                     WPHQIVQYYYGLYPAATVSAVVGMSLLALISIFASCYMLVAARSKCLTPYALTPGAAV
                     PWTLGILCCAPRAHAASVAETMAYLWDQNQALFWLEFAAPVACILIITYCLRNVLCCC
                     KSLSFLVLLSLGATARAYEHSTVMPNVVGFPYKAHIERPGYSPLTLQMQVVETSLEPT
                     LNLEYITCEYKTVVPSPYVKCCGASECSTKEKPDYQCKVYTGVYPFMWGGAYCFCDSE
                     NTQLSEAYVDRSDVCRHDHASAYKAHTASLKAKVRVMYGNVNQTVDVYVNGDHAVTIG
                     GTQFIFGPLSSAWTPFDNKIVVYKDEVFNQDFPPYGSGQPGRFGDIQSRTVESNDLYA
                     NTALKLARPSPGMVHVPYTQTPSGFKYWLKEKGTALNTKAPFGCQIKTNPVRAMNCAV
                     GNIPVSMNLPDSAFTRIVEAPTIIDLTCTVATCTHSSDFGGVLTLTYKTDKNGDCSVH
                     SHSNVATLQEATAKVKTAGKVTLHFSTASASPSFVVSLCSARATCSASCEPPKDHIVP
                     YAASHSNVVFPDMSGTALSWVQKISGGLGAFAIGAILVLVVVTCIGLRR"
     mat_peptide     7421. .8221
                     /product="capsid protein"
     mat_peptide     8222. .8419
                     /product="E3 protein"
     mat_peptide     8420. .9685
                     /product="E2 protein"
     mat_peptide     9686. .9865
                     /product="6 kDa protein"
     mat_peptide     9866. .11179
                     /product="E1 protein"
     3'UTR           11180. .11406
ORIGIN

Alignment Scores:
Pred. No.:                395           Length:          11406
Score:                    67.00         Matches:         14
Percent Similarity:       51.2%         Conservative:    7
Best Local Similarity:    34.1%         Mismatches:      20
Query Match:              29.3%         Indels:          0
DB:                       13            Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AY112987 (1-11406)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||                ||:::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                     |||      |||   |||      |||   ||||||||:::    |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101

RESULT 67
ALSFV42S
LOCUS       ALSFV42S               11442 bp    RNA     linear   VRL 13-MAR-2001
DEFINITION  Semliki forest virus 42S RNA genome.
ACCESSION   X04129 J02361 J02362 L00018 V01400 V01401
VERSION     X04129.1  GI:58614
KEYWORDS    capsid protein; genome; nonstructural protein; polyprotein;
            structural protein.
SOURCE      Semliki forest virus
  ORGANISM  Semliki forest virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
            Alphavirus; SFV complex.
REFERENCE   1  (bases 7387 to 8335)
  AUTHORS   Garoff,H., Frischauf,A.M., Simons,K., Lehrach,H. and Delius,H.
  TITLE     The capsid protein of Semliki Forest virus has clusters of basic
            amino acids and prolines in its amino-terminal region
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 77 (11), 6376-6380 (1980)
   PUBMED   6935652
```

```
REFERENCE   2  (bases 8212 to 11442)
  AUTHORS   Garoff,H., Frischauf,A.M., Simons,K., Lehrach,H. and Delius,H.
  TITLE     Nucleotide sequence of cdna coding for Semliki Forest virus
            membrane glycoproteins
  JOURNAL   Nature 288 (5788), 236-241 (1980)
   PUBMED   6985476
REFERENCE   3  (bases 7121 to 7480)
  AUTHORS   Riedel,H., Lehrach,H. and Garoff,H.
  TITLE     Nucleotide sequence at the junction between the nonstructural and
            the structural genes of the semliki forest virus genome
  JOURNAL   J. Virol. 42 (2), 725-729 (1982)
   PUBMED   7086974
REFERENCE   4  (bases 1 to 7520)
  AUTHORS   Takkinen,K.
  TITLE     Complete nucleotide sequence of the nonstructural protein genes of
            Semliki Forest virus
  JOURNAL   Nucleic Acids Res. 14 (14), 5667-5682 (1986)
   PUBMED   3488539
COMMENT     On Jul 24, 2002 this sequence version replaced gi:334036.
            Data kindly reviewed (23-MAR-1987) by K. Takkinen.
FEATURES             Location/Qualifiers
     source          1. .11442
                     /organism="Semliki forest virus"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:11033"
     CDS             86. .7381
                     /note="(aa 1-2431)"
                     /codon_start=1
                     /product="polyprotein of non-structural proteins"
                     /protein_id="CAA27741.1"
                     /db_xref="GI:58615"
                     /db_xref="GOA:P08411"
                     /db_xref="UniProt/Swiss-Prot:P08411"
                     /translation="MAAKVHVDIEADSPFIKSLQKAFPSFEVESLQVTPNDHANARAF
                     SHLATKLIEQETDKDTLILDIGSAPSRRMMSTHKYHCVCPMRSAEDPERLDSYAKKLA
                     AASGKVLDREIAGKITDLQTVMATPDAESPTFCLHTDVTCRTAAEVAVYQDVYAVHAP
                     TSLYHQAMKGVRTAYWIGFDTTPFMFDALAGAYPTYATNWADEQVLQARNIGLCAASL
                     TEGRLGKLSILRKKQLKPCDTVMFSVGSTLYTESRKLLRSWHLPSVFHLKGKQSFTCR
                     CDTIVSCEGYVVKKITMCPGLYGKTVGYAVTYHAEGFLVCKTTDTVKGERVSFPVCTY
                     VPSTICDQMTGILATDVTPEDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPIVAVAF
                     SKWAREYKADLDDEKPLGVRERSLTCCCLWAFKTRKMHTMYKKPDTQTIVKVPSEFNS
                     FVIPSLWSTGLAIPVRSRIKMLLAKKTKRELIPVLDASSARDAEQEEKERLEAELTRE
                     ALPPLVPIAPAETGVVDVDVEELEYHAGAGVVETPRSALKVTAQPNDVLLGNYVVLSP
                     QTVLKSSKLAPVHPLAEQVKIITHNGRAGGYQVDGYDGRVLLPCGSAIPVPEFQALSE
                     SATMVYNEREFVNRKLYHIAVHGPSLNTDEENYEKVRAERTDAEYVFDVDKKCCVKRE
                     EASGLVLVGELTNPPFHEFAYEGLKIRPSAPYKTTVVGVFGVPGSGKSAIIKSLVTKH
                     DLVTSGKKENCQEIVNDVKKHRGKGTSRENSDSILLNGCRRAVDILYVDEAFACHSGT
                     LLALIALVKPRSKVVLCGDPKQCGFFNMMQLKVNFNHNICTEVCHKSISRRCTRPVTA
                     IVSTLHYGGKMRTTNPCNKPIIIDTTGQTKPKPGDIVLTCFRGWAKQLQLDYRGHEVM
                     TAAASQGLTRKGVYAVRQKVNENPLYAPASEHVNVLLTRTEDRLVWKTLAGDPWIKVL
                     SNIPQGNFTATLEEWQEEHDKIMKVIEGPAAPVDAFQNKANVCWAKSLVPVLDTAGIR
                     LTAEEWSTIITAFKEDRAYSPVVALNEICTKYYGVDLDSGLFSAPKVSLYYENNHWDN
                     RPGGRMYGFNAATAARLEARHTFLKGQWHTGKQAVIAERKIQPLSVLDNVIPINRRLP
                     HALVAEYKTVKGSRVEWLVNKVRGYHVLLVSEYNLALPRRRVTWLSPLNVTGADRCYD
                     LSLGLPADAGRFDLVFVNIHTEFRIHHYQQCVDHAMKLQMLGGDALRLLKPGGILMRA
                     YGYADKISEAVVSSLSRKFSSARVLRPDCVTSNTEVFLLFSNFDNGKRPSTLHQMNTK
                     LSAVYAGEAMHTAGCAPSYRVKRADIATCTEAAVVNAANARGTVGDGVCRAVAKKWPS
                     AFKGAATPVGTIKTVMCGSYPVIHAVAPNFSATTEAEGDRELAAVYRAVAAEVNRLSL
                     SSVAIPLLSTGVFSGGRDRLQQSLNHLFTAMDATDADVTIYCRDKSWEKKIQEAIDMR
                     TAVELLNDDVELTTDLVRVHPDSSLVGRKGYSTTDGSLYSYFEGTKFNQAAIDMAEIL
                     TLWPRLQEANERICLYALGETMDNIGSKCPVNDSDSSTPPRTVPCLCRYAMTAERIAR
                     LRSHQVKSMVVCSSFPLPKYHVDGVQKVKCEKVLLFDPTVPSVVSPRKYAASTTDHSD
                     RSLRGFDLDWTTDSSSTASDTMSLPSLQSCDIDSIYEPMAPIVVTADVHPEPAGIADL
                     AADVHPEPADHVDLENPIPPPRPKRAAYLASRAAERPVPAPRKPTPAPRTAFRNKLPL
                     TFGDFDEHEVDALASGITFGDFDDVLRLGRAGAYIFSSDTGSGHLQQKSVRQHNLQCA
                     QLDAVQEEKMYPPKLDTEREKLLLLKMQMHPSEANKSRYQSRKVENMKATVVDRLTSG
                     ARLYTGADVGRIPTYAVRYPRPVYSPTVIERFSSPDVAIAACNEYLSRNYPTVASYQI
                     TDEYDAYLDMVDGSDSCLDRATFCPAKLRCYPKHHAYHQPTVRSAVPSPFQNTLQNVL
                     AAATKRNCNVTQMRELPTMDSAVFNVECFKRYACSGEYWEEYAKQPIRITTENITTYV
                     TKLKGPKAAALFAKTHNLVPLQEVPMDRFTVDMKRDVKVTPGTKHTEERPKVQVIQAA
                     EPLATAYLCGIHRELVRRLNAVLRPNVHTLFDMSAEDFDAIIASHFHPGDPVLETDIA
                     SFDKSQDDSLALTGLMILEDLGVDQYLLDLIEAAFGEISSCHLPTGTRFKFGAMMKSG
                     MFLTLFINTVLNITIASRVLEQRLTDSACAAFIGDDNIVHGVISDKLMAERCASWVNM
                     EVKIIDAVMGEKPPYFCGGFIVFDSVTQTACRVSDPLKRLFKLGKPLTAEDKQDEDRR
                     RALSDEVSKWFRTGLGAELEVALTSRYEVEGCKSILIAMTTLARDIKAFKKLRGPVIH
                     LYGGPRLVR"
     mat_peptide     86. .1696
                     /product="nonstructural protein nsP1"
     mat_peptide     1697. .4090
                     /product="nonstructural protein nsP2"
     mat_peptide     4091. .5536
                     /product="nonstructural protein nsP3"
     mat_peptide     5537. .7378
                     /product="nonstructural protein nsP4"
     misc_feature    5225. .5248
                     /note="imperfect repeat 1"
     misc_feature    5267. .5290
                     /note="imperfect repeat 1"
     misc_feature    5378. .5389
                     /note="imperfect repeat 2"
     misc_feature    5399. .5410
                     /note="imperfect repeat 2"
     misc_feature    5438. .5455
                     /note="imperfect repeat 3"
     misc_feature    5489. .5506
                     /note="imperfect repeat 3"
     misc_feature    7369. .11442
                     /note="26S RNA"
     CDS             7420. .11181
                     /codon_start=1
                     /product="capsid protein"
                     /protein_id="CAA27742.1"
                     /db_xref="GI:1627514"
                     /db_xref="GOA:P03315"
                     /db_xref="UniProt/Swiss-Prot:P03315"
                     /translation="MNYIPTQTFYGRRWRPRPAARPWPLQATPVAPVVPDFQAQQMQQ
                     LISAVNALTMRQNAIAPARPPKPKKKKTTKPKPKTQPKKINGKTQQQKKKDKQADKKK
                     KKPGKRERMCMKIENDCIFEVKHEGKVTGYACLVGDKVMKPAHVKGVIDNADLAKLAF
                     KKSSKYDLECAQIPVHMRSDASKYTHEKPEGHYNWHHGAVQYSGGRFTIPTGAGKPGD
                     SGRPIFDNKGRVVAIVLGGANEGSRTALSVVTWNKDMVTRVTPEGSEEWSAPLITAMC
                     VLANATFPCFQPPCVPCCYENNAEATLRMLEDNVDRPGYYDLLQAALTCRNGTRHRRS
                     VSQHFNVYKATRPYIAYCADCGAGHSCHSPVAIEAVRSEATDGMLKIQFSAQIGIDKS
                     DNHDYTKIRYADGHAIENAVRSSLKVATSGDCPVHGTMGHFILAKCPPGEFLQVSIQD
                     TRNAVRACRIQYHHDPQPVGREKFTIRPHYGKEIPCTTYQQTTAETVEEIDMHMPPDT
                     PDRTLLSQQSGNVKITVGGKKVKYNCTCGTGNVGTTNSDMTINTCLIEQCHVSVTDHK
                     KWQFNSPFVPRADEPARKGKVHIPFPLDNITCRVPMAREPTVIHGKREVTLHLHPDHP
                     TLFSYRTLGEDPQYHEEWVTAAVERTIPVPVDGMEYHWGNNDPVRLWSQLTTEGKPHG
                     WPHQIVQYYYGLYPAATVSAVVGMSLLALISIFASCYMLVAARSKCLTPYALTPGAAV
                     PWTLGILCCAPRAHAASVAETMAYLWDQNQALFWLEFAAPVACILIITYCLRNVLCCC
                     KSLSFLVLLSLGATARAYEHSTVMPNVVGFPYKAHIERPGYSPLTLQMQVVETSLEPT
                     LNLEYITCEYKTVVPSPYVKCCGASECSTKEKPDYQCKVYTGVYPFMWGGAYCFCDSE
                     NTQLSEAYVDRSDVCRHDHASAYKAHTASLKAKVRVMYGNVNQTVDVYVNGDHAVTIG
                     GTQFIFGPLSSAWTPFDNKIVVYKDEVFNQDFPPYGSGQPGRFGDIQSRTVESNDLYA
                     NTALKLARPSPGMVHVPYTQTPSGFKYWLKEKGTALNTKAPFGCQIKTNPVRAMNCAV
                     GNIPVSMNLPDSAFTRIVEAPTIIDLTCTVATCTHSSDFGGVLTLTYKTNKNGDCSVH
                     SHSNVATLQEATAKVKTAGKVTLHFSTASASPSPVVSLCSARATCSASCEPPKDHIVP
                     YAASHSNVVFPDMSGTALSWVQKISGGLGAFAIGAILVLVVVTCIGLRR"

ORIGIN

Alignment Scores:
Pred. No.:              396             Length:         11442
Score:                  67.00           Matches:        14
Percent Similarity:     51.2%           Conservative:   7
Best Local Similarity:  34.1%           Mismatches:     20
Query Match:            29.3%           Indels:         0
DB:                     13              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x ALSFV42S (1-11442)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||      |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||     |||   |||      |||   |||||||||:::   |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097
```

```
Qy          41 Val 41
               |||
Db        3098 GTG 3100

RESULT 68
A18788
LOCUS       A18788                 11517 bp    RNA     linear   PAT 06-MAY-1994
DEFINITION  complete nucleotide sequence of pSP6-SFV4 RNA transcript.
ACCESSION   A18788
VERSION     A18788.1  GI:513416
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 11517)
  AUTHORS   .
  TITLE     DNA EXPRESSION SYSTEMS BASED ON ALPHAVIRUSES
  JOURNAL   Patent: WO 9210578-A 14 25-JUN-1992;
FEATURES             Location/Qualifiers
     source          1. .11517
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
     CDS             87. .7382
                     /codon_start=1
                     /transl_table=11
                     /product="non-structural polyprotein of pSP6-SFV4"
                     /protein_id="CAA01422.1"
                     /db_xref="GI:513417"
                     /translation="MAAKVHVDIEADSPFIKSLQKAFPSFEVESLQVTPNDHANARAF
                     SHLATKLIEQETDKDTLILDIGSAPSRRMMSTHKYHCVCPMRSAEDPERLDSYAKKLA
                     AASGKVLDREIAGKITDLQTVMATPDAESPTFCLHTDVTCRTAAEVAVYQDVYAVHAP
                     TSLYHQAMKGVRTAYWIGFDTTPFMFDALAGAYPTYATNWADEQVLQARNIGLCAASL
                     TEGRLGKLSILRKKQLKPCDTVMFSVGSTLYTESRKLLRSWHLPSVFHLKGKQSFTCR
                     CDTIVSCEGYVVKKITMCPGLYGKTVGYAVTYHAEGFLVCKTTDTVKGERVSFPVCTY
                     VPSTICDQMTGILATDVTPEDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPIVAVAF
                     SKWAREYKADLDDEKPLGVRERSLTCCCLWAFKTRKMHTMYKKPDTQTIVKVPSEFNS
                     FVIPSLWSTGLAIPVRSRIKMLLAKKTKRELIPVLDASSARDAEQEEKERLEAELTRE
                     ALPPLVPIAPAETGVVDVDVEELEYHAGAGVVETPRSALKVTAQPNDVLLGNYVVLSP
                     QTVLKSSKLAPVHPLAEQVKIITHNGRAGGYQVDGYDGRVLLPCGSAIPVPEFQALSE
                     SATMVYNEREFVNRKLYHIAVHGPSLNTDEENYEKVRAERTDAEYVFDVDKKCCVKRE
                     EASGLVLVGELTNPPFHEFAYEGLKIRPSAPYKTTVVGVFGVPGSGKSAIIKSLVTKH
                     DLVTSGKKENCQEIVNDVKKHRGKGTSRENSDSILLNGCRRAVDILYVDEAFACHSGT
                     LLALIALVKPRSKVVLCGDPKQCGFFNMMQLKVNFNHNICTEVCHKSISRRCTRPVTA
                     IVSTLHYGGKMRTTNPCNKPIIIDTTGQTKPKPGDIVLTCFRGWAKQLQLDYRGHEVM
                     TAAASQGLTRKGVYAVRQKVNENPLYAPASEHVNVLLTRTEDRLVWKTLAGDPWIKVL
                     SNIPQGNFTATLEEWQEEHDKIMKVIEGPAAPVDAFQNKANVCWAKSLVPVLDTAGIR
                     LTAEEWSTIITAFKEDRAYSPVVALNEICTKYYGVDLDSGLFSAPKVSLYYENNHWDN
                     RPGGRMYGFNAATAARLEARHTFLKGQWHTGKQAVIAERKIQPLSVLDNVIPINRRLP
                     HALVAEYKTVKGSRVEWLVNKVRGYHVLLVSEYNLALPRRRVTWLSPLNVTGADRCYD
                     LSLGLPADAGRFDLVFVNIHTEFRIHHYQQCVDHAMKLQMLGGDALRLLKPGGILMRA
                     YGYADKISEAVVSSLSRKFSSARVLRPDCVTSNTEVFLLFSNFDNGKRPSTLHQMNTK
                     LSAVYAGEAMHTAGCAPSYRVKRADIATCTEAAVVNAANARGTVGDGVCRAVAKKWPS
                     AFKGAATPVGTIKTVMCGSYPVIHAVAPNFSATTEAEGDRELAAVYRAVAAEVNRLSL
                     SSVAIPLLSTGVFSGGRDRLQQSLNHLFTAMDATDADVTIYCRDKSWEKKIQEAIDMR
                     TAVELLNDDVELTTDLVRVHPDSSLVGRKGYSTTDGSLYSYFEGTKFNQAAIDMAEIL
                     TLWPRLQEANEQICLYALGETMDNIRSKCPVNDSDSSTPPRTVPCLCRYAMTAERIAR
                     LRSHQVKSMVVCSSFPLPKYHVDGVQKVKCEKVLLFDPTVPSVVSPRKYAASTTDHSD
                     RSLRGFDLDWTTDSSSTASDTMSLPSLQSCDIDSIYEPMAPIVVTADVHPEPAGIADL
                     AADVHPEPADHVDLENPIPPPRPKRAAYLASRAAERPVPAPRKPTPAPRTAFRNKLPL
                     TFGDFDEHEVDALASGITFGDFDDVLRLGRAGAYIFSSDTGSGHLQQKSVRQHNLQCA
                     QLDAVQEEKMYPPKLDTEREKLLLLKMQMHPSEANKSRYQSRKVENMKATVVDRLTSG
                     ARLYTGADVGRIPTYAVRYPRPVYSPTVIERFSSPDVAIAACNEYLSRNYPTVASYQI
                     TDEYDAYLDMVDGSDSCLDRATFCPAKLRCYPKHHAYHQPTVRSAVPSPFQNTLQNVL
                     AAATKRNCNVTQMRELPTMDSAVFNVECFKRYACSGEYWEEYAKQPIRITTENITTYV
                     TKLKGPKAAALFAKTHNLVPLQEVPMDRFTVDMKRDVKVTPGTKHTEERPKVQVIQAA
                     EPLATAYLCGIHRELVRRLNAVLRPNVHTLFDMSAEDFDAIIASHFHPGDPVLETDIA
                     SFDKSQDDSLALTGLMILEDLGVDQYLLDLIEAAFGEISSCHLPTGTRFKFGAMMKSG
                     MFLTLFINTVLNITIASRVLEQRLTDSACAAFIGDDNIVHGVISDKLMAERCASWVNM
                     EVKIIDAVMGEKPPYFCGGFIVFDSVTQTACRVSDPLKRLFKLGKPLTAEDKQDEDRR
                     RALSDEVSKWFRTGLGAELEVALTSRYEVEGCKSILIAMTTLARDIKAFKKLRGPVIH
                     LYGGPRLVR"
```

```
     CDS             7421. .11182
                     /codon_start=1
                     /transl_table=11
                     /product="structural polyprotein of pSP6-SFV4"
                     /protein_id="CAA01423.1"
                     /db_xref="GI:513418"
                     /translation="MNYIPTQTFYGRRWRPRPAARPWPLQATPVAPVVPDFQAQQMQQ
                     LISAVNALTMRQNAIAPARPPKPKKKKTTKPKPKTQPKKINGKTQQQKKKDKQADKKK
                     KKPGKRERMCMKIENDCIFEVKHEGKVTGYACLVGDKVMKPAHVKGVIDNADLAKLAF
                     KKSSKYDLECAQIPVHMRSDASKYTHEKPEGHYNWHHGAVQYSGGRFTIPTGAGKPGD
                     SGRPIFDNKGRVVAIVLGGANEGSRTALSVVTWNKDMVTRVTPEGSEEWSAPLITAMC
                     VLANATFPCFQPPCVPCCYENNAEATLRMLEDNVDRPGYYDLLQAALTCRNGTRHRRS
                     VSQHFNVYKATRPYIAYCADCGAGHSCHSPVAIEAVRSEATDGMLKIQFSAQIGIDKS
                     DNHDYTKIRYADGHAIENAVRSSLKVATSGDCFVHGTMGHFILAKCPPGEFLQVSIQD
                     TRNAVRACRIQYHHDPQPVGREKFTIRPHYGKEIPCTTYQQTTAKTVEEIDMHMPPDT
                     PDRTLLSQQSGNVKITVGGKKVKYNCTCGTGNVGTTNSDMTINTCLIEQCHVSVTDHK
                     KWQFNSPFVPRADEPARKGKVHIPFPLDNITCRVPMAREPTVIHGKREVTLHLHPDHP
                     TLFSYRTLGEDPQYHEEWVTAAVERTIPVPVDGMEYHWGNNDPVRLWSQLTTEGKPHG
                     WPHQIVQYYYGLYPAATVSAVVGMSLLALISIFASCYMLVAARSKCLTPYALTPGAAV
                     PWTLGILCCAPRAHAASVAETMAYLWDQNQALFWLEFAAPVACILIITYCLRNVLCCC
                     KSLSFLVLLSLGATARAYEHSTVMPNVVGFPYKAHIERPGYSPLTLQMQVVETSLEPT
                     LNLEYITCEYKTVVPSPYVKCCGASECSTKEKPDYQCKVYTGVYPFMWGGAYCFCDSE
                     NTQLSEAYVDRSDVCRHDHASAYKAHTASLKAKVRVMYGNVNQTVDVYVNGDHAVTIG
                     GTQFIFGPLSSAWTPFDNKIVVYKDEVFNQDFPPYGSGQPGRFGDIQSRTVESNDLYA
                     NTALKLARPSPGMVHVPYTQTPSGFKYWLKEKGTALNTKAPFGCQIKTNPVRAMNCAV
                     GNIPVSMNLPDSAFTRIVEAPTIIDLTCTVATCTHSSDFGGVLTLTYKTNKNGDCSVH
                     SHSNVATLQEATAKVKTAGKVTLHFSTASASPSFVVSLCSARATCSASCEPPKDHIVP
                     YAASHSNVVFPDMSGTALSWVQKISGGLGAFAIGAILVLVVVTCIGLRR"
ORIGIN

Alignment Scores:
Pred. No.:              399             Length:         11517
Score:                  67.00           Matches:        14
Percent Similarity:     51.2%           Conservative:   7
Best Local Similarity:  34.1%           Mismatches:     20
Query Match:            29.3%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x A18788 (1-11517)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               |||||||||      :::::::::   ::::::|||                |||:::
Db        2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||     |||   |||     |||   |||||||||:::   |||
Db        3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy          41 Val 41
               |||
Db        3099 GTG 3101

RESULT 69
AR001293
LOCUS       AR001293               11517 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 1 from patent US 5739026.
ACCESSION   AR001293
VERSION     AR001293.1  GI:3963360
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11517)
  AUTHORS   Garoff,H. and Liljestrom,P.
  TITLE     DNA expression systems based on alphaviruses
  JOURNAL   Patent: US 5739026-A 1 14-APR-1998;
FEATURES             Location/Qualifiers
     source          1. .11517
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
```

```
Pred. No.:              399             Length:         11517
Score:                  67.00           Matches:        14
Percent Similarity:     51.2%           Conservative:   7
Best Local Similarity:  34.1%           Mismatches:     20
Query Match:            29.3%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AR001293 (1-11517)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::::||                ||||:::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||       |||   |||      |||   |||||||||:::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101

RESULT 70
AR130391
LOCUS       AR130391               11517 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6190666.
ACCESSION   AR130391
VERSION     AR130391.1  GI:14118716
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11517)
  AUTHORS   Garoff,H. and Liljestrom,P.
  TITLE     DNA expression systems based on alphaviruses
  JOURNAL   Patent: US 6190666-A 1 20-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .11517
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              399             Length:         11517
Score:                  67.00           Matches:        14
Percent Similarity:     51.2%           Conservative:   7
Best Local Similarity:  34.1%           Mismatches:     20
Query Match:            29.3%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AR130391 (1-11517)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::|||                ||||:::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||       |||   |||      |||   |||||||||:::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101

RESULT 71
AR569539
LOCUS       AR569539               11517 bp    RNA     linear   PAT 14-DEC-2004
DEFINITION  Sequence 1 from patent US 6770283.
ACCESSION   AR569539
VERSION     AR569539.1  GI:56570146
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11517)
  AUTHORS   Garoff,H. and Liljestrom,P.
  TITLE     DNA expression systems based on alphaviruses
  JOURNAL   Patent: US 6770283-A 1 03-AUG-2004;
            Bioption AB; Tullinge;
            SEX;
FEATURES             Location/Qualifiers
     source          1. .11517
                     /organism="unknown"
                     /mol_type="genomic RNA"
ORIGIN

Alignment Scores:
Pred. No.:              399             Length:         11517
Score:                  67.00           Matches:        14
Percent Similarity:     51.2%           Conservative:   7
Best Local Similarity:  34.1%           Mismatches:     20
Query Match:            29.3%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AR569539 (1-11517)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   :::::|||                ||||:::
Db       2979 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3038

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||       |||   |||      |||   |||||||||:::   |||
Db       3039 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3098

Qy         41 Val 41
              |||
Db       3099 GTG 3101

RESULT 72
BD263357
LOCUS       BD263357               12464 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compositions and methods for use in recombinational cloning of
            nucleic acids.
ACCESSION   BD263357
VERSION     BD263357.1  GI:33073125
KEYWORDS    JP 2002537790-A/135.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 12464)
  AUTHORS   Hartley,J.L., Brasch,M.A., Temple,G.F. and Cheo,D.
  TITLE     Compositions and methods for use in recombinational cloning of
            nucleic acids
  JOURNAL   Patent: JP 2002537790-A 135 12-NOV-2002;
            INVITROGEN CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002537790-A/135
            PD   12-NOV-2002
            PF   02-MAR-2000 JP 2000602252
            PR   02-MAR-1999 US    60/122389,23-MAR-1999 US    60/126049 PR
            28-MAY-1999 US    60/136744
            PI   JAMES L HARTLEY,MICHAEL A BRASCH,GARY F TEMPLE,DAVID CHEO PC
            C12N15/09,C07K14/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/ PC
             00,C12N5/00
            CC   pDEST9
            CC   attR1
            CC   CmR
            CC   inactivated ccdA
            CC   ccdB
            CC   attR2
            CC   ori
            CC   ampR
            CC   SP6 promoter
            CC   nsP1:non-structural protein 1
            CC   nsP2:non-structural protein 2
            CC   nsP3:non-structural protein 3
            CC   nsP4:non-structural protein 4
```

```
            FH   Key             Location/Qualifiers
            FT   gene            (232). .(355)
            FT   gene            (605). .(1264)
            FT   gene            (1384). .(1468)
            FT   gene            (1606). .(1911)
            FT   gene            (1952). .(2078)
            FT   gene            (2532). .(2782)
            FT   gene            (3482). .(4282)
            FT   gene            (5232). .(5365)
            FT   gene            (5365). .(6965)
            FT   gene            (6965). .(9265)
            FT   gene            (9265). .(10865)
            FT   gene            (161). .(10865).
FEATURES             Location/Qualifiers
     source          1. .12464
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Alignment Scores:
Pred. No.:                433           Length:         12464
Score:                    67.00         Matches:        14
Percent Similarity:       51.2%         Conservative:   7
Best Local Similarity:    34.1%         Mismatches:     20
Query Match:              29.3%         Indels:         0
DB:                       6             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BD263357 (1-12464)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||              |||:::
Db       8244 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 8303

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                     |||      |||   |||      |||   ||||||||||:::   |||
Db       8304 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 8363

Qy         41 Val 41
              |||
Db       8364 GTG 8366

RESULT 73
SFVRNAIS
LOCUS       SFVRNAIS               14529 bp    RNA     linear   VRL 18-APR-2005
DEFINITION  Semliki forest virus genomic RNA for non-structural polyprotein and
            structural polyprotein, strain A7.
ACCESSION   Z48163 AY242846
VERSION     Z48163.2  GI:30519795
KEYWORDS    non-structural polyprotein; structural polyprotein.
SOURCE      Semliki forest virus
  ORGANISM  Semliki forest virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
            Alphavirus; SFV complex.
REFERENCE   1  (bases 1 to 14529)
  AUTHORS   Tarbatt,C.J., Glasgow,G.M., Mooney,D.A., Sheahan,B.J. and
            Atkins,G.J.
  TITLE     Sequence analysis of the avirulent, demyelinating A7 strain of
            Semliki Forest virus
  JOURNAL   J. Gen. Virol. 78 (PT 7), 1551-1557 (1997)
   PUBMED   9225028
REFERENCE   2
  AUTHORS   Tarbatt,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1995) Catherine J Tarbatt BSc., Microbiology, The
            Moyne Institute, Trinity College, Dublin 2, Ireland
  REMARK    revised by [3]
REFERENCE   3  (bases 1 to 14529)
  AUTHORS   Logue,C.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2003) Christopher H. Logue, Virology, The Moyne
            Institute, Trinity College, Dublin 2, Ireland
COMMENT     On May 10, 2003 this sequence version replaced gi:663260.
```

```
FEATURES             Location/Qualifiers
     source          1. .14529
                     /organism="Semliki forest virus"
                     /mol_type="genomic RNA"
                     /strain="A7"
                     /db_xref="taxon:11033"
                     /note="infectious clone"
     CDS             86. .5500
                     /codon_start=1
                     /product="non-structural polyprotein"
                     /protein_id="CAD90833.1"
                     /db_xref="GI:30519796"
                     /db_xref="GOA:Q809B7"
                     /db_xref="InterPro:IPR000606"
                     /db_xref="InterPro:IPR002589"
                     /db_xref="InterPro:IPR002620"
                     /db_xref="InterPro:IPR010916"
                     /db_xref="UniProt/TrEMBL:Q809B7"
                     /translation="MAAKVHVDIEADSPFIKSLQKAFPSFEVESLQVTPNDHANARAF
                     SHLATKLIEQETDKDTLILDIGSAPSRRMMSTHKYHCVCPMRSAEDPERLVCYAKKLA
                     AASGKVLDREIAGKITDLQTVMATPDAESPTFCLHTDVTCRTAAEVAVYQDVYAVHAP
                     TSLYHQAMKGVRTAYWIGFDTTPFMFDALAGAYPTYATNWADEQVLQARNIGLCAASL
                     TEGRLGKLSILRKKQLKPSDTVMFSVGSTLYTESRKLLRSWHLPSVFHLKGKQSFTCR
                     CDTIVSCEGYVVKKITMCPGLYGKTVGYAVTHHAEGFLVCKTTDTVKGERVSFPVCTY
                     VPSTICDQMTGILATDVTPEDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPVVAVAF
                     SKWAREYKADLDDEKPLGVRERSLTCCCLWAFKTKKMHTMYKKPDTQTIVKVPSEFNS
                     FVIPSLWSTGLAIPVRSRIKMLLAKKTKRELIPALDASSARDAEQEEKERLEAELTRE
                     ALPPLVPIAPAETGVVDVDVEELEYRAGAGVVETPRSALKVTAQPNDVLLGNYVVLSP
                     QTVLKSSKLAPVHPLAEQVKIITHNGRAGRYQVDGYDGRVLLPCGSAIPVPEFQALSE
                     SATMVYNEREFVNRKLYHIAVHGPSLNTDEENYEKVRAERTDAEYVFDVDKKCCIKRE
                     EASGLVLVGELTNPPFHEFAYEGLKIRPSAPYKTTVVGVFGVPGSGKSAIIKSLVTKH
                     DLVTSGKKENCQEIVNDVKKHRGLDIQAKTVDSILLNGCRRAVDILYVDEAFACHSGT
                     LLALIALVKPRSKVVLCGDPKQCGFFNMMQLKVNFNHNICTEVCHKSISRRCTRPVTA
                     IVSTLHYGGKMRTTNPCNKPIIIDTTGQTKPKPGDIVLTCFRGWVKQLQLDYRGHEVM
                     TAAASQGLTRKGVYAVRQKVNENPLYAPASEHVNVLLTRTEDRLVWKTLAGDPWIKVL
                     SNIPQGNFTATLEEWQEEHDKIMKVIEGPAAPVDAFQNKANVCWAKSLVPVLDTAGIR
                     LTAEEWSTIITAFKEDRAYSPEVALNEICTKYYGVDLDSGLFSAPKVSLYYENNHWDN
                     RPGGRMYGFNAATAARLEARHTFLKGQWHTGKQAVIAERKIQPLSVLDNVIPINRRLP
                     HALVAEYKTVKGSRVEWLVNKVRGYHVLLVSEYNLALPRRRVTWLSPLNVTGADRCYD
                     LGLGLPADAGRYDLVFVNIHTEFRIHHYQQCVDHAMKLQMLGGDALRLLKPGGNLLMR
                     AYGYADKISEAVVSSLSRKFSSARVLRPDCVTSNTEVFLLFSNFDNGKRPSTLHQMNT
                     KLSAVYAGEAMHTAGCAPSYRVKRADIATCTEAAVVNAANARGTVGDGVCRAVAKKWP
                     SAFKGAATPVGTIKTVMCGSYPVIHAVGPNFSATTEAEGDRELAAAVYRAVAAEVNRLS
                     LSSVAIPLLSTGVFSGGRDRLQQSLNHLFTAMDATDADVTIYCRDKSWEKKIQEAIDM
                     RTAVELLNDDVELTTDLVRVHPDSSLVGRKGYSTTDGSLYSYLEGTKFNQAAIDMAEI
                     LTLWPRLQEANEQICLYALGETMDNIRSKCPVDDSDSSTPPRTVPCLCRYAMTAERIA
                     RLRSHQVKSMVVCSSFPLPKYHVDGVQKVKCEKVLLFDPTVPSVVSPRKYAASTTDHS
                     DRSLRGFDLDWTTDSSSTASDTMSLPSLQSCDIDSIYEPMAPIVVTADVHPEPAAVHP
                     EPADHVDLENPIPPPRPKRAAYLASRAAERPVPAPRKPTPAPRTTFRNKLPFTFGDFD
                     EHEVDALASGITFGDFDDVL"
     CDS             7402. .11163
                     /codon_start=1
                     /product="structural polyprotein"
                     /protein_id="CAD90834.1"
                     /db_xref="GI:30519797"
                     /db_xref="GOA:Q809B6"
                     /db_xref="InterPro:IPR000930"
                     /db_xref="InterPro:IPR000936"
                     /db_xref="InterPro:IPR002533"
                     /db_xref="InterPro:IPR002548"
                     /db_xref="UniProt/TrEMBL:Q809B6"
                     /translation="MNYIPTQTFYGRRWRPRPAARPWPLQATPVAPVVPDFQAQQMQQ
                     LISAVNALTMRQNAIAPTRPPKPKKKKTTKPKPKTQPKKIKGKTQQQKKKDKQADKKK
                     KKPGKRERMCMKIENDCIFEVKHEGKVTGYACLVGDKVMKPAHVKGVIDNADLAKLAF
                     KKSSKYDLECAQIPVHMRSDASKYTHEKPEGHYNWHHGAVQYSGGRFTIPTGAGKPGD
                     SGRPIFDNKGRVVAIVLGGANEGSRTALSVVTWNKDMVTRVTPEGSEEWSAPLITAMC
                     VLTNATFPCFQPPCAPCCYENNAEATLRMLEDNVDRPGYYDLLQAALTCRNGTRHRRS
                     VSQHFNVYKATRPYIAYCADCGAGHSCHSPVAIEAIRSEATDGMLKIQFSAQIGIDKS
                     DNHDYTKIRYADGHAIENAVRSSLKVATSGDCFVHGTMGHFILATCPPGEFLQVSIQD
                     TRNAVRACRIQYHHDPQPVGREKFTIRPHYGKEIPCTTYQQTTAETVEEIDMHMPPDT
                     PDRTLLSQQSGNVKITVGGKKVKYNCTCGTGNVGTTSSDKTINTCLIEQCHVSVTDHK
                     KWQFNSPFVPRADEPARKGKVHIPFPLDNITCRVPMAREPTVIHGKREVTLHLHPDHP
                     TLFSYRTLGEDPQYHEEWVTAAVERTIPVPVDGMEYHWGNNDPVRLWSQLTTEGKPHG
```

WPHQIVQYYYGLYPAATASAVAGMSLLALISIFASCYMLAAARSKCLTPYALTPGAAV
PWTLGILCCAPRAHAASVAETMAYLWDQNQALFWLEFAAPVACILIITYCLRNVLCCC
KSLSFLVLLSLGATARAYEHSTVMPNVVGFPYKAHIERPGYSPLTLQMQVVETSLEPT
LNLEYITCEYKTVVPSPYVKCCGSSECSTKEKPDYQCKVYTGVYPFMWGGAYCFCDSE
NTQLSEAYVDRSDVCKHDHASAYKAHTASLKAKVRVMYGNVNQTVDVYVNGDHAVTIG
GTQFIFGPLSSAWTPFDNKIVVYKDEVFNQDFPPYGSGQPGRFGDIQSRTVESNDLYA
NTALKLARPSPGTVHVPYTQTPSGFKYWLKEKGTALNTKAPFGCQIKTNPVRAMNCAV
GNIPVSMNLPDSAFTRIVEAPTITDLTCTVATCTHSSDFGGVLTLKYKTDKNGDCSVH
SHSNVATLQEATAKVKTAGKVTLHFSTASASPSFVVSLCSAKATCSASCEPPKDHIVP
YAASHSNVVFPDMSGTALSWVQKISGGLGAFAIGAILVLVVVTCIGLRR"

ORIGIN

Alignment Scores:
Pred. No.: 507 Length: 14529
Score: 67.00 Matches: 14
Percent Similarity: 51.2% Conservative: 7
Best Local Similarity: 34.1% Mismatches: 20
Query Match: 29.3% Indels: 0
DB: 13 Gaps: 0

US-10-757-745-6_COPY_2_45 (1-44) x SFVRNAIS (1-14529)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||                |||:::
Db       2978 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 3037

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||     |||  |||     |||   |||||||||:::   |||
Db       3038 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 3097

Qy         41 Val 41
              |||
Db       3098 GTG 3100
```

RESULT 74
BD226816
LOCUS BD226816 15418 bp DNA linear PAT 17-JUL-2003
DEFINITION Alphavirus vectors for paramyxovirus vaccines.
ACCESSION BD226816
VERSION BD226816.1 GI:33036586
KEYWORDS JP 2002512003-A/1.
SOURCE Semliki forest virus
ORGANISM Semliki forest virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
Alphavirus; SFV complex.
REFERENCE 1 (bases 1 to 15418)
AUTHORS Parrington,M. and Li,X.
TITLE Alphavirus vectors for paramyxovirus vaccines
JOURNAL Patent: JP 2002512003-A 1 23-APR-2002;
CONNAUGHT LABORATORIES LTD
COMMENT OS Semliki forest virus (SFV)
PN JP 2002512003-A/1
PD 23-APR-2002
PF 13-NOV-1998 JP 2000521221
PR 14-NOV-1997 US 60/065791
PI MARK PARRINGTON,XIAOMAO LI
PC C12N15/09,A61K31/711,A61K39/155,A61P31/12,C07K14/115,C07K14/
PC 135,C12N15/00
CC Alphavirus vectors for paramyxovirus vaccines FH Key
Location/Qualifiers
FT source 1. .15538
FT /organism='Semliki forest virus (SFV)'.
FEATURES Location/Qualifiers
source 1. .15418
/organism="Semliki forest virus"
/mol_type="genomic DNA"
/db_xref="taxon:11033"

ORIGIN

Alignment Scores:
Pred. No.: 540 Length: 15418
Score: 67.00 Matches: 14
Percent Similarity: 51.2% Conservative: 7
Best Local Similarity: 34.1% Mismatches: 20
Query Match: 29.3% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-6_COPY_2_45 (1-44) x BD226816 (1-15418)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||                |||:::
Db       5525 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 5584

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||     |||  |||     |||   |||||||||:::   |||
Db       5585 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 5644

Qy         41 Val 41
              |||
Db       5645 GTG 5647
```

RESULT 75
AR243299
LOCUS AR243299 15538 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6475780.
ACCESSION AR243299
VERSION AR243299.1 GI:27290489
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 15538)
AUTHORS Parrington,M., Li,X. and Klein,M.H.
TITLE Alphavirus vectors for paramyxovirus vaccines
JOURNAL Patent: US 6475780-A 1 05-NOV-2002;
Aventis Pasteur Limited; Toronto;
CAX;
FEATURES Location/Qualifiers
source 1. .15538
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 544 Length: 15538
Score: 67.00 Matches: 14
Percent Similarity: 51.2% Conservative: 7
Best Local Similarity: 34.1% Mismatches: 20
Query Match: 29.3% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-6_COPY_2_45 (1-44) x AR243299 (1-15538)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||      :::::::::   ::::::|||                |||:::
Db       5525 GGCGATCCCTGGATTAAGGTCCTATCAAACATTCCACAGGGTAACTTTACGGCCACATTG 5584

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||     |||  |||     |||   |||||||||:::   |||
Db       5585 GAAGAATGGCAAGAAGAACACGACAAAATAATGAAGGTGATTGAAGGACCGGCTGCGCCT 5644

Qy         41 Val 41
              |||
Db       5645 GTG 5647
```

Search completed: March 10, 2006, 23:18:55
Job time : 1728.85 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:40:54 ; Search time 11935.2 Seconds
(without alignments)
1724.091 Million cell updates/sec

Title: US-10-757-745-2
Perfect score: 1902
Sequence: 1 MELGSCLEGGREAAEEEGEP.....DCGRFPSDHWGLLCNLDIIL 362

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170327_4011/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_5954_@runat_09032006_170327_4011 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1902	100.0	1261	8	AF201687	AF201687 Homo sapi
2	1902	100.0	1920	6	BD205490	BD205490 CD40-Inte
3	1902	100.0	1920	6	AR594294	AR594294 Sequence
4	1902	100.0	1920	6	AX011599	AX011599 Sequence
5	1902	100.0	1921	8	HSA269473	AJ269473 Homo sapi
6	1902	100.0	1948	6	AR339398	AR339398 Sequence
7	1902	100.0	1952	8	BC017553	BC017553 Homo sapi
8	1902	100.0	2499	6	E23195	E23195 Topoisomera
9	1897	99.7	1936	8	AF223469	AF223469 Homo sapi
10	1893	99.5	1898	6	BD157138	BD157138 Primer fo
11	1893	99.5	1898	6	AX878304	AX878304 Sequence
12	1893	99.5	1898	8	AK002168	AK002168 Homo sapi
c 13	1590	83.6	150344	14	AC152027	AC152027 Dasypus n
14	1464.5	77.0	2151	8	HSA420495	AJ420495 Homo sapi
15	1452.5	76.4	858	6	CQ726600	CQ726600 Sequence
16	1258	66.1	1312	6	BD205491	BD205491 CD40-Inte
17	1258	66.1	1312	6	AR594295	AR594295 Sequence
18	1258	66.1	1312	6	AX011601	AX011601 Sequence
19	1258	66.1	1943	9	MMU251328	AJ251328 Mus muscu
20	1188.5	62.5	1236	5	CR761829	CR761829 Xenopus t
21	1182	62.1	752	6	BD150065	BD150065 Primer fo
22	1182	62.1	752	6	AX870003	AX870003 Sequence
c 23	1177	61.9	1079	6	BD139846	BD139846 Compounds
c 24	1177	61.9	1079	6	AR202950	AR202950 Sequence
c 25	1177	61.9	1079	6	AR208051	AR208051 Sequence
c 26	1177	61.9	1079	6	AX429924	AX429924 Sequence
27	946	49.7	1451	5	BC083404	BC083404 Danio rer
28	923	48.5	1455	5	BC097117	BC097117 Danio rer
29	837	44.0	483	6	BD076937	BD076937 5' EST of
30	590	31.0	400	6	CQ663156	CQ663156 Sequence
31	536	28.2	256608	9	AL589699	AL589699 Mouse DNA
32	518	27.2	20938	8	AY613922	AY613922 Homo sapi
33	518	27.2	102200	8	HS30M3	AL031775 Human DNA
34	518	27.2	167847	8	CR925830	CR925830 Human DNA
35	518	27.2	176819	8	CR942205	CR942205 Human DNA
c 36	454	23.9	245394	14	AC125756	AC125756 Rattus no
37	376	19.8	258	6	BD077628	BD077628 5'EST of
38	376	19.8	258	6	AX939201	AX939201 Sequence
39	333	17.5	201	10	BV203236	BV203236 sqnm21148
40	298	15.7	176	6	AX351374	AX351374 Sequence
c 41	298	15.7	179	6	AX339803	AX339803 Sequence
c 42	280.5	14.7	411	10	HS133N16S	AL110544 H.sapiens
43	277.5	14.6	198752	5	BX957279	BX957279 Zebrafish
44	276.5	14.5	200076	5	BX511258	BX511258 Zebrafish
c 45	265.5	14.0	110000	14	AP006486_2	Continuation (3 of
46	231	12.1	1326	15	BT000599	BT000599 Arabidops
47	231	12.1	1557	15	AY085643	AY085643 Arabidops
48	231	12.1	1582	15	AY128286	AY128286 Arabidops
c 49	223.5	11.8	110000	14	CR954207_1	Continuation (2 of
c 50	221	11.6	268	6	AR240556	AR240556 Sequence
51	186.5	9.8	91566	15	F25C20	AC007296 Arabidops
52	182.5	9.6	444	6	CQ069444	CQ069444 Sequence
53	182.5	9.6	691	6	CQ078617	CQ078617 Sequence
c 54	181	9.5	33899	2	CEY63D3A	AL032652 Caenorhab
55	165	8.7	2036	2	AK113716	AK113716 Ciona int
56	158	8.3	110000	15	AP008218_128	Continuation (129
c 57	158	8.3	111678	15	AC149575	AC149575 Populus t
c 58	158	8.3	156724	15	CNS08CB2	AL844879 Oryza sat
59	154	8.1	86100	15	AC151708	AC151708 Medicago
c 60	154	8.1	133028	15	AC134967	AC134967 Medicago
61	148	7.8	2709	2	AY228171	AY228171 Trypanoso
c 62	148	7.8	103129	15	AP004494	AP004494 Lotus cor
c 63	148	7.8	231985	14	AC156673	AC156673 Bos tauru
64	147	7.7	9737	9	MMU70139	U70139 Mus musculu
65	146	7.7	156973	2	AC159426	AC159426 Trypanoso
66	145	7.6	379	6	BD112459	BD112459 EST and e
67	145	7.6	379	6	AR416906	AR416906 Sequence
68	145	7.6	379	6	AX977600	AX977600 Sequence
69	140	7.4	2561	9	AF199491	AF199491 Mus muscu
70	140	7.4	3037	9	AF183960	AF183960 Mus muscu
71	139.5	7.3	2745	2	AY437500	AY437500 Trypanoso
72	139	7.3	2778	2	AY437501	AY437501 Trypanoso
73	139	7.3	152420	14	BX927200	BX927200 Danio rer
74	139	7.3	187464	5	AL929007	AL929007 Zebrafish
c 75	139	7.3	208135	5	BX004778	BX004778 Zebrafish

ALIGNMENTS

```
RESULT 1
AF201687
LOCUS       AF201687                1261 bp    mRNA    linear   PRI 16-MAY-2003
DEFINITION  Homo sapiens ETS1-associated protein 2 (EAP2) mRNA, complete cds.
ACCESSION   AF201687
VERSION     AF201687.1  GI:11493668
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1261)
  AUTHORS   Pei,H., Yordy,J.S., Leng,Q., Zhao,Q., Watson,D.K. and Li,R.
  TITLE     EAPII interacts with ETS1 and modulates its transcriptional
            function
  JOURNAL   Oncogene 22 (18), 2699-2709 (2003)
   PUBMED   12743594
REFERENCE   2  (bases 1 to 1261)
  AUTHORS   Li,R., Pei,H. and Papas,T.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1999) Center for Molecular and Structural
            Biology, Department of Medicine, and Hollings Cancer Center,
            Medical University of South Carolina, 86 Jonathan Lucas St.,
            Charleston, SC 29425, USA
FEATURES             Location/Qualifiers
     source          1. .1261
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22.1-p22.3"
                     /cell_line="Jurkat"
     gene            1. .1261
                     /gene="EAP2"
     CDS             26. .1114
                     /gene="EAP2"
                     /codon_start=1
                     /product="ETS1-associated protein 2"
                     /protein_id="AAG35600.1"
                     /db_xref="GI:11493669"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                4.16e-163       Length:          1261
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2 (1-362) x AF201687 (1-1261)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
              ||||||
Db       1106 ATATTG 1111

RESULT 2
BD205490
LOCUS       BD205490                1920 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  CD40-Interacting and TRAF-interacting protein.
ACCESSION   BD205490
VERSION     BD205490.1  GI:33015260
KEYWORDS    JP 2002512796-A/1.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1920)
  AUTHORS   Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting protein
  JOURNAL   Patent: JP 2002512796-A 1 08-MAY-2002;
            VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002512796-A/1
            PD   08-MAY-2002
            PF   28-APR-1999 JP 2000546003
            PR   29-APR-1998 EP   98201392.2
            PI   STEFAN MARIA CHRISTIAAN PYPE,
            PI   JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI   DANNY
           FRANCOIS EVELINE HUYLEBROECK
            PC   C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
             A61P37/02,
            PC   A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
            PC   G01N33/566,
            PC   C12N15/00,A61K37/02
            CC   CD40-Interacting and TRAF-interacting protein. FH   Key
                 Location/Qualifiers
            FT   CDS             (20). .(1108).
FEATURES                Location/Qualifiers
     source             1. .1920
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              6.81e-163       Length:         1920
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2 (1-362) x BD205490 (1-1920)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 79

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 139

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 199

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 259

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 319

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 379

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1099

Qy        361 IleLeu 362
              ||||||
Db       1100 ATATTG 1105

RESULT 3
AR594294
LOCUS       AR594294                1920 bp    DNA     linear   PAT 15-DEC-2004
DEFINITION  Sequence 1 from patent US 6812203.
ACCESSION   AR594294
VERSION     AR594294.1  GI:56643900
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1920)
  AUTHORS   Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting proteins
  JOURNAL   Patent: US 6812203-A 1 02-NOV-2004;
            Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
            Zwijnaarde;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              6.81e-163       Length:         1920
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
```

```
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2 (1-362) x AR594294 (1-1920)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 79

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 139

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 199

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 259

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 319

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 379

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1099

Qy        361 IleLeu 362
              ||||||
Db       1100 ATATTG 1105

RESULT 4
AX011599
LOCUS       AX011599                1920 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 1 from Patent WO9955859.
ACCESSION   AX011599
VERSION     AX011599.1  GI:9998123
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 1 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             20. .1108
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07555.1"
                     /db_xref="GI:9998124"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:              6.81e-163       Length:         1920
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2 (1-362) x AX011599 (1-1920)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 79

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 139

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 199

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 259
```

```
Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 319

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 379

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1099

Qy        361 IleLeu 362
              ||||||
Db       1100 ATATTG 1105
```

```
RESULT 5
HSA269473
LOCUS       HSA269473              1921 bp    mRNA    linear   PRI 27-MAY-2000
DEFINITION  Homo sapiens mRNA for TRAF and TNF receptor associated protein
            (ttrap gene).
ACCESSION   AJ269473
VERSION     AJ269473.1  GI:8247253
KEYWORDS    TRAF and TNF receptor associated protein; ttrap gene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van
            Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,
            Huylebroeck,D. and Remacle,J.E.
  TITLE     TTRAP, a novel protein that associates with CD40, tumor necrosis
            factor (TNF) receptor-75 and TNF receptor-associated factors
            (TRAFs), and that inhibits nuclear factor-kappa B activation
  JOURNAL   J. Biol. Chem. 275 (24), 18586-18593 (2000)
   PUBMED   10764746
REFERENCE   2  (bases 1 to 1921)
  AUTHORS   Pype,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1999) Pype S., VIB07, Dept. Cell Growth,
            Differentiation and Development, Flanders Interuniversity Institute
            for Biotechnology, CELGEN, K.U.Leuven, Gasthuisberg Campus,
            Herestraat 49, B-3000 Leuven, BELGIUM
COMMENT     Related sequence: AL031775.
FEATURES             Location/Qualifiers
     source          1. .1921
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .1921
                     /gene="ttrap"
     5'UTR           <1. .21
                     /gene="ttrap"
     CDS             22. .1110
                     /gene="ttrap"
                     /function="putative role in signal transduction"
                     /codon_start=1
                     /product="TRAF and TNF receptor associated protein"
                     /protein_id="CAB92966.1"
                     /db_xref="GI:8247254"
                     /db_xref="GOA:O95551"
                     /db_xref="UniProt/TrEMBL:O95551"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
     3'UTR           1111. .1921
                     /gene="ttrap"
     polyA_signal    1900. .1905
                     /gene="ttrap"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                6.81e-163       Length:          1921
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2 (1-362) x HSA269473 (1-1921)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 81

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 141

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 201

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 261
```

```
Qy        81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
Db       262 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 321

Qy       101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
Db       322 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 381

Qy       121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
Db       382 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 441

Qy       141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
Db       442 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 501

Qy       161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
Db       502 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 561

Qy       181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
Db       562 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 621

Qy       201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
Db       622 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 681

Qy       221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
Db       682 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 741

Qy       241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
Db       742 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 801

Qy       261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
Db       802 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 861

Qy       281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
Db       862 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 921

Qy       301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
Db       922 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 981

Qy       321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
Db       982 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1041

Qy       341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
Db      1042 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1101

Qy       361 IleLeu 362
Db      1102 ATATTG 1107
```

```
RESULT 6
AR339398
LOCUS       AR339398                1948 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 889 from patent US 6569662.
ACCESSION   AR339398
VERSION     AR339398.1  GI:33726255
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Tang,Y.T., Zhou,P. and Drmanac,R.T.
  TITLE     Nucleic acids and polypeptides
  JOURNAL   Patent: US 6569662-A 889 27-MAY-2003;
            Hyseq, Inc.; Sunnyvale, CA
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                  6.93e-163       Length:          1948
Score:                      1902.00         Matches:         362
Percent Similarity:         100.0%          Conservative:    0
Best Local Similarity:      100.0%          Mismatches:      0
Query Match:                100.0%          Indels:          0
DB:                         6               Gaps:            0

US-10-757-745-2 (1-362) x AR339398 (1-1948)

Qy         1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
Db        39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy        21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
Db        99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy        41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
Db       159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy        61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
Db       219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy        81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
Db       279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy       101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
Db       339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy       121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
Db       399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy       141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
Db       459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy       161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
Db       519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy       181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
Db       579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy       201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
Db       639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy       221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
Db       699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy       241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
Db       759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy       261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
```

```
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124
```

RESULT 7
BC017553

```
LOCUS       BC017553                1952 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein, mRNA (cDNA
            clone MGC:9099 IMAGE:3920790), complete cds.
ACCESSION   BC017553
VERSION     BC017553.2  GI:34782842
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1952)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1952)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Sep 16, 2003 this sequence version replaced gi:17028464.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 15 Row: m Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 23510347.
FEATURES             Location/Qualifiers
     source          1. .1952
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:9099 IMAGE:3920790"
                     /tissue_type="Skin, melanotic melanoma."
                     /clone_lib="NIH_MGC_72"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1952
                     /gene="TTRAP"
                     /note="synonyms: AD022, EAP2, MGC9099, dJ30M3.3"
                     /db_xref="GeneID:51567"
                     /db_xref="MIM:605764"
     CDS             38. .1126
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TRAF and TNF receptor-associated protein"
                     /protein_id="AAH17553.1"
                     /db_xref="GI:17028465"
                     /db_xref="GeneID:51567"
                     /db_xref="MIM:605764"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMPSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                6.94e-163       Length:          1952
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2 (1-362) x BC017553 (1-1952)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         38 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 97

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 157

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 217

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 277
```

```
Qy          81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         278 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 337

Qy         101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         338 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 397

Qy         121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         398 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 457

Qy         141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         458 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 517

Qy         161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         518 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 577

Qy         181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         578 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 637

Qy         201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         638 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 697

Qy         221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         698 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 757

Qy         241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         758 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 817

Qy         261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         818 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 877

Qy         281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         878 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 937

Qy         301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         938 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 997

Qy         321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         998 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1057

Qy         341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1058 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1117

Qy         361 IleLeu 362
               ||||||
Db        1118 ATATTG 1123
```

```
RESULT 8
E23195
LOCUS       E23195                  2499 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Topoisomerase binding protein.
ACCESSION   E23195
VERSION     E23195.1  GI:13024277
KEYWORDS    JP 1999075856-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2499)
  AUTHORS   Takashi,T. and Kazuhiko,Y.
  TITLE     Topoisomerase binding protein
  JOURNAL   Patent: JP 1999075856-A 1 23-MAR-1999;
            TAKASHI TSURUO,CHUGAI PHARMACEUT CO LTD
COMMENT     OS   Unidentified
            PN   JP 1999075856-A/1
            PD   23-MAR-1999
            PF   17-SEP-1997 JP 1997251544
            PR
            PI   TAKASHI TSURUO,KAZUHIKO YAMANE
            PC   C12N15/09,C07K14/47,C12N9/90,C12N15/00
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .2499
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2499
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                9.28e-163      Length:          2499
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2 (1-362) x E23195 (1-2499)

Qy           1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         482 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 541

Qy          21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         542 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 601

Qy          41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         602 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 661

Qy          61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         662 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 721

Qy          81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         722 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 781

Qy         101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         782 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 841

Qy         121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         842 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 901

Qy         141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         902 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 961

Qy         161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         962 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1021

Qy         181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1022 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1081

Qy         201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
```

```
Db       1082 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1141

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240

Db       1142 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1201

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260

Db       1202 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1261

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280

Db       1262 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1321

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db       1322 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1381

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db       1382 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1441

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340

Db       1442 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1501

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360

Db       1502 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1561

Qy        361 IleLeu 362

Db       1562 ATATTG 1567
```

```
RESULT 9
AF223469
LOCUS       AF223469               1936 bp    mRNA    linear   PRI 17-APR-2000
DEFINITION  Homo sapiens AD022 protein (AD022) mRNA, complete cds.
ACCESSION   AF223469
VERSION     AF223469.1  GI:7578788
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1936)
  AUTHORS   Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
  TITLE     A novel gene expressed in human adrenal gland
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1936)
  AUTHORS   Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) Chinese National Human Genome Center at
            Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
            Shanghai 201203, China
FEATURES             Location/Qualifiers
     source          1. .1936
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="adrenal gland"
     gene            1. .1936
                     /gene="AD022"
     CDS             17. .1105
                     /gene="AD022"
                     /codon_start=1
                     /product="AD022 protein"
                     /protein_id="AAF64144.1"
                     /db_xref="GI:7578789"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVRFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPPPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              1.96e-162       Length:         1936
Score:                  1897.00         Matches:        361
Percent Similarity:     99.7%           Conservative:   0
Best Local Similarity:  99.7%           Mismatches:     1
Query Match:            99.7%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2 (1-362) x AF223469 (1-1936)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20

Db         17 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 76

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40

Db         77 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGAGGTTTGCCTCGGTCGCAAGCTGCGATGCC 136

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60

Db        137 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 196

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db        197 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 256

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        257 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 316

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        317 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 376

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        377 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 436

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        437 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 496

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db        497 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 556

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db        557 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 616

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220

Db        617 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 676

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240

Db        677 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 736

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260

Db        737 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 796

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280

Db        797 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 856
```

```
Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 916

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 976

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1036

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1096

Qy        361 IleLeu 362
              ||||||
Db       1097 ATATTG 1102

RESULT 10
BD157138
LOCUS       BD157138                1898 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD157138
VERSION     BD157138.1  GI:27862896
KEYWORDS    JP 2002191363-A/11981.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1898)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 11981 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/11981
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                     Location/Qualifiers
            FT   CDS             (1). .(1086).
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.41e-162      Length:          1898
Score:                    1893.00        Matches:         361
Percent Similarity:       99.7%          Conservative:    0
Best Local Similarity:    99.7%          Mismatches:      1
Query Match:              99.5%          Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2 (1-362) x BD157138 (1-1898)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1080

Qy        361 IleLeu 362
              ||||||
Db       1081 ATATTG 1086
```

```
RESULT 11
AX878304
LOCUS       AX878304               1898 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 13209 from Patent EP1074617.
ACCESSION   AX878304
VERSION     AX878304.1  GI:40033040
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 13209 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .1089
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE90095.1"
                     /db_xref="GI:40033041"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                4.41e-162      Length:          1898
Score:                    1893.00        Matches:         361
Percent Similarity:       99.7%          Conservative:    0
Best Local Similarity:    99.7%          Mismatches:      1
Query Match:              99.5%          Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2 (1-362) x AX878304 (1-1898)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1080

Qy        361 IleLeu 362
              ||||||
Db       1081 ATATTG 1086

RESULT 12
AK002168
LOCUS       AK002168               1898 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ11306 fis, clone PLACE1010031.
ACCESSION   AK002168
VERSION     AK002168.1  GI:7023882
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
```

Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.

TITLE Complete sequencing and characterization of 21,243 full-length human cDNAs

JOURNAL Nat. Genet. 36 (1), 40-45 (2004)

PUBMED 14702039

REFERENCE 2

AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1898)

AUTHORS Isogai,T. and Otsuki,T.

TITLE Direct Submission

JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES Location/Qualifiers

source 1. .1898
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1010031"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="cloning vector: pME18SFL3"

CDS 1. .1089
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA92119.1"
/db_xref="GI:7023883"
/translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
DLLGLEKLDCGRFPSDHWGLLCNLDIIL"

ORIGIN

Alignment Scores:

Pred. No.:	4.41e-162	Length:	1898
Score:	1893.00	Matches:	361
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	99.5%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2 (1-362) x AK002168 (1-1898)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020
```

```
Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1080

Qy        361 IleLeu 362
              ||||||
Db       1081 ATATTG 1086

RESULT 13
AC152027/c
LOCUS       AC152027              150344 bp    DNA     linear   HTG 11-NOV-2004
DEFINITION  Dasypus novemcinctus clone VMRC5-12D21, WORKING DRAFT SEQUENCE, 5
            ordered pieces.
ACCESSION   AC152027
VERSION     AC152027.2  GI:55700150
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Dasypus novemcinctus (nine-banded armadillo)
  ORGANISM  Dasypus novemcinctus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Xenarthra; Dasypodidae; Dasypus.
REFERENCE   1  (bases 1 to 150344)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
            Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
            Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Laric,P.,
            Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L.,
            Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A.,
            Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,
            Shah,K., Sison,C., Stantripop,S., Stephen,E., Thomas,J.W.,
            Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
            Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625
            Fishers Lane, Rockville, MD 20852, USA
COMMENT     On Nov 11, 2004 this sequence version replaced gi:54312159.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: ifi
                 Center clone name: 012D21

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            -------------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 149480 bases at least Q40
                 Consensus quality: 149754 bases at least Q30
                 Consensus quality: 149909 bases at least Q20
                 Insert size: 155000; agarose-fp
                 Insert size: 149944; sum-of-contigs
                 Quality coverage: 9.52x in Q20 bases; agarose-fp
                 Quality coverage: 9.84x in Q20 bases; sum-of-contigs
            --------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    36273: contig of 36273 bp in length
            *    36274    36373: gap of unknown length
            *    36374    55049: contig of 18676 bp in length
            *    55050    55149: gap of unknown length
            *    55150   101939: contig of 46790 bp in length
            *   101940   102039: gap of unknown length
            *   102040   114806: contig of 12767 bp in length
            *   114807   114906: gap of unknown length
            *   114907   150344: contig of 35438 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150344
                     /organism="Dasypus novemcinctus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9361"
                     /clone="VMRC5-12D21"
                     /clone_lib="VMRC5"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .36273
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    1. .23436
                     /note="clone overlaps with GenBank Accession Number
                     AC152029 clone VMRC5-168B11 (center project name ifj)"
     gap             36274. .36373
                     /estimated_length=unknown
     misc_feature    36374. .55049
                     /note="assembly_fragment"
     gap             55050. .55149
                     /estimated_length=unknown
     misc_feature    55150. .101939
                     /note="assembly_fragment"
     gap             101940. .102039
                     /estimated_length=unknown
     misc_feature    102040. .114806
                     /note="assembly_fragment"
     gap             114807. .114906
                     /estimated_length=unknown
     misc_feature    114907. .150344
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:              2.46e-132      Length:          150344
Score:                  1590.00        Matches:         303
Percent Similarity:     90.7%          Conservative:    19
Best Local Similarity:  85.4%          Mismatches:      33
Query Match:            83.6%          Indels:          0
DB:                     14             Gaps:            0

US-10-757-745-2 (1-362) x AC152027 (1-150344)

Qy          8 GluGlyGlyArgGluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeu 27
              |||     |||                 |||||||||   ||||||||||||||||||
Db      53576 GAGCCGAGGCGGCGCGAGACGACGGGGAGCGGGGAGCCCCGGGTGAAGAAGCGGCGACTC 53517
```

```
Qy        28 LeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeu 47
Db     53516 TTGTGTGCGGAGTTTGCCTCGGTCGCGAATTGCGACCCGGCCGTGGCCCAGTGCTTCCTG 53457

Qy        48 AlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGlu 67
Db     53456 GCCGAGAACGACTGGGAGATGGAAAGAGCGCAGAACTCCTACTTTGAGCCGTGCGTGGAG 53397

Qy        68 GluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeu 87
Db     53396 GAGAGCGCCCCGGAGAGCCGGCCTCGGACCCCCTGCCAGCCCGAGGCCTGTGTTGACCTA 53337

Qy        88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 107
Db     53336 AGCAACGAAGATACGACTGATTCCAATTGTTCTAAAATCACCTCATCTGAAAATATTCAG 53277

Qy       108 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn 127
Db     53276 CAAGAAGATGGCAGTATGTTCTCTTTAATTACCTGGAATATTGATGGATTGGATCTAAAC 53217

Qy       128 AsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspVal 147
Db     53216 AACCTGCAAGAGAGGGCTCGAGGGGTGTGTTCTTACTTAACTTTGTACAGTCCAGATGTG 53157

Qy       148 IlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsn 167
Db     53156 ATATTTCTACAGGAAGTTATTCCTCCGTACTACAGTTACCTAAAGAAGAGAGCAAGTAAT 53097

Qy       168 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 187
Db     53096 TACGAGATTATTACAGGTCGTGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCA 53037

Qy       188 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 207
Db     53036 AGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAAC 52977

Qy       208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
Db     52976 CTATTGTGTGTGCATGTGAGTGTGTCAGGAAATGAACTTTACCTTATGACTTCCCATTTA 52917

Qy       228 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 247
Db     52916 GAGAGCACCAGAGGACATGCTAAGGAACGAATCAATCAGTTAAAAGTGGTTTTAAAGAAA 52857

Qy       248 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 267
Db     52856 ATGCAAGAGGCTCCAGAGACAGCTACGGTTATATTTGCAGGAGATACGAATTTAAGGGAT 52797

Qy       268 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
Db     52796 CAAGAAGTTACCAGATGTGGTGGTTTGCCCAACAACATTTTGGATGTCTGGGAGTTTTTG 52737

Qy       288 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 307
Db     52736 GGCAAGCCTGAGCATTGCCAGTTTACATGGGATACGCAAATGAACTTTAACCTTGGAATA 52677

Qy       308 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 327
Db     52676 CCTGCTACTTGTAAGCTTCGTTTTGATTGAATCTTTTTCAGAGCAGCAGCAGAAGAGGGC 52617

Qy       328 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 347
Db     52616 CATATCATTCCCCGAAGTATGGACCTTCTTGGGTTGGAAAAACTGGACTGTGGCAGATTT 52557

Qy       348 ProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
Db     52556 CCTAGTGATCACTGGGGTCTTTTGTGCAGCTTAGATATAATATTG 52512

RESULT 14
HSA420495
LOCUS       HSA420495              2151 bp    mRNA    linear   PRI 23-NOV-2001
DEFINITION  Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1761756.
ACCESSION   AJ420495
VERSION     AJ420495.1  GI:17066359
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Auffray,C., Ansorge,W., Ballabio,A., Estivill,X., Gibson,K.,
            Lehrach,H., Poustka,A. and Lundeberg,J.
  TITLE     The European IMAGE consortium for integrated Molecular analysis of
            human gene transcripts
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2151)
  AUTHORS   Persson,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2001) Persson A., Center for Molecular
            Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagsvagen
            30B, 106 91 Stockholm, SWEDEN
COMMENT     This clone is available royalty-free through IMAGE Consortium
            Distributors. IMPORTANT: This sequence represents the full insert
            of this IMAGE cDNA clone. No attempt has been made to verify
            whether this corresponds to the full-length of the original mRNA
            from which it was derived.
FEATURES             Location/Qualifiers
     source          1. .2151
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="EUROIMAGE 1761756"
                     /clone_lib="NCI_CGAP_Kid3"
     polyA_signal    2105. .2110
ORIGIN

Alignment Scores:
Pred. No.:                4.2e-123      Length:          2151
Score:                    1464.50       Matches:         298
Percent Similarity:       68.3%         Conservative:    0
Best Local Similarity:    68.3%         Mismatches:      2
Query Match:              77.0%         Indels:          137
DB:                       8             Gaps:            1

US-10-757-745-2 (1-362) x HSA420495 (1-2151)

Qy        63 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 82
Db         6 GAGGCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAG 65

Qy        82 ---------------------------------------------------------- 82
Db        66 ACCTAGTAAGTGATGAGGGGAGCGGGAGCCAGTCCGCACGGAGGCAGTTTGATACCAGGT 125

Qy        82 ---------------------------------------------------------- 82
Db       126 AGCCCTGCGCGTCCTGGGACCTGCCGTGGAGCTGGGCGGACGATGCTGAGGCAGATCTCC 185

Qy        82 ---------------------------------------------------------- 82
Db       186 AAGCTGCGTAGCCAGAGGAGCAGGGCTTGCGTTGCTGCTGCTGCGCGGAGTTGCGCCCCA 245

Qy        82 ---------------------------------------------------------- 82
Db       246 GCAGCGGATCTAACACTGCACCAGACTGTCGCTTCGAACGTGATTTGGGGTCCAGGGAGG 305

Qy        82 ---------------------------------------------------------- 82
Db       306 AAAAGATGGAAATGTAATGTTTAACTCAAATGAAAGGAAATAAGTAAGCGAGAGGTTTGA 365

Qy        82 ---------------------------------------------------------- 82
Db       366 AACAAATGAACGAAGGCTAAACCTTCTATCGAAAAATGGCAACTTGGAAGTCTCCTTGCT 425
```

```
Qy          83 ----------------------------------------------------ThrTyrValAsp  86
                                                                   |||   ||||||
Db         426 TCCTGTATTGTTTATTTATTCTAGGCCCTAGGCTGACCGGCCAGAATAACAGT-GTTGAC 484

Qy          87 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 106
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         485 CTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACT 544

Qy         107 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 126
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         545 CAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTA 604

Qy         127 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAsp 146
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         605 AACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGAT 664

Qy         147 ValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSer 166
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         665 GTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGT 724

Qy         167 AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         725 AATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAA 784

Qy         187 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 206
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         785 TCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGA 844

Qy         207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 226
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         845 AACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCAT 904

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         905 TTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAG 964

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         965 AAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGG 1024

Qy         267 AspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPhe 286
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1025 GATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTT 1084

Qy         287 LeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGly 306
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1085 TTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGA 1144

Qy         307 IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 326
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1145 ATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAG 1204

Qy         327 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 346
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1205 GGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGA 1264

Qy         347 PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
               ||||||||||||||||||||||||||||||||||||||||||||||||
Db        1265 TTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 1312
```

RESULT 15
CQ726600
LOCUS CQ726600 858 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 12534 from Patent WO02068579.
ACCESSION CQ726600
VERSION CQ726600.1 GI:42290170
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1

AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 12534 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1. .858
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:

Pred. No.:	1.76e-122	Length:	858
Score:	1452.50	Matches:	285
Percent Similarity:	78.7%	Conservative:	0
Best Local Similarity:	78.7%	Mismatches:	0
Query Match:	76.4%	Indels:	77
DB:	6	Gaps:	1

US-10-757-745-2 (1-362) x CQ726600 (1-858)

```
Qy           1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro  20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT  60

Qy          21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla  40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy          41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy          61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu  80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy          81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy         101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy         121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
               ||||||||||||||||||||||||||||||||||||||||||||
Db         361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGG---------------- 405

Qy         141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db         405 ---------------------------------------------------------- 405

Qy         161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db         405 ---------------------------------------------------------- 405

Qy         181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db         405 ---------------------------------------------------------- 405

Qy         201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
                                           ||||||||||||||||||||||||||||
Db         406 ----------------------------------GTGAATGTGTCAGGAAATGAGCTT 429

Qy         221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         430 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 489

Qy         241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        490 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 549

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280

Db        550 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 609

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        610 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 669

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        670 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 729

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340

Db        730 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 789

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360

Db        790 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 849

Qy        361 IleLeu 362

Db        850 ATATTG 855
```

```
RESULT 16
BD205491
LOCUS       BD205491               1312 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  CD40-Interacting and TRAF-interacting protein.
ACCESSION   BD205491
VERSION     BD205491.1  GI:33015261
KEYWORDS    JP 2002512796-A/2.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting protein
  JOURNAL   Patent: JP 2002512796-A 2 08-MAY-2002;
            VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT     OS   Mus musculus (mouse)
            PN   JP 2002512796-A/2
            PD   08-MAY-2002
            PF   28-APR-1999 JP 2000546003
            PR   29-APR-1998 EP   98201392.2
            PI   STEFAN MARIA CHRISTIAAN PYPE,
            PI   JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI   DANNY
            FRANCOIS EVELINE HUYLEBROECK
            PC   C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
              A61P37/02,
            PC   A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
            PC   G01N33/566,
            PC   C12N15/00,A61K37/02
            CC   CD40-Interacting and TRAF-interacting protein. FH   Key
                 Location/Qualifiers
            FT   CDS             (122). .(1234).
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
ORIGIN

Alignment Scores:
Pred. No.:              1.34e-104       Length:          1312
Score:                  1258.00         Matches:         240
Percent Similarity:     80.6%           Conservative:    43
Best Local Similarity:  68.4%           Mismatches:      64
Query Match:            66.1%           Indels:          4
DB:                     6               Gaps:            3
```

US-10-757-745-2 (1-362) x BD205491 (1-1312)

```
Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31

Db        191 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 247

Qy         32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51

Db        248 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 307

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71

Db        308 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 367

Qy         72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91

Db        368 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 427

Qy         92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111

Db        428 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 484

Qy        112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131

Db        485 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 544

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151

Db        545 AGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 604

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171

Db        605 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 664

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191

Db        665 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTT 724

Qy        192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211

Db        725 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 784

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231

Db        785 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 844

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251

Db        845 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 904

Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271

Db        905 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 964

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291

Db        965 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 1024

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311

Db       1025 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1084

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331

Db       1085 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1138

Qy        332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351

Db       1139 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1198

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
```

```
Db       1199 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1231

RESULT 17
AR594295
LOCUS       AR594295               1312 bp    DNA     linear   PAT 15-DEC-2004
DEFINITION  Sequence 3 from patent US 6812203.
ACCESSION   AR594295
VERSION     AR594295.1  GI:56643901
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting proteins
  JOURNAL   Patent: US 6812203-A 3 02-NOV-2004;
            Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
            Zwijnaarde;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.34e-104      Length:          1312
Score:                    1258.00        Matches:         240
Percent Similarity:       80.6%          Conservative:    43
Best Local Similarity:    68.4%          Mismatches:      64
Query Match:              66.1%          Indels:          4
DB:                       6              Gaps:            3

US-10-757-745-2 (1-362) x AR594295 (1-1312)

Qy          12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31
               ||||||||||:::   |||   |||       ||||||:::|||||||||   |||:::
Db         191 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 247

Qy          32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51
               ||||||   |||       ||||||      :::             |||   |||||||||
Db         248 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 307

Qy          52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
               |||:::   ::::::||||||::::::|||||||||   |||       :::
Db         308 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 367

Qy          72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91
               |||:::|||   |||       :::   :::   |||||||||||||||||||:::
Db         368 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 427

Qy          92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111
                  |||:::|||       :::   |||||||||       |||       |||:::
Db         428 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 484

Qy         112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131
               |||       |||   |||||||||||||||||||||||||||       ||||||   |||
Db         485 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 544

Qy         132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
               ||||||||||||||||||||   |||||||||||||||||||||||:::||||||||||
Db         545 AGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 604

Qy         152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
               ||||||||||||||||   :::|||||||||||||||:::::::::|||   ||||||
Db         605 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 664

Qy         172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191
               ||||||:::|||||||||||||||||||||||||:::|||||||||   ||||||||||
Db         665 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTT 724

Qy         192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211
               ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||

Db         725 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 784

Qy         212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
               :::|||:::::   |||||||||   ||||||||||||||||||||||||||||||||
Db         785 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 844

Qy         232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
               |||:::|||||||||:::   |||||||||   ||||||   ||||||||||||||||
Db         845 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 904

Qy         252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271
               |||:::|||   ||||||||||||||||||||||||||||||||||:::||||||
Db         905 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 964

Qy         272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
               :::|||||||||||||||:::|||:::   |||   |||||||||||||||||||||
Db         965 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 1024

Qy         292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
               ||||||||||||||||||||||:::   |||:::||||||   |||   ||||||
Db        1025 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1084

Qy         312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
               |||   ||||||||||||||||||||||||       ||||||||||||||:::||||||
Db        1085 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1138

Qy         332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
               :::|||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        1139 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1198

Qy         352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
               ||||||||||||||   |||:::::::::|||
Db        1199 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1231

RESULT 18
AX011601
LOCUS       AX011601               1312 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 3 from Patent WO9955859.
ACCESSION   AX011601
VERSION     AX011601.1  GI:9998125
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 3 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
     CDS             122. .1234
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07556.1"
                     /db_xref="GI:9998126"
                     /translation="MASGSSSDAAEPAGPAGRAASAPEAAQAEEDRVKRRRLQCLGFA
                     LVGGCDPTMVPSVLRENDWQTQKALSAYFELPENDQGWPRQPPTSFKSEAYVDLTNED
                     ANDTTILEASPSGTPLEDSSTISFITWNIDGLDGCNLPERARGVCSCLALYSPDVVFL
                     QEVIPPYCAYLKKRAASYTIITGNEEGYFTAILLKKGRVKFKSQEIIPFPNTKMMRNL
                     LCVNVSLGGNEFCLMTSHLESTREHSAERIRQLKTVLGKMQEAPDSTTVIFAGDTNLR
                     DQEVIKCGGLPDNVFDAWEFLGKPKHCQYTWDTKANNNLRIPAAYKHRFDRIFFRAEE
                     GHLIPQSLDLVGLEKLDCGRFPSDHWGLLCTLNVVL"
ORIGIN

Alignment Scores:
```

Pred. No.:	1.34e-104	Length:	1312
Score:	1258.00	Matches:	240
Percent Similarity:	80.6%	Conservative:	43
Best Local Similarity:	68.4%	Mismatches:	64
Query Match:	66.1%	Indels:	4
DB:	6	Gaps:	3

US-10-757-745-2 (1-362) x AX011601 (1-1312)

```
Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31
              ||||||||||::  |||  |||      ||||||:::|||||||||   |||:::
Db        191 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 247

Qy         32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51
              ||||||   |||      ||||||       :::           |||   |||||||||
Db        248 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 307

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
              |||:::   :::::||||||:::::|||||||||  |||       :::
Db        308 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 367

Qy         72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91
              |||:::|||   |||      :::   :::   ||||||||||||||||||||:::
Db        368 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 427

Qy         92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111
              |||:::|||      :::   |||||||||      |||      |||:::
Db        428 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 484

Qy        112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131
              |||      |||  |||||||||||||||||||||||||      ||||||   |||
Db        485 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 544

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
              ||||||||||||||||||||  ||||||||||||||||||||:::|||||||||
Db        545 AGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 604

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
              ||||||||||||||||||   :::|||||||||||||:::::::::|||   ||||||
Db        605 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 664

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191
              ||||||:::|||||||||||||||||||||||:::|||||||||   |||||||||
Db        665 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTT 724

Qy        192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db        725 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 784

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
              :::|||::::::  |||||||||  ||||||||||||||||||||||||||||||||
Db        785 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 844

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
              |||:::|||||||||:::  |||||||||  ||||||  ||||||||||||||||
Db        845 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 904

Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271
              |||:::|||  ||||||||||||||||||||||||||||||||||:::|||||||
Db        905 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 964

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
              :::|||||||||||||:::|||:::   |||  ||||||||||||||||||||||
Db        965 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 1024

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
              ||||||||||||||||||||||:::   |||:::||||||  |||  ||||||
Db       1025 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1084

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              |||  ||||||||||||||||||||||||  ||||||||||||||:::||||||
Db       1085 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1138

Qy        332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHis 351
              :::||||||||||||:::|||||||||||||||||||||||||||||||||||||||
Db       1139 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1198

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||||||  |||:::::::::|||
Db       1199 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1231
```

RESULT 19
MMU251328
LOCUS MMU251328 1943 bp mRNA linear ROD 27-MAY-2000
DEFINITION Mus musculus mRNA for putative TRAF and TNF receptor associated protein (TTRAP gene).
ACCESSION AJ251328
VERSION AJ251328.1 GI:8247282
KEYWORDS ttrap gene.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P., Huylebroeck,D. and Remacle,J.E.
TITLE TTRAP, a novel protein that associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor-associated factors (TRAFs), and that inhibits nuclear factor-kappa B activation
JOURNAL J. Biol. Chem. 275 (24), 18586-18593 (2000)
PUBMED 10764746
REFERENCE 2 (bases 1 to 1943)
AUTHORS Pype,S.M.C.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-1999) Pype S.M.C., VIB07, Department Cell Growth, Differentiation and Developme, Flanders Interuniversity Institute for Biotechnology, K.U.Leuven, Campus Gasthuisberg, Herestraat 49, 3000 Leuven, BELGIUM
COMMENT Related sequence: AJ269473.
FEATURES Location/Qualifiers
source 1. .1943
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/note="IMAGE cDNA EST clone 1262914"
gene 1. .1943
/gene="TTRAP"
CDS 48. .1160
/gene="TTRAP"
/codon_start=1
/product="putative TRAF and TNF receptor associated protein"
/protein_id="CAB92971.1"
/db_xref="GI:8247283"
/db_xref="GOA:Q9JJX7"
/db_xref="UniProt/TrEMBL:Q9JJX7"
/translation="MASGSSSDAAEPAGPAGRAASAPEAAQAEEDRVKRRRLQCLGFA
LVGGCDPTMVPSVLRENDWQTQKALSAYFELPENDQGWPRQPPTSFKSEAYVDLTNED
ANDTTILEASPSGTPLEDSSTISFITWNIDGLDGCNLPERARGVCSCLALYSPDVVFL
QEVIPPYCAYLKKRAASYTIITGNEEGYFTAILLKKGRVKFKSQEIIPFPNTKMMRNL
LCVNVSLGGNEFCLMTSHLESTREHSAERIRQLKTVLGKMQEAPDSTTVIFAGDTNLR
DQEVIKCGGLPDNVFDAWEFLGKPKHCQYTWDTKANNNLRIPAAYKHRFDRIFFRAEE
GHLIPQSLDLVGLEKLDCGRFPSDHWGLLCTLNVVL"
ORIGIN

Alignment Scores:

Pred. No.:	2.12e-104	Length:	1943
Score:	1258.00	Matches:	240
Percent Similarity:	80.6%	Conservative:	43
Best Local Similarity:	68.4%	Mismatches:	64
Query Match:	66.1%	Indels:	4
DB:	9	Gaps:	3

US-10-757-745-2 (1-362) x MMU251328 (1-1943)

```
Qy        12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31
             |||||||||:::   |||  |||      ||||||:::|||||||||   |||:::
Db       117 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 173

Qy        32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51
             ||||||   |||     ||||||     :::              |||   |||||||||
Db       174 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 233

Qy        52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
             |||:::   ::::::||||||::::::|||||||||   |||      :::
Db       234 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 293

Qy        72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91
                |||:::|||  |||      :::  :::  ||||||||||||||||||||||:::
Db       294 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 353

Qy        92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111
                |||:::|||     :::   |||||||||      |||      |||:::
Db       354 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 410

Qy       112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131
             |||      |||  ||||||||||||||||||||||||||||      ||||||   |||
Db       411 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 470

Qy       132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
             ||||||||||||||||||  |||||||||||||||||||||||:::|||||||||
Db       471 AGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 530

Qy       152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
             ||||||||||||||||   :::|||||||||||||||:::::::|||   ||||||
Db       531 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 590

Qy       172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191
             ||||||:::|||||||||||||||||||||||:::|||||||||   |||||||||
Db       591 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTT 650

Qy       192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211
             ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db       651 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 710

Qy       212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
             :::|||::::::  |||||||||  ||||||||||||||||||||||||||||||||||
Db       711 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 770

Qy       232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
                |||:::|||||||||:::   |||||||||   ||||||   ||||||||||||||||
Db       771 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 830

Qy       252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271
             |||:::|||   ||||||||||||||||||||||||||||||||||:::|||||||
Db       831 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 890

Qy       272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
             :::|||||||||||||||:::|||:::   |||  ||||||||||||||||||||||||
Db       891 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 950

Qy       292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
             ||||||||||||||||||||||:::   |||:::||||||   |||   ||||||
Db       951 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1010

Qy       312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
             |||   ||||||||||||||||||||||      ||||||||||||||:::||||||
Db      1011 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1064

Qy       332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHis 351
             :::||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db      1065 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1124

Qy       352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
             ||||||||||||||   |||:::::::::|||
Db      1125 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1157
```

```
RESULT 20
CR761829
LOCUS       CR761829               1236 bp    mRNA    linear   VRT 15-SEP-2004
DEFINITION  Xenopus tropicalis finished cDNA, clone TGas102g02.
ACCESSION   CR761829
VERSION     CR761829.1  GI:51967738
KEYWORDS    .
SOURCE      Xenopus tropicalis (Silurana tropicalis)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1236)
  AUTHORS   Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
            Francis,M.D., Garrett,N., Gilchrist,M.J., Grafham,D.V.,
            McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
            Voigt,J. and Zorn,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
COMMENT     Sanger Xenopus tropicalis EST/cDNA project.
            This sequence is from a Xenopus Gene Collection (XGC) library, from
            a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
            extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
            ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'
            end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1-blue.
FEATURES             Location/Qualifiers
     source          1. .1236
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TGas102g02"
                     /clone_lib="XGC-gastrula"
                     /dev_stage="gastrula (stage 10.5-13 mixed)"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              2.55e-98        Length:         1236
Score:                  1188.50         Matches:        227
Percent Similarity:     75.7%           Conservative:   53
Best Local Similarity:  61.4%           Mismatches:     71
Query Match:            62.5%           Indels:         19
DB:                     5               Gaps:           4

US-10-757-745-2 (1-362) x CR761829 (1-1236)

Qy         9 GlyGlyArgGluAlaAlaGluGluGluGlyGluProGluVal---LysLysArgArgLeu 27
             |||      :::          |||||||||      |||      :::|||
Db        67 GGAGCTGTACAGTGTGGGACTGGGGAGGGCGAGGCAGTGGTGCTCAATGAGAGGAGCGAA 126

Qy        28 LeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeu 47
                |||      ||||||||:::|||   ||||||   ||||||||||||||||||||
Db       127 CAGTGTTCTGCGTTTGCTTCCATCGCTGGCTGTGATGAGGCAGTTGCGCAATGTTTCCTA 186

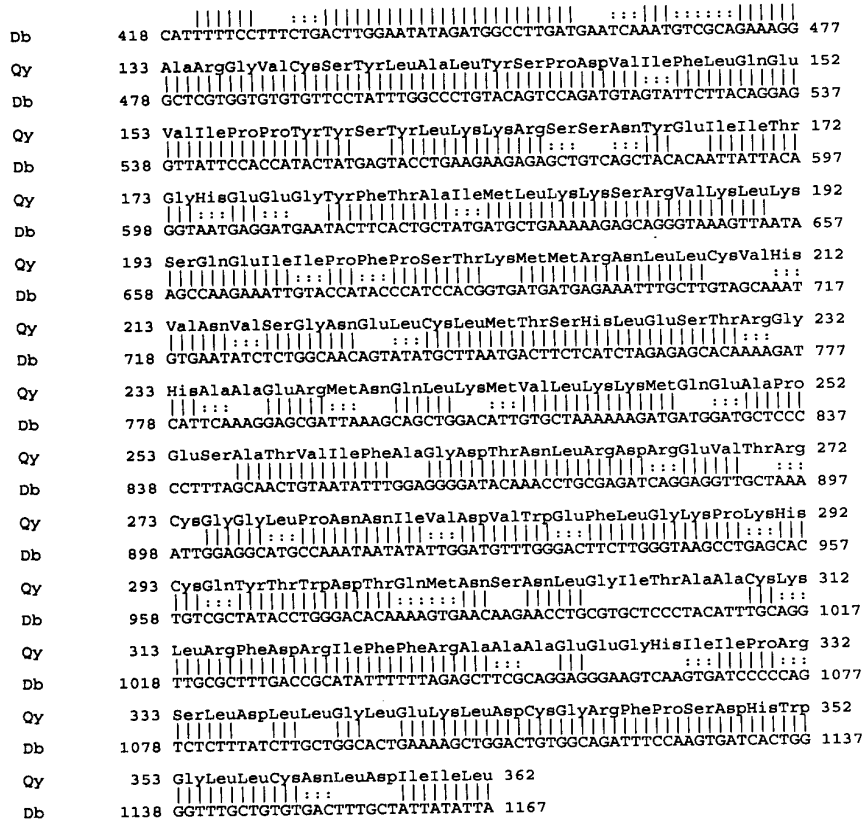
Qy        48 AlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGlu 67
             ||||||||||||||:::||||||||||||:::||||||:::|||||||||   ||||||
Db       187 GCTGAGAATGACTGGGATATGGAAAGGGCGATTAATTCTTTCTTTGAGCCTGGGGTGGAG 246

Qy        68 GluSerAlaLeuGluArgArgProGluThrIleSerGluPro---------------- 81
                :::          :::      |||::::::|||
Db       247 TCAGCCTTGCAAAACAAAGCTGCAGCTGATATAGCAGATCCTTTAAAGCAAGAGACTATG 306

Qy        82 ---------------LysThrTyrValAspLeuThrAsnGluGlu------------Thr 92
                           :::   :::|||||||||   ::::::
Db       307 TCTGGAACTGCTTCCGATTCATGTATTGACCTCACCGGCGATGACTTAGTGGTTACCAAA 366

Qy        93 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 112
             :::::::::|||||||||      |||      :::|||||:::   |||
Db       367 TCAGAAGCAACTACCAGCAAT---------TCCTCCACTGTGAAGCAAGAAGATGAGAGC 417

Qy       113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132
```

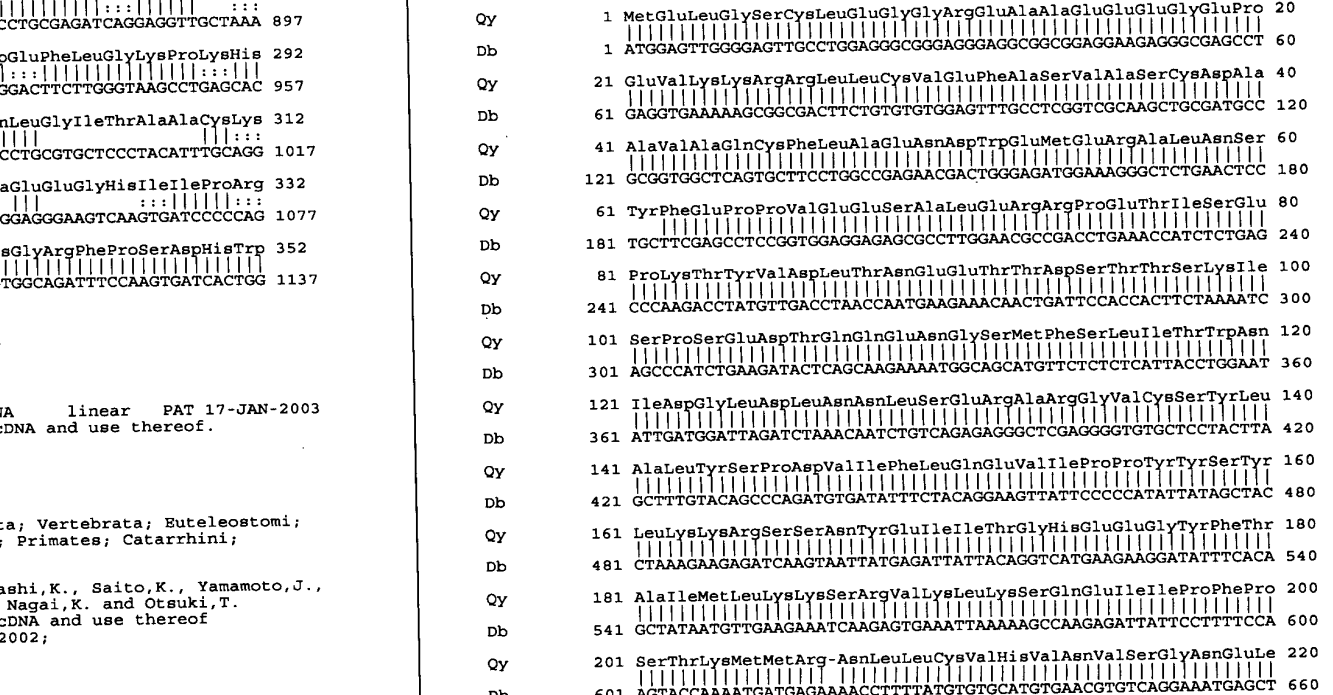


RESULT 21
BD150065
LOCUS BD150065 752 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD150065
VERSION BD150065.1 GI:27855823
KEYWORDS JP 2002191363-A/4908.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 752)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 4908 09-JUL-2002;
HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/4908
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers
FT source 1. .752
FT /organism='Homo sapiens (human)'.
FEATURES Location/Qualifiers
source 1. .752
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:

Pred. No.:	5.55e-98	Length:	752
Score:	1182.00	Matches:	238
Percent Similarity:	94.8%	Conservative:	1
Best Local Similarity:	94.4%	Mismatches:	10
Query Match:	62.1%	Indels:	3
DB:	6	Gaps:	0

US-10-757-745-2 (1-362) x BD150065 (1-752)



```
Qy        220 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 240
              ||||||| ||||||  |||   :::|||  ||||||     |||||||||||||
Db        661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy        240 lnLeuLysMetValLeu-LysLysMetGln 249
              ||  ||||||  |||   |||||||||
Db        721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750

RESULT 22
AX870003
LOCUS       AX870003                752 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 4908 from Patent EP1074617.
ACCESSION   AX870003
VERSION     AX870003.1  GI:40024866
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 4908 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              5.55e-98        Length:         752
Score:                  1182.00         Matches:        238
Percent Similarity:     94.8%           Conservative:   1
Best Local Similarity:  94.4%           Mismatches:     10
Query Match:            62.1%           Indels:         3
DB:                     6               Gaps:           0

US-10-757-745-2 (1-362) x AX870003 (1-752)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArg-AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 660

Qy        220 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 240
              ||||||| ||||||  |||   :::|||  ||||||     |||||||||||||
Db        661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy        240 lnLeuLysMetValLeu-LysLysMetGln 249
              ||  ||||||  |||   |||||||||
Db        721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750

RESULT 23
BD139846/c
LOCUS       BD139846                1079 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Compounds for immunotherapy and diagnosis of breast cancer and
            methods for their use.
ACCESSION   BD139846
VERSION     BD139846.1  GI:23234791
KEYWORDS    JP 2002507387-A/63.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for immunotherapy and diagnosis of breast cancer and
            methods for their use
  JOURNAL   Patent: JP 2002507387-A 63 12-MAR-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002507387-A/63
            PD   12-MAR-2002
            PF   22-DEC-1998 JP 2000526543
            PR   24-DEC-1997 US    08/998253,24-DEC-1997 US    08/998255 PR
            17-JUL-1998 US    09/118627,17-JUL-1998 US    09/118554 PI   STEVEN
            G REED,JIANGCHUN XU
            PC   C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,
            PC   C07K14/47,
            PC
            C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/ PC
             53,
            PC   G01N33/577,C12N15/00,A61K37/02,C12N5/00
            CC   Compounds for immunotherapy and diagnosis
            of breast cancer and
            CC            methods for
            CC   their use
            FH   Key             Location/Qualifiers
            FT   source          1. .1079
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.41e-97        Length:         1079
Score:                  1177.00         Matches:        226
Percent Similarity:     99.6%           Conservative:   0
Best Local Similarity:  99.6%           Mismatches:     1
```

```
Query Match:            61.9%       Indels:        1
DB:                     6           Gaps:          0

US-10-757-745-2 (1-362) x BD139846 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398

RESULT 24
AR202950/c
LOCUS       AR202950               1079 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 63 from patent US 6365348.
ACCESSION   AR202950
VERSION     AR202950.1  GI:21499210
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for diagnosis of Breast cancer and methods for their use
  JOURNAL   Patent: US 6365348-A 63 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              2.41e-97    Length:        1079
Score:                  1177.00     Matches:       226
Percent Similarity:     99.6%       Conservative:  0
Best Local Similarity:  99.6%       Mismatches:    1
Query Match:            61.9%       Indels:        1
DB:                     6           Gaps:          0

US-10-757-745-2 (1-362) x AR202950 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398

RESULT 25
AR208051/c
LOCUS       AR208051               1079 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 63 from patent US 6379951.
ACCESSION   AR208051
VERSION     AR208051.1  GI:21507966
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for immunotherapy of breast cancer and methods for their
            use
  JOURNAL   Patent: US 6379951-A 63 30-APR-2002;
```

```
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                2.41e-97       Length:          1079
Score:                    1177.00        Matches:         226
Percent Similarity:       99.6%          Conservative:    0
Best Local Similarity:    99.6%          Mismatches:      1
Query Match:              61.9%          Indels:          1
DB:                       6              Gaps:            0

US-10-757-745-2 (1-362) x AR208051 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398

RESULT 26
AX429924/c
LOCUS       AX429924               1079 bp    DNA     linear   PAT 21-JUN-2002
DEFINITION  Sequence 63 from Patent WO0198339.
ACCESSION   AX429924
VERSION     AX429924.1  GI:21541088
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.
  TITLE     Compositions and methods for the therapy and diagnosis of breast
            cancer
  JOURNAL   Patent: WO 0198339-A 63 27-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                2.41e-97       Length:          1079
Score:                    1177.00        Matches:         226
Percent Similarity:       99.6%          Conservative:    0
Best Local Similarity:    99.6%          Mismatches:      1
Query Match:              61.9%          Indels:          1
DB:                       6              Gaps:            0

US-10-757-745-2 (1-362) x AX429924 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398
```

```
RESULT 27
BC083404
LOCUS       BC083404               1451 bp    mRNA    linear   VRT 27-MAY-2005
DEFINITION  Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7238334), partial
            cds.
ACCESSION   BC083404
VERSION     BC083404.1  GI:53734140
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 1451)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1451)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: John Postlethwait, University of Oregon
            cDNA Library Preparation: Dr. Sumio Sugano
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 192 Row: c Column: 18.
FEATURES             Location/Qualifiers
     source          1. .1451
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:7238334"
                     /tissue_type="Ovary"
                     /clone_lib="NIH_ZGC_5"
                     /lab_host="DH10B"
                     /note="Vector: pME18S-FL3"
     gene            <1. .1451
                     /gene="TTRAP"
                     /db_xref="GeneID:553516"
     CDS             <1. .1141
                     /gene="TTRAP"
                     /codon_start=2
                     /product="TTRAP protein"
                     /protein_id="AAH83404.1"
                     /db_xref="GI:53734141"
                     /db_xref="GeneID:553516"
                     /translation="FLSGNQTDQQMSALEESRTRLCDQFAFVSGSDSAVAQCYLAENE
                     WDMERALNSFFEAHMDSVFDEEAEKTEVTGNKRKDDTAEASGTKKKLKTDNADCIDLT
                     AEEPTCSITVNSKENQAENGTAKSEVEDSKLSIISWNVDGLDTLNLADRARGLCSYLA
                     LYTPDVVFLQELIPAYVQYLKKRAVSYLFFEGSDDGYFTGIMLRKSRVKFLESEIICF
                     PTTQMMRNLLIAQVTFSGQKLYLMTSHLESCKNQSQERTKQLRVVLQKIKEAPEDAIV
                     IFAGDTNLRDAEVANVGGLPAGVCDVWEQLGKQEHCRYTWDTKANSNKTVPYVSRCRF
                     DRIFLRSAKTGPPVTPDHMALIGMEKLDCGRYTSDHWGIYCTFNT"
ORIGIN

Alignment Scores:
Pred. No.:                3.25e-76      Length:          1451
Score:                    946.00        Matches:         188
Percent Similarity:       65.8%         Conservative:    53
Best Local Similarity:    51.4%         Mismatches:      93
Query Match:              49.7%         Indels:          32
DB:                       5             Gaps:            4

US-10-757-745-2 (1-362) x BC083404 (1-1451)

Qy         23 LysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaVal 42
              :::  |||       ||||||  :::|||||  |||:::     |||:::||||||
Db         47 GAATCCAGGACGAGACTTTGTGATCAGTTTGCCTTTGTATCTGGATCGGACAGCGCCGTG 106

Qy         43 AlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPhe 62
              |||||||||:::|||||||||||||:::|||:::||||||||||||||||||||:::|||
Db        107 GCACAGTGCTATCTAGCAGAAAATGAGTGGGACATGGAGAGAGCTTTAAACTCTTTCTTT 166

Qy         63 Glu--------------------------------ProProValGluGluSerAlaLeu 71
              |||                                          |||   :::
Db        167 GAGGCTCATATGGATTCAGTCTTTGATGAGGAAGCAGAGAAAACTGAAGTGACTGGGAAT 226

Qy         72 GluArgArgProGluThrIleSerGluProLysThr------------------------ 83
              :::|||:::   :::|||              |||
Db        227 AAGAGAAAAGATGACACTGCTGAAGCTTCAGGGACGAAGAAGAAATTGAAGACGGACAAT 286

Qy         84 ------TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSer 101
                    :::||||||||||  ||||||  |||  |||  |||
Db        287 GCGGATTGCATTGACTTGACAGCAGAGGAGCCCACCTGCTCTATTACTGTAAAC------ 340

Qy        102 ProSerGluAspThrGlnGlnGluAsnGly------------------------SerMet 113
              |||::::::  |||  |||||||||                          |||
Db        341 ---TCTAAAGAAAATCAGGCTGAAAATGGCACAGCCAAATCTGAAGTGGAGGACAGCAAG 397

Qy        114 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 133
                 |||:::|||:::|||||||:::|||||||||||||      ||||||:::::|||||
Db        398 CTCTCCATCATCAGCTGGAATGTGGACGGTTTGGACACATTAAATCTTGCAGATCGTGCC 457

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
              |||||||:::|||||||||||||||||||:::|||||||:::||||||||||||:::
Db        458 AGGGGCTTGTGCTCATATCTAGCTCTATACACACCAGATGTGGTTTTTCTTCAAGAGCTC 517

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
              ||||||  |||          ||||||||||||||:::  :::|||          |||
Db        518 ATTCCAGCTTATGTTCAGTATCTTAAGAAACGTGCCGTCAGCTACTTGTTTTTTGAAGGA 577

Qy        174 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 193
              ::::::||||||||||||   ||||||||||:::||||||||||||||||
Db        578 AGCGATGATGGATACTTTACCGGGATAATGTTGAGGAAATCAAGAGTCAAATTTTTGGAA 637

Qy        194 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 213
              |||||||||  ||||||:::|||:::|||||||||||||||||||          |||
Db        638 AGTGAGATCATCTGCTTTCCAACAACACAAATGATGAGGAATCTGCTCATTGCTCAGGTG 697

Qy        214 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 233
              ||||||  :::|||  ||||||||||||||||||||||||||  :::
Db        698 ACTTTCTCAGGTCAGAAGCTATACTTGATGACATCCCATTTAGAAAGCTGTAAAAACCAA 757
```

```
Qy        234 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 253
              :::   ||||||      ||||||:::::||||||:::|||::::::||||||||||||
Db        758 TCCCAGGAGAGAACGAAACAATTGCGGGTTGTGCTTCAAAAGATTAAGGAAGCGCCAGAA 817

Qy        254 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 273
                 |||   ||||||||||||||||||||||||||||||||||||   ||||||
Db        818 GATGCCATTGTCATTTTTGCAGGTGACACAAATCTCAGAGACGCAGAGGTGGCCAATGTT 877

Qy        274 GlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCys 293
              ||||||||||||      :::   ||||||||||||   |||||||||   :::||||||
Db        878 GGAGGTTTGCCTGCAGGTGTGTGTGATGTGTGGGAACAGCTGGGCAAACAAGAACACTGC 937

Qy        294 GlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeu 313
              :::||||||||||||||:::   |||||||||      :::         :::
Db        938 AGGTACACCTGGGACACCAAAGCTAACAGCAATAAGACTGTGCCTTACGTCAGCAGATGT 997

Qy        314 ArgPheAspArgIlePhePheArgAlaAlaGluGluGlyHisIleIleProArgSer 333
              ||||||||||||||||   |||:::|||          :::   |||
Db        998 CGCTTCGACCGCATATTCCTGAGGTCTGCTAAAACCGGTCCGCCGGTCACTCCTGACCAC 1057

Qy        334 LeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGly 353
              :::   |||:::|||:::|||||||||||||||||||:::   |||||||||||||||
Db       1058 ATGGCCTTGATTGGAATGGAGAAACTGGACTGCGGTCGTTATACCAGTGATCACTGGGGT 1117

Qy        354 LeuLeuCysAsnLeuAsp 359
              :::   |||     :::
Db       1118 ATTTACTGCACTTTCAAC 1135
```

RESULT 28
BC097117

```
LOCUS       BC097117               1455 bp    mRNA    linear   VRT 23-JUN-2005
DEFINITION  Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7431816), partial
            cds.
ACCESSION   BC097117
VERSION     BC097117.1  GI:66910444
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 1455)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1455)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Will Talbot
            cDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 226 Row: g Column: 18
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 53734140.
FEATURES              Location/Qualifiers
     source           1. .1455
                      /organism="Danio rerio"
                      /mol_type="mRNA"
                      /db_xref="taxon:7955"
                      /clone="IMAGE:7431816"
                      /tissue_type="Embryo, whole, 2-8 hr, mixed wild type, pool
                      of 2500 embryos"
                      /clone_lib="NIH_ZGC_15"
                      /lab_host="DH10B"
                      /note="Vector: pME18S-FL3"
     gene             <1. .1455
                      /gene="TTRAP"
                      /db_xref="GeneID:553516"
     CDS              <1. .438
                      /gene="TTRAP"
                      /codon_start=1
                      /product="TTRAP protein"
                      /protein_id="AAH97117.1"
                      /db_xref="GI:66910445"
                      /db_xref="GeneID:553516"
                      /translation="VFSGNQTDQQMSALGESRTRLCDQFAFVSGSDSAVAQCYLAENE
                      WDMERALNSFFEAHMDSVFDEEAEKTEVTGNKRKDDTAEASGTKKKLKTDNADCIDLT
                      AEEPTCSITVNSKKIRLKMAQPNLKWRTASSPSSAGMWTVWTH"
ORIGIN
```

Alignment Scores:

Pred. No.:	4e-74	Length:	1455
Score:	923.00	Matches:	188
Percent Similarity:	65.6%	Conservative:	52
Best Local Similarity:	51.4%	Mismatches:	94
Query Match:	48.5%	Indels:	33
DB:	5	Gaps:	4

US-10-757-745-2 (1-362) x BC097117 (1-1455)

```
Qy         23 LysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaVal 42
              :::   |||      ||||||   :::||||||   |||:::     |||:::||||||
Db         46 GAATCCAGGACGAGACTTTGTGATCAGTTTGCCTTTGTATCTGGATCGGACAGCGCCGTG 105

Qy         43 AlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPhe 62
              |||||||||:::|||||||||||:::|||:::||||||||||||||||||||:::|||
Db        106 GCACAGTGCTATCTAGCAGAAAATGAGTGGGACATGGAGAGAGCTTTAAACTCTTTCTTT 165

Qy         63 Glu--------------------------------ProProValGluGluSerAlaLeu 71
              |||                                     |||   :::
Db        166 GAGGCTCATATGGATTCAGTCTTTGATGAGGAAGCAGAGAAAACTGAAGTGACTGGGAAT 225

Qy         72 GluArgArgProGluThrIleSerGluProLysThr------------------------ 83
              :::|||:::   :::|||          |||
Db        226 AAGAGAAAAGATGACACTGCTGAAGCTTCAGGGACGAAGAAGAAATTGAAGACGGACAAT 285

Qy         84 ------TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSer 101
                    :::|||||||||   ||||||   |||   |||   |||   |||:::
Db        286 GCGGATTGCATTGACTTGACAGCAGAGGAGCCCACCTGCTCTATTACTGT-AAACTC--- 341
```

```
Qy        102 ProSerGluAspThrGlnGlnGluAsnGly----------------------SerMet 113
Db        342 ------TAAGAAAATCAGGCTGAAAATGGCACAGCCAAATCTGAAGTGGAGGACAGCAAG 395

Qy        114 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 133
Db        396 CTCTCCATCATCAGCTGGAATGTGGACGGTTTGGACACATTAAATCTTGCAGATCGTGCC 455

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
Db        456 AGGGGCTTGTGCTCATATCTAGCTCTATACACACCAGATGTGGTTTTTCTTCAAGAGCTC 515

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
Db        516 ATTCCAGCTTATGTTCAGTATCTTAAGAAACGTGCCGTCAGCTACTTGTTTTTTGAAGGA 575

Qy        174 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 193
Db        576 AGCGATGATGGATACTTTACCGGGATAATGTTAAGGAAATCAAGAGTAAAATTTTTGGAA 635

Qy        194 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 213
Db        636 AGCGAGATCATCTGTTTTCCAACAACACAAATGATGAGGAATCTGCTCATTGCTCAGGTG 695

Qy        214 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 233
Db        696 ACTTTCTCAGGTCAGAAGCTATACTTGATGACATCCCATTTAGAAAGCTGTAAAAACCAA 755

Qy        234 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 253
Db        756 TCCCAGGAGAGAACGAAACAATTGCGGGTTGTGCTTCAAAAGATTAAGGAAGCGCCAGAA 815

Qy        254 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 273
Db        816 GATGCCATTGTCATTTTTGCAGGCGACACAAATCTCAGAGACGCAGAGGTGGCCAATGTT 875

Qy        274 GlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCys 293
Db        876 GGAGGTTTGCCTGCAGGTGTGTGTGATGTGTGGGAACAGCTGGGCAAACAAGAACACTGC 935

Qy        294 GlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeu 313
Db        936 AGGTACACCTGGGACACCAAAGCTAACAGCAATAAGACTGTGCCTTACGTCAGCAGATGT 995

Qy        314 ArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSer 333
Db        996 CGCTTCGACCGCATATTCCTGAGGTCTGCTAAAACCGCTCCGCCGGTCACTCCTGACCAC 1055

Qy        334 LeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHisTrpGly 353
Db       1056 AAGGCCTTGATTGGAATGGAGAAACTGGACTGCGGTCGTTATACCAGTGATCACTGGGGT 1115

Qy        354 LeuLeuCysAsnLeuAsp 359
Db       1116 ATTTACTGCACTTTCAAC 1133
```

RESULT 29
BD076937

```
LOCUS       BD076937                483 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  5' EST of secretory protein expressed in prostate.
ACCESSION   BD076937
VERSION     BD076937.1  GI:22622540
KEYWORDS    JP 2001512013-A/184.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
  TITLE     5' EST of secretory protein expressed in prostate
  JOURNAL   Patent: JP 2001512013-A 184 21-AUG-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001512013-A/184
            PD   21-AUG-2001
            PF   31-JUL-1998 JP 2000505291
            PR   01-AUG-1997 US     08/905144
            PI   JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO  PI
            LACROIX
            PC   C12N15/09,C07K14/47,C12P21/02,C12Q1/02,C12Q1/68,C12N15/00 CC
            blastn
            CC   identity 99
            CC   region 1. .310
            CC   id T26956
            CC   est
            CC   blastn
            CC   identity 98
            CC   region 1. .315
            CC   id T31666
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 137. .267
            CC   id R14990
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 63. .137
            CC   id R14990
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 1. .50
            CC   id R14990
            CC   est
            CC   Von Heijne matrix
            CC   score 5.2
            CC   seq LCVEFASVASCDA/AV
            FH   Key             Location/Qualifiers
            FT   misc_feature    105. .414
            FT   misc_feature    45. .359
            FT   misc_feature    202. .332
            FT   misc_feature    127. .201
            FT   misc_feature    65. .114
            FT   sig_peptide     1. .120.
FEATURES                Location/Qualifiers
     source             1. .483
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN
```

Alignment Scores:

Pred. No.:	7.09e-67	Length:	483
Score:	837.00	Matches:	161
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	44.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2 (1-362) x BD076937 (1-483)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
Db        121 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 Leu 161
               |||
Db        481 CTA 483

RESULT 30
CQ663156
LOCUS       CQ663156                400 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 8082 from Patent WO02070737.
ACCESSION   CQ663156
VERSION     CQ663156.1  GI:42140949
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 8082 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.54e-44        Length:         400
Score:                  590.00          Matches:        111
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  99.1%           Mismatches:     0
Query Match:            31.0%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2 (1-362) x CQ663156 (1-400)

Qy        223 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 242
               :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GTGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAA 124

Qy        243 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGAT 184

Qy        263 ThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 282
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGAT 244

Qy        283 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 302
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 GTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 304

Qy        303 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAla 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCA 364

Qy        323 AlaAlaGluGluGlyHisIleIleProArgSerLeu 334
               ||||||||||||||||||||||||||||||||||
Db        365 GCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 400

RESULT 31
AL589699
LOCUS       AL589699                256608 bp    DNA     linear   ROD 17-DEC-2004
DEFINITION  Mouse DNA sequence from clone RP23-92G13 on chromosome 13 Contains
            the 3' end of the gene for a novel protein similar to KIAA0386,
            five novel genes, the Gmnn gene for geminin, gene PNAS-27, the
            Ttrap gene for Traf and Tnf receptor associated protein, the gene
            for a novel protein similar to KIAA0319, the gene for the ortholog
            of human and rat aldehyde dehydrogenase 5A1 (succinate-semialdehyde
            dehydrogenase) (ALDH5A1) (SSADH), the 5' end of the Gpld1 gene for
            glycosylphosphatidylinositol specific phospholipase D1 and four CpG
            islands, complete sequence.
ACCESSION   AL589699
VERSION     AL589699.4  GI:13810107
KEYWORDS    HTG; Aldh5a1; CpG island; geminin; Gmnn; Gpld1; KIAA0319; KIAA0386;
            phospholipase; SSADH; Ttrap.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 256608)
  AUTHORS   Pelan,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 26, 2001 this sequence version replaced gi:13751568.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep ----------
            Genome Center
            Center: UK Medical Research Council
            Center code: UK-MRC
            Web site: http://mrcseq.har.mrc.ac.uk
            Contact: mouseq@har.mrc.ac.uk
            ----------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence.  Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
            RP23-92G13 is from the RPCI-23 Mouse BAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .256608
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="13"
                     /clone="RP23-92G13"
                     /clone_lib="RPCI-23"
     gene            join(AL513014.9:11808. .79376,4770. .28457)
                     /locus_tag="RP23-461O24.1-003"
     mRNA            join(AL513014.9:11808. .12037,AL513014.9:43061. .43187,
```

```
                AL513014.9:44856. .45011,AL513014.9:46133. .46211,
                AL513014.9:47748. .47771,AL513014.9:48876. .48929,
                AL513014.9:50825. .50974,AL513014.9:53729. .53792,
                AL513014.9:62567. .62610,AL513014.9:66867. .66992,
                AL513014.9:67218. .67366,AL513014.9:68674. .68803,
                AL513014.9:74163. .74831,AL513014.9:79192. .79376,
                4770. .4932,8263. .8398,9931. .10092,12188. .12346,
                16272. .16474,18350. .18433,19587. .19677,26226. .28457)
                /locus_tag="RP23-461O24.1-003"
                /note="match: cDNAs: AK007145.1 AB002384.1
                match: ESTs: AW822704.1 BB674120.1 AA041998.1 BM144846.1
                Z42557.1 R20596.1"
CDS             join(AL513014.9:11989. .12037,AL513014.9:43061. .43187,
                AL513014.9:44856. .45011,AL513014.9:46133. .46211,
                AL513014.9:47748. .47771,AL513014.9:48876. .48929,
                AL513014.9:50825. .50974,AL513014.9:53729. .53792,
                AL513014.9:62567. .62610,AL513014.9:66867. .66992,
                AL513014.9:67218. .67366,AL513014.9:68674. .68803,
                AL513014.9:74163. .74831,AL513014.9:79192. .79376,
                4770. .4932,8263. .8398,9931. .10092,12188. .12346,
                16272. .16474,18350. .18433,19587. .19677,26226. .26326)
                /locus_tag="RP23-461O24.1-003"
                /standard_name="OTTMUSP00000000566"
                /note="match: proteins: Q9D9C2 Q9Y4F9"
                /codon_start=1
                /protein_id="CAI26080.1"
                /db_xref="GI:56237839"
                /translation="MPPGTKRLRALGAFSAGLPTRLPEIMLVGSQSFSPGGPNGIIRS
                QSFAGFSGLQERRSRCNSFIENASALKKPQAKLKKMHNLGHKNNNTPKEPQPKRVEEV
                YRALKNGLDEYLEFHQTELDKLTAQLKDMKRNSRLGVLYDLDKQIKTIERYMRRLEFH
                ISKVDELYEAYCIQRRLQDGASKMKQAFATSPASKAARESLSEINRSYKEYTENMCAI
                EAELESLLGEFSIKMKGLAGFARLCPGDQYEIFMKYGRQRWKLKGKIEVNGKQSWDGA
                ETVFLPLIVGFISIKVTELKGLATHILVGSVTCETKELFAARPQVVAVDINDLGTIKL
                NLEITWYPFDVEDTTPSSGPGNKTAALQRRMSMYSQGTPETPTFKDQSFFSNLPDDIF
                ESGKAAEEKRPLSLSFSDLQDGDCVFTSSSATSPSSSHSAHPEITITPAELTHSSLSS
                QNEGTEDSSSASSRNSLGEDHEPKSHPKSDTVEPGKPGVATRSGTESLFLESSVAEAL
                LQESDEASELKPVELDTFEGNITKQLVKRLTSAEGPVTTDKLFFEGSVGSESEAGRSF
                LDGSLEDAFNGLFLALDPHKKQYKEFQDLNQEVTHLDDVLKCKPAGSRSRSSSLSLTV
                ESALESFDFLNTSDFDEEEEDGDDVCHVGGGADSVFSDTETEKSGSVHPEARGHLSEA
                LTEDTGVGTSVAGSPLPLTTGNESLDITIVKHLQYCTQLIQQIVFSSKTPFVARSLLE
                KLSRQVLVMQKLAAVSDENLGNITSVVEAIPEFHKKLSLLAFWTKCCSPSGVYHSSAA
                RLIKQLEASFARSINKDYPGLAEPVFRTLVSQILDRAEPLLSSSLSSEVITVFQYYSF
                FTSHGVSDLETYLGQLTRQVAMVQTLQSLRDEKLLQTMSDLAPSNLPAQQEVLRTLAL
                LLTRDDNEVSEAVTLYLAAASKNEHFREKALLYYCEALTKANLQLQKAACLALKSLEA
                TESIKMLVTLCQSDTEEIRTVASETLLSLGEDGRLAYEQLDKFPRDCVKVGGRHGTEV
                ATAF"
polyA_signal    26424. .26429
                /locus_tag="RP23-461O24.1-003"
polyA_site      26446
                /locus_tag="RP23-461O24.1-003"
polyA_site      26447
                /locus_tag="RP23-461O24.1-003"
polyA_site      26457
                /locus_tag="RP23-461O24.1-003"
gene            join(35873. .36201,40700. .41233)
                /locus_tag="RP23-92G13.3-001"
mRNA            join(35873. .36201,40700. .41233)
                /locus_tag="RP23-92G13.3-001"
                /note="match: cDNAs: AK006029.1
                match: ESTs: AV045211 BF018417 BB012811 BB014319 BB020000
                AV263251 BP319523 AV206925 AV041162 AV259155 AV041846"
polyA_signal    39211. .39216
                /locus_tag="RP23-92G13.3-001"
polyA_site      39233
                /locus_tag="RP23-92G13.3-001"
polyA_site      complement(44486)
polyA_site      complement(44487)
polyA_site      complement(44488)
polyA_signal    complement(44511. .44516)
polyA_signal    complement(44515. .44520)
gene            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56454. .56564))
                /gene="Gmnn"
mRNA            /locus_tag="RP23-92G13.2-001"
                complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56454. .56564))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-001"
                /note="match: ESTs: BI789536.1 BQ178448.1 BF722692.1
                AA510015.1 BG797776.1 BB868504.1 AA014584.1 AA032471.1
                BI454155.1 BE448331.1 W13399.1 BF556131.1 BG085437.1
                AI528457.1 AA277821.1 AA681520.1 W99871.1 AA004051.1"
gene            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56228. .56502))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-002"
mRNA            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56228. .56502))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-002"
                /note="match: cDNAs: AF068780.1 BC020061.1
                match: ESTs: BF160648.1 AA437903.1"
CDS             complement(join(46710. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54851))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-001"
                /standard_name="OTTMUSP00000000569"
                /note="match: proteins: O88513 O75496"
                /codon_start=1
                /protein_id="CAI26081.1"
                /db_xref="GI:56237840"
                /translation="MNLSMKQKQEGAQENVKNSPVPRRTLKMIQPSADGSLVGRENEL
                PKGLFKRKLWDDQLASQTSSCGPEANENKDVGDLTQEAFDLISKENPSSQYWKEVAEQ
                RRKALYEALKENEKLHKEIEQKDSEIARLRKENKDLAEVAEHVQYMAEVIERLSNEPL
                DNFESPDSQEFDSEEEAVEYSELEDSGAGTCAEETVSSSTDARPCT"
gene            96664. .108654
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
mRNA            join(96664. .98150,99423. .99599,100503. .100625,
                104527. .104661,107278. .108654)
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
                /note="match: cDNAs: BC002217.1 BC005537.1 AL136632.1
                AK022681.1 AJ420475.1
                match: ESTs: BI078093.1 BG295445.1 BI696541.1 BG717389.1
                BI108344.1 BG975946.1 BI452750.1 BG168325.1 BG295139.1
                BG914583.1 BI753721.1"
CDS             join(98022. .98150,99423. .99599,100503. .100625,
                104527. .104661,107278. .107403)
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
                /standard_name="OTTMUSP00000000563"
                /note="match: proteins: Q9H1Z2 Q99LU8 Q9GZU0 O95550"
                /codon_start=1
                /protein_id="CAI26082.1"
                /db_xref="GI:56237841"
                /translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
                KKSALFEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
                DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEFHHGDYEKQFLHV
                LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
                IEDHLRPYMPE"
polyA_signal    106607. .106612
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
polyA_site      106654
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
```

Alignment Scores:

Pred. No.:	2.43e-36	Length:	256608
Score:	536.00	Matches:	115
Percent Similarity:	40.8%	Conservative:	18
Best Local Similarity:	35.3%	Mismatches:	15
Query Match:	28.2%	Indels:	178
DB:	9	Gaps:	2

US-10-757-745-2 (1-362) x AL589699 (1-256608)

```
Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              |||::::::   |||||||||  ||||||||||||||||||||||||||||||||||||
Db     135029 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 135088

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              |||:::|||||||||:::   |||||||||  ||||||  ||||||||||||||||||
Db     135089 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 135148

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr--- 271
              :::|||   ||||||||||||||||||||||||||||||:::||||||:::
Db     135149 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTAAGTAAC 135208

Qy        271 ------------------------------------------------------------ 271

Db     135209 AGAATTGATTCAATGTTTGTGGGGAGAGTATTGCTATGGAAATTAATTTTGATTCATTTT 135268

Qy        271 ------------------------------------------------------------ 271

Db     135269 ATCCAAAGTCTTATGGACTTGAAATACAAGTTAGAAAACACTGTTGTAGCTTTTGTTTAA 135328

Qy        271 ------------------------------------------------------------ 271

Db     135329 AATATGCACAAAAGCTATTAGGAGAAATCCTCTTATATGTTAGATAAACAAGTATTAAAA 135388

Qy        271 ------------------------------------------------------------ 271

Db     135389 TAAGATAATAATTATAAGAATTCTAATTTTAAAAGTTGAGGGTATTTGTAGCATCCCTGT 135448

Qy        271 ------------------------------------------------------------ 271

Db     135449 GTATAAGAAGATCTAATTCTCTTCTGAGTTGAAATCAACTTTTTAAAATACTTTCTTGCT 135508

Qy        271 ------------------------------------------------------------ 271

Db     135509 GGGTGTGGTGGCCCATGCCTTAATCTCAGCACTTGGGAGGCGGAGACAGGCAGATCTCTG 135568

Qy        271 ------------------------------------------------------------ 271

Db     135569 AGTTACAGAGTGAATTCCAGCACAGCTCAGATACACAGAAACCCTGTCTTGAAAAACAAC 135628

Qy        271 ------------------------------------------------------------ 271

Db     135629 AACAAAAGCTGTACTTACAAATGATAATATCAGACACCTTTTAAATTGGGGGCAGGGAGT 135688

Qy        272 ----------------------------------------------ArgCysGlyGlyLeu 276
                                                             :::||||||||||||
Db     135689 AAGTATATAAACCTTTCATAATATTCTTTCTCTTGCCAGGTTATCAAATGTGGTGGTTTA 135748

Qy        277 ProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 296
              |||:::|||:::   |||  ||||||||||||||||||||||||||||||||||||||
Db     135749 CCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACA 135808

Qy        297 TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAsp 316
              |||||||||:::   |||:::||||||   |||   ||||||   |||   ||||||||
Db     135809 TGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGAT 135868

Qy        317 ArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeu 336
              ||||||||||||||      |||||||||||||||:::||||||:::||||||||||||
Db     135869 CGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTT 135922

Qy        337 LeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCys 356
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135923 GTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGC 135982

Qy        357 AsnLeuAspIleIleLeu 362
                 |||::::::::::|||
Db     135983 ACCTTGAATGTAGTATTG 136000
```

RESULT 32
AY613922

```
LOCUS       AY613922               20938 bp    DNA     linear   PRI 11-MAY-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein (TTRAP) gene,
            complete cds.
ACCESSION   AY613922
VERSION     AY613922.1  GI:47060296
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 20938)
  AUTHORS   Rieder,M.J., Daniels,R.L., da Ponte,S.H., Hastings,N.C.,
            Ahearn,M.O., Rajkumar,N., Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2004) Genome Sciences, University of Washington,
            1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  SeattleSNPs. NHLBI HL66682 Program
            for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.gs.washington.edu).
FEATURES             Location/Qualifiers
     source          1. .20938
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     repeat_region   1. .32
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     variation       81
                     /frequency="0.67"
                     /replace=""
     repeat_region   87. .368
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     repeat_region   369. .673
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       449
                     /frequency="0.96"
                     /replace=""
     variation       540
                     /frequency="0.01"
                     /replace="t"
     variation       586
                     /frequency="0.01"
                     /replace="a"
     variation       650
                     /frequency="0.03"
                     /replace="g"
     variation       658. .660
                     /frequency="0.01"
                     /replace=""
     repeat_region   683. .989
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       695
                     /frequency="0.18"
                     /replace="t"
     variation       922
                     /frequency="0.02"
                     /replace="t"
     variation       1319
                     /frequency="0.15"
                     /replace="t"
     variation       1456
                     /frequency="0.19"
                     /replace="a"
     variation       1532
                     /frequency="0.01"
                     /replace="c"
     variation       1711. .1713
                     /frequency="0.04"
                     /replace=""
     variation       1795
```

```
                /frequency="0.10"
                /replace="a"
     variation  1821
                /frequency="0.02"
                /replace="t"
     variation  1830. .1862
                /frequency="0.04"
                /replace=""
     variation  1913. .1914
                /frequency="0.13"
                /replace=""
     gene       <2005. .18938
                /gene="TTRAP"
     mRNA       join(<2005. .2194,2281. .2366,10172. .10345,11003. .11094,
                14376. .14494,15753. .15923,17846. .18938)
                /gene="TTRAP"
                /product="TRAF and TNF receptor associated protein"
     CDS        join(2030. .2194,2281. .2366,10172. .10345,11003. .11094,
                14376. .14494,15753. .15923,17846. .18127)
                /gene="TTRAP"
                /codon_start=1
                /product="TRAF and TNF receptor associated protein"
                /protein_id="AAT09764.1"
                /db_xref="GI:47060297"
                /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
     variation  2381
                /gene="TTRAP"
                /frequency="0.08"
                /replace="a"
     variation  2767
                /gene="TTRAP"
                /frequency="0.08"
                /replace="g"
     variation  2800
                /gene="TTRAP"
                /frequency="0.79"
                /replace=""
     repeat_region 2919. .3323
                /rpt_family="L2"
                /rpt_type=dispersed
     variation  2927
                /gene="TTRAP"
                /frequency="0.13"
                /replace="g"
     variation  3004
                /gene="TTRAP"
                /frequency="0.04"
                /replace="t"
     variation  3061
                /gene="TTRAP"
                /frequency="0.13"
                /replace="t"
     variation  3062
                /gene="TTRAP"
                /frequency="0.13"
                /replace="t"
     variation  3114
                /gene="TTRAP"
                /frequency="0.21"
                /replace="c"
     variation  3418
                /gene="TTRAP"
                /frequency="0.03"
                /replace="c"
     variation  3781
                /gene="TTRAP"
                /frequency="0.15"
                /replace="g"
     variation  3988
                /gene="TTRAP"
                /frequency="0.01"
                /replace="a"
     variation  4191
                /gene="TTRAP"
                /frequency="0.01"
                /replace="c"
     variation  4393
                /gene="TTRAP"
                /frequency="0.97"
                /replace=""
     variation  4414
                /gene="TTRAP"
                /frequency="0.03"
                /replace="c"
     variation  4437
                /gene="TTRAP"
                /frequency="0.07"
                /replace="g"
     repeat_region 4446. .4530
                /rpt_family="MER1_type"
                /rpt_type=dispersed
     repeat_region 4559. .4609
                /rpt_family="MER1_type"
                /rpt_type=dispersed
     variation  4565
                /gene="TTRAP"
                /frequency="0.52"
                /replace=""
     repeat_region 4625. .4760
                /rpt_family="MER1_type"
                /rpt_type=dispersed
     variation  4658
                /gene="TTRAP"
                /frequency="0.01"
                /replace="a"
     variation  4861
                /gene="TTRAP"
                /frequency="0.02"
                /replace="a"
     variation  4941
                /gene="TTRAP"
                /frequency="0.07"
                /replace="g"
     variation  4983
                /gene="TTRAP"
                /frequency="0.07"
                /replace="t"
     variation  5024
                /gene="TTRAP"
                /frequency="0.07"
```

Alignment Scores:

Pred. No.:	5.53e-36	Length:	20938
Score:	518.00	Matches:	93
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	98.9%	Mismatches:	0
Query Match:	27.2%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2 (1-362) x AY613922 (1-20938)

```
Qy        269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17843 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 17902

Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17903 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 17962

Qy        309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17963 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACAC 18022

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18023 ATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCT 18082

Qy        349 SerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
                ||||||||||||||||||||||||||||||||||||||||
Db      18083 AGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 18124

RESULT 33
HS30M3
LOCUS       HS30M3                 102200 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP1-30M3 on chromosome 6p22.1-22.3
            Contains the 5' end of gene KIAA0319, the gene for TRAF and TNF
            receptor-associated protein (TTRAP)(EAP2, AD022, MGC9099), gene
            FLJ20501, the 3' end of gene FLJ12619 (DKFZP564G182) and two CpG
            islands, complete sequence.
ACCESSION   AL031775
VERSION     AL031775.1  GI:4071041
KEYWORDS    HTG; AD022; CpG island; DKFZP564G182; FLJ12619; FLJ20501; KIAA0319;
            TNF; TRAF; TTRAP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 102200)
  AUTHORS   Phillips,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 29, 1998 this sequence version replaced gi:4007165.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP1-30M3 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            ---------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            ----------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES              Location/Qualifiers
     source           1. .102200
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="RZPD:RPCIP704M0330"
                      /db_xref="taxon:9606"
                      /chromosome="6"
                      /map="p22.1-22.3"
                      /clone="RP1-30M3"
                      /clone_lib="RPCI-1"
     gene             join(complement(AL133264.10:4611. .6907),
                      complement(AL133264.10:2219. .2395),
                      complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                      /gene="C6orf62"
                      /locus_tag="RP1-30M3.4-001"
     mRNA             join(complement(AL133264.10:4611. .6907),
                      complement(AL133264.10:2219. .2395),
                      complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                      /gene="C6orf62"
                      /locus_tag="RP1-30M3.4-001"
                      /product="chromosome 6 open reading frame 62"
                      /note="match: ESTs: AA084220 AA093151 AA168907 AA197971
                      AA200765 AA237701 AA292313 AA295166 AA402388 AA478637
                      AA591228 AA612170 AA638148 AA647423 AA647435 AA647742
                      AA672357 AA798123 AA798467 AA815944 AA833812 AA863774
                      AI053561 AI054330 AI122327 AI156738 AI191279 AI194277
                      AI284633 AI316639 AI390324 AI458673 AI688647 AI694146
                      AI694199 AI788240 AI802319 AI807413 AI831412 AI932363
                      AI935446 AI982078 AJ396992 AJ398211 AL529211 AL563388.1
                      AL601897.1 AU051650 AU125496 AU135776 AU137791 AU138464
                      AV656408 AV686324 AV698092 AV703346 AV708342 AV708873
                      AV709145 AV709526 AV739602 AV742411 AV743612 AV744587
                      AV745018 AW050065 AW063931 AW270998 AW300107 AW300121
                      AW301984 AW302318 AW302718 AW302841 AW302861 AW303029
                      AW355281 AW822881 AW859606 AW959243 BB067326 BB123906
                      BB609972 BB624305 BB628344 BB654156 BB823708 BE007675
                      BE049029 BE049182 BE120619 BE178692 BE178693 BE244058
                      BE246922 BE287246 BE289710 BE311050 BE369398 BE377181
                      BE386264 BE465032 BE484573 BE554360 BE571717 BE625094
                      BE766876 BE766886 BE770856 BE785982.1 BE813002 BE817780
                      BE829341 BE829344 BE829355 BE829359 BE893564 BE895461
                      BE896141 BE908143 BE913030 BE945891 BE991271 BE992722
                      BF015474 BF021751 BF022025 BF029248 BF037603 BF060909
                      BF064161 BF109576 BF111215 BF122334 BF137583 BF139381
                      BF140373 BF145038 BF148549 BF150174 BF165619 BF207643
                      BF326037 BF326041 BF326042 BF326045 BF334390 BF341149
                      BF376614 BF378829 BF380539 BF380550 BF465841 BF535383
                      BF578793 BF582307 BF748822 BF748938 BF748963 BF765346
                      BF766403 BF784199 BF789335 BF789696 BF885392 BF896088
                      BF940327 BF984391 BG030354 BG035327 BG088552 BG116557
                      BG148023 BG167018 BG168325 BG247693 BG252842 BG295272
                      BG295445 BG342500 BG610013 BG655426 BG667600 BG677384.1
                      BG678705 BG692520 BG717389 BG801183 BG802707 BG803992
                      BG863302 BG914583 BG917993 BG940215 BG943615 BG975946
                      BI080618 BI092275 BI106992 BI107793 BI108344 BI222784
                      BI258149 BI326830 BI651746 BI659175 BI666111 BI688353
                      BI689042 BI694119 BI696541 BI696575 BI696638 BI733610
                      BI753721 BI754669 BI851772 BI856689.1 BI872006 BI872428
                      BI903238 BI903250 BI960823 BI986171 BM016286 BM069622
                      BM069876 BM147734 BM237580 BM238719 BM238848 BM423047
                      BM423107 F11344 F20959 H17162 H98057 N56057 R25494 T11219
                      T51727 W47233
                      match: cDNAs: AB046018.1 AJ420475.1 AK022681.1 AL136632.1
                      BC002217.1 BC005537.1"
     gene             join(complement(AL133264.10:101. .211),5119. .5253,
                      7768. .8194)
                      /gene="C6orf62"
                      /locus_tag="RP1-30M3.4-002"
     mRNA             join(complement(AL133264.10:101. .211),5119. .5253,
                      7768. .8194)
                      /gene="C6orf62"
                      /locus_tag="RP1-30M3.4-002"
                      /product="chromosome 6 open reading frame 62"
                      /note="match: ESTs: BG772628 BG772628.1"
     gene             join(complement(AL133264.10:5931. .6069),
                      complement(AL133264.10:2219. .2395),
                      complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
                      /gene="C6orf62"
                      /locus_tag="RP1-30M3.4-003"
     mRNA             join(complement(AL133264.10:5931. .6069),
                      complement(AL133264.10:2219. .2395),
                      complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
                      /gene="C6orf62"
                      /locus_tag="RP1-30M3.4-003"
                      /product="chromosome 6 open reading frame 62"
```

```
                     /note="match: ESTs: BG342500.1 BG692520.1 BG803992.1"
     CDS             join(complement(AL133264.10:4611. .4739),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-003"
                     /standard_name="OTTHUMP00000016093"
                     /note="match: proteins: O95550 Q961D0 Q99K03 Q99LU8 Q9GZU0
                     Q9H1Z2 Q9N0B2 Q9QEJ3 Q9W384 Q9XWI7"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 62"
                     /protein_id="CAI20135.1"
                     /db_xref="GI:56202469"
                     /translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
                     KKSALFEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
                     DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEFHHGDYEKQFLHV
                     LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
                     IEDHLRPYMPE"
     CDS             join(complement(AL133264.10:5931. .5972),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-003"
                     /standard_name="OTTHUMP00000016094"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 62"
                     /protein_id="CAI20136.1"
                     /db_xref="GI:56202470"
                     /db_xref="UniProt/TrEMBL:Q5JVZ2"
                     /translation="MSENLSDPVSPVVRKKKSALFEVSEVIPVMTNNYEENILKGVRD
                     SSYSLESSLELLQKDVVQLHAPRYQSMRRDVIGCTQEMDFILWPRNDIEKIVCLLFSR
                     WKESDEPFRPVQAKFEFHHGDYEKQFLHVLSRKDKTGIVVNNPNQSVFLFIDRQHLQT
                     PKNKATIFKLCSICLYLPQEQLTHWAVGTIEDHLRPYMPE"
     polyA_signal    9124. .9129
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-001"
     gene            complement(join(12316. .12571,15963. .16147,46686. .46959))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-001"
     mRNA            complement(join(12316. .12571,15963. .16147,46686. .46959))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-001"
                     /product="thioesterase superfamily member 2"
                     /note="match: ESTs: AA614605 AA928745 AI017897 AI052396
                     BF407527 BF475709 BF594161
                     match: cDNAs: AF155649.1 AF220186.1 AF274952.1 AK000508.1
                     AK002261.1 AK008624.1 BC000894.1"
     gene            complement(join(12316. .12571,15963. .16147,26337. .26628))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-002"
     mRNA            complement(join(12316. .12571,15963. .16147,26337. .26628))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-002"
                     /product="thioesterase superfamily member 2"
                     /note="match: ESTs: BG219890 BG256069"
     polyA_site      12318
     polyA_signal    12336. .12341
     CDS             complement(join(12415. .12571,15963. .16147,46686. .46766))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-001"
                     /standard_name="OTTHUMP00000016090"
                     /note="match: proteins: AAA28024 AAH18165 AAL48890 O95549
                     P34419 P87304 P93828 P95914 Q18187 Q9ASQ1 Q9NPJ3 Q9VZZ5
                     Q9VZZ6"
                     /codon_start=1
                     /product="thioesterase superfamily member 2"
                     /protein_id="CAD92509.1"

Alignment Scores:
Pred. No.:                3.55e-35        Length:          102200
Score:                    518.00          Matches:         93
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    98.9%           Mismatches:      0

Query Match:              27.2%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2 (1-362) x HS30M3 (1-102200)

Qy        269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62958 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 63017

Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63018 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 63077

Qy        309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63078 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACAC 63137

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63138 ATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCT 63197

Qy        349 SerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||||||||||||||||||||||||
Db      63198 AGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 63239

RESULT 34
CR925830
LOCUS       CR925830               167847 bp    DNA     linear   PRI 08-JAN-2005
DEFINITION  Human DNA sequence from clone RP11-195J19 on chromosome 6, complete
            sequence.
ACCESSION   CR925830
VERSION     CR925830.3  GI:56798018
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 167847)
  AUTHORS   Palmer,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 24, 2004 this sequence version replaced gi:56542150.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
```

```
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-195J19 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES              Location/Qualifiers
     source           1. .167847
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="6"
                      /clone="RP11-195J19"
                      /clone_lib="RPCI-11.1"
ORIGIN

Alignment Scores:
Pred. No.:              6.36e-35        Length:         167847
Score:                  518.00          Matches:        93
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  98.9%           Mismatches:     0
Query Match:            27.2%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2 (1-362) x CR925830 (1-167847)

Qy        269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43201 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 43260

Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43261 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 43320

Qy        309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43321 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACAC 43380

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      43381 ATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCT 43440

Qy        349 SerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||||||||||||||||||||||||
Db      43441 AGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 43482

RESULT 35
CR942205
LOCUS       CR942205               176819 bp    DNA     linear   PRI 29-APR-2005
DEFINITION  Human DNA sequence from clone RP11-948M1 on chromosome 6, complete
            sequence.
ACCESSION   CR942205
VERSION     CR942205.4  GI:62989869
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 176819)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:62551221.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-948M1 is from the library RPCI-11.4 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES              Location/Qualifiers
     source           1. .176819
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="6"
                      /clone="RP11-948M1"
                      /clone_lib="RPCI-11.4"
ORIGIN

Alignment Scores:
Pred. No.:              6.76e-35        Length:         176819
Score:                  518.00          Matches:        93
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  98.9%           Mismatches:     0
Query Match:            27.2%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2 (1-362) x CR942205 (1-176819)

Qy        269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37184 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 37243

Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37244 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 37303

Qy        309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37304 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACAC 37363

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37364 ATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCT 37423

Qy        349 SerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||||||||||||||||||||||||
Db      37424 AGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 37465

RESULT 36
AC125756/c
LOCUS       AC125756               245394 bp    DNA     linear   HTG 13-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-42H3, WORKING DRAFT SEQUENCE, 2
            unordered pieces.
ACCESSION   AC125756
VERSION     AC125756.5  GI:30578987
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
```

SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 245394)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 245394)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 245394)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On May 13, 2003 this sequence version replaced gi:23265926.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GJVU
Center clone name: CH230-42H3
---------- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 238380 bases at least Q40
Consensus quality: 239660 bases at least Q30
Consensus quality: 240760 bases at least Q20
Estimated insert size: 255581; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 242009: contig of 242009 bp in length
* 242010 242109: gap of unknown length
* 242110 245394: contig of 3285 bp in length.

FEATURES Location/Qualifiers

source	1. .245394	/organism="Rattus norvegicus" /mol_type="genomic DNA" /db_xref="taxon:10116" /clone="CH230-42H3"
misc_feature	1. .1025	/note="wgs_end_extension clone_end:T7"
misc_feature	1076. .2080	/note="wgs_end_extension clone_end:T7"
misc_feature	3723. .4536	/note="clone_boundary clone_end:T7 site:EcoRI end_sequence:BH348771"
misc_feature	128656. .128715	/note="clone_boundary clone_end:Sp6 site:EcoRI end_sequence:BH260088"
misc_feature	236171. .238412	/note="wgs_end_extension clone_end:Sp6"
misc_feature	239605. .242009	/note="wgs_end_extension clone_end:Sp6"
gap	242010. .242109	/estimated_length=unknown

ORIGIN

Alignment Scores:
Pred. No.: 6.44e-29 Length: 245394

```
Score:                  454.00          Matches:        113
Percent Similarity:     27.0%           Conservative:   20
Best Local Similarity:  23.0%           Mismatches:     15
Query Match:            23.9%           Indels:         344
DB:                     14              Gaps:           2

US-10-757-745-2 (1-362) x AC125756 (1-245394)

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              |||::::::   |||||||||  ||||||||||||||||||||||||||||||||
Db     141402 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACAAGAAAA 141343

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              |||:::|||||||||:::|||||||||||||   |||   :::|||||||||||||||
Db     141342 CATTCTGCTGAACGAATAAATCAATTAAAAACTGTTTTTCAAAAAATGCAAGAGGCTACA 141283

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr--- 271
              :::|||  ||||||||||||||||||||||||||||||||:::|||||||:::
Db     141282 GATTCAACTACTGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTAAGTAAC 141223

Qy        271 ------------------------------------------------------------ 271

Db     141222 AGAATTGATTCAGTGTTTGTGGGGAGAGTATTGCTATTTAATTTGGATTCATTTTATTCA 141163

Qy        271 ------------------------------------------------------------ 271

Db     141162 AGGTATTATGGACTTGAAATACAAGGTAGGAAACACTGTTGTATCTTTTGTATAAAATGT 141103

Qy        271 ------------------------------------------------------------ 271

Db     141102 GCACAAAAGAGGCCTTTACACTATTAGGGGAAACCCTGCTATATTATATTAGTTAGATAA 141043

Qy        271 ------------------------------------------------------------ 271

Db     141042 ATGAGTATTAGAATAAGATAAGAATTATAAGAATTTTAATTAAAGTTAAAATAGGGTATT 140983

Qy        271 ------------------------------------------------------------ 271

Db     140982 TGTAGCATTCCTACGTATAAGAAGACCTAATCCTCTTATGACTTGAAATCAACTTTTTAA 140923

Qy        271 ------------------------------------------------------------ 271

Db     140922 AATACTTTATTGCTGGGTGCTGTGGCCCGTGCCTTTAATCATAGCACTTGGGTGTCAGAG 140863

Qy        271 ------------------------------------------------------------ 271

Db     140862 ACAGGCAGATCTCTGAATTCAAGGCCAGCCTGATATGCAGAGTGAATTGTAGTACAGCCA 140803

Qy        271 ------------------------------------------------------------ 271

Db     140802 GAATTACACAGAAACCCTGTCTTGAAAAACAATAGTCCTGAGAGACACATCCAGAGCATG 140743

Qy        271 ------------------------------------------------------------ 271

Db     140742 TCCAATACAGAGGTCAGTGCTAGCAGCAAACCACCGAACTGAGAACAGGACCCCCTTGGG 140683

Qy        271 ------------------------------------------------------------ 271

Db     140682 GGGAATTAGAGAAAGGACTGAAAGAGCTAAAGGGGCTTGCAACCCCATAAAAACAACAAT 140623

Qy        271 ------------------------------------------------------------ 271

Db     140622 GCCAACCAACCAGAGCTTCCAGGGACTAAACCACTACGGAAAGACTATATGTGGACTGAC 140563

Qy        271 ------------------------------------------------------------ 271

Db     140562 CCAGGGCTCCAACTGCATATGTAGCAGAGAATAGCCTTGTTGGGGCACCAGTGGAAGGGG 140503

Qy        271 ------------------------------------------------------------ 271

Db     140502 AAGCCCTTGGTCCTGCCAAGGTTAGACCCCCAGTGCAGGGGATTTGGGGGGGGGGCAATA 140443

Qy        271 ------------------------------------------------------------ 271

Db     140442 AGGGAAATGTATAGGGGAATACCAGTATGGGAGCAGGAGGAGGGAATGGGGACTTATGGA 140383

Qy        271 ------------------------------------------------------------ 271

Db     140382 CAGGAAACTGGGAAGGGGAATGACATTTGAAATGTAAGTAAAGAAATATATCTAATAAAC 140323

Qy        271 ------------------------------------------------------------ 271

Db     140322 AATTAAAAAAAAAGAAAAACAACAAAAGTTGCACTTACAAGTGATAATATCAAACACCTTT 140263

Qy        271 ------------------------------------------------------------ 271

Db     140262 TAAGTTGAGGTGGGGAGTAAATATATAAACCTTACATAATATTTTTTCTCTTGCCAGGTT 140203

Qy        272 ---ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 290
                  :::|||||||||||||||:::|||:::   |||   |||||||||||||||||||
Db     140202 ATCAAATGTGGTGGTTTACCTGACAATGTTTTTGATGCCTGGGAATTTTTGGGCAAACCT 140143

Qy        291 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 310
              |||||||||:::|||||||||||||||:::   |||   ||||||   |||   ||||||
Db     140142 AAGCATTGCCGGTATACATGGGATACGAAAGCAAACGATAATCTCAGGATTCCTGCCGCT 140083

Qy        311 CysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 330
              ||||||   |||||||||||||||||||||||||||||||||||||||||||:::|||
Db     140082 TGTAAGCATCGTTTTGATCGAATATTCTTCAGA------GCAGAAGAGGGACACCTTATT 140029

Qy        331 ProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAsp 350
              |||:::|||||||||||||:::|||||||||:::|||||||||||||||||||||||||
Db     140028 CCCCAAAGTTTAGACCTAATTGGGTTGGAAAGACTGGACTGTGGTAGATTTCCTAGTGAT 139969

Qy        351 HisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||||||||   |||:::::::::|||
Db     139968 CACTGGGGGCTCCTGTGTACCTTGAATGTAGTATTG 139933
```

```
RESULT 37
BD077628
LOCUS       BD077628                258 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  5'EST of secreted protein expressed in muscles and other mesodermal
            tissues.
ACCESSION   BD077628
VERSION     BD077628.1  GI:22623231
KEYWORDS    JP 2001512016-A/214.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 258)
  AUTHORS   Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
  TITLE     5'EST of secreted protein expressed in muscles and other mesodermal
            tissues
  JOURNAL   Patent: JP 2001512016-A 214 21-AUG-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001512016-A/214
            PD   21-AUG-2001
            PF   31-JUL-1998 JP 2000505295
            PR   01-AUG-1997 US    08/905134
            PI   JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO  PI
            LACROIX
            PC   C12N15/09,C12N15/09,C07K14/47,C12M1/00,C12N15/00,C12N15/00 CC
             blastn
            CC   identity 97
            CC   region 1. .180
            CC   id T31666
            CC   est
            CC   blastn
            CC   identity 92
            CC   region 88. .141
            CC   id R58665
            CC   est
            CC   blastn
```

```
          CC   identity 90
          CC   region 39. .93
          CC   id R58665
          CC   est
          CC   blastn
          CC   identity 98
          CC   region 63. .137
          CC   id R14990
          CC   est
          CC   blastn
          CC   identity 94
          CC   region 1. .50
          CC   id R14990
          CC   est
          CC   blastn
          CC   identity 99
          CC   region 1. .120
          CC   id T26956
          CC   est
          CC   Von Heijne matrix
          CC   score 6
          CC   seq LCXEFXSVASCDA/AV
          FH   Key             Location/Qualifiers
          FT   misc_feature    75. .254
          FT   misc_feature    73. .126
          FT   misc_feature    23. .77
          FT   misc_feature    157. .231
          FT   misc_feature    95. .144
          FT   misc_feature    135. .254
          FT   sig_peptide     31. .150.
FEATURES             Location/Qualifiers
     source          1. .258
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.:                2.5e-25        Length:          258
Score:                    376.00         Matches:         72
Percent Similarity:       96.0%          Conservative:    0
Best Local Similarity:    96.0%          Mismatches:      3
Query Match:              19.8%          Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2 (1-362) x BD077628 (1-258)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 90

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||  ||||||  ||||||||||||||||||||
Db         91 GAGGTGAAAAAGCGGCGACTTCTGTGTSTRGAGTTTRCCTCGGTCGCAAGCTGCGATGCC 150

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GCAGTGGCTCAGTGCTTCCTGGCCGAKAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 210

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 75
              |||||||||||||||||||||||||||||||||||||||||||||
Db        211 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCA 255

RESULT 38
AX939201
LOCUS       AX939201                  258 bp    DNA     linear   PAT 07-JAN-2004
DEFINITION  Sequence 220 from Patent EP1367124.
ACCESSION   AX939201
VERSION     AX939201.1  GI:40733556
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   dumas milne Edwards,J.B., Duclert,A. and Lacroix,B.
  TITLE     5' ests for secreted proteins expressed in muscle and other
            mesodermal tissues
  JOURNAL   Patent: EP 1367124-A 220 03-DEC-2003;
            GENSET (FR)
FEATURES             Location/Qualifiers
     source          1. .258
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /tissue_type="KIDNEY"
                     /dev_stage="FETAL"
     sig_peptide     31. .150
ORIGIN

Alignment Scores:
Pred. No.:                2.5e-25        Length:          258
Score:                    376.00         Matches:         72
Percent Similarity:       96.0%          Conservative:    0
Best Local Similarity:    96.0%          Mismatches:      3
Query Match:              19.8%          Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2 (1-362) x AX939201 (1-258)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 90

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||  ||||||  ||||||||||||||||||||
Db         91 GAGGTGAAAAAGCGGCGACTTCTGTGTSTRGAGTTTRCCTCGGTCGCAAGCTGCGATGCC 150

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GCAGTGGCTCAGTGCTTCCTGGCCGAKAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 210

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 75
              |||||||||||||||||||||||||||||||||||||||||||||
Db        211 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCA 255

RESULT 39
BV203236
LOCUS       BV203236                  201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm211480 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV203236
VERSION     BV203236.1  GI:48172559
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
```

```
          STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN

Alignment Scores:
Pred. No.:                1.5e-21        Length:          201
Score:                    333.00         Matches:         66
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              17.5%          Indels:          0
DB:                       10             Gaps:            0

US-10-757-745-2 (1-362) x BV203236 (1-201)

Qy        182 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 62

Qy        202 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCys 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTTTGC 122

Qy        222 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTA 182

Qy        242 LysMetValLeuLysLys 247
              ||||||||||||||||||
Db        183 AAAATGGTTTTAAAGAAA 200

RESULT 40
AX351374
LOCUS       AX351374                176 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 121 from Patent WO0196390.
ACCESSION   AX351374
VERSION     AX351374.1  GI:18616721
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196390-A 121 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .176
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                1.94e-18       Length:          176
Score:                    298.00         Matches:         58
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              15.7%          Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2 (1-362) x AX351374 (1-176)

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 62

Qy        164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 122

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176

RESULT 41
AX339803/c
LOCUS       AX339803                179 bp    DNA     linear   PAT 10-JAN-2002
DEFINITION  Sequence 50 from Patent WO0196388.
ACCESSION   AX339803
VERSION     AX339803.1  GI:18135784
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Harlocker,S.L. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196388-A 50 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .179
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                1.98e-18       Length:          179
Score:                    298.00         Matches:         58
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              15.7%          Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2 (1-362) x AX339803 (1-179)

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy        164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 42
HS133N16S/c
LOCUS       HS133N16S               411 bp    DNA     linear   STS 02-SEP-1999
DEFINITION  H.sapiens STS from genomic clone 133N16, sequence tagged site.
ACCESSION   AL110544
VERSION     AL110544.1  GI:5830260
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   Mungall,A.J., Huckle,E., Taylor,R. and Hunt,S.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-1999) E-mail contact: humquery@sanger.ac.uk
```

```
COMMENT     Marker stdJ133N16SP6 (Primer A : TGTTCCCTCTTTGCTGCAG; Primer B :
            AACAGTTCAAGGTTGTGGGC; amplimer size : 127 bp) is from sequence
            generated from the SP6 end of PAC 133N16. 133N16 is part of the
            bacterial clone contigs constructed by the Chromosome 6 Mapping
            Group.(http://www.sanger.ac.uk/HGP/Chr6/) 133N16 is from the
            library  constructed at the Roswell Park Cancer Institute by the
            group of Pieter de Jong.
            For further details see http://bacpac.med.buffalo.edu/.
FEATURES             Location/Qualifiers
     source          1. .411
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="133N16"
ORIGIN

Alignment Scores:
Pred. No.:                2.04e-16      Length:          411
Score:                    280.50        Matches:         59
Percent Similarity:       82.4%         Conservative:    2
Best Local Similarity:    79.7%         Mismatches:      10
Query Match:              14.7%         Indels:          3
DB:                       10            Gaps:            1

US-10-757-745-2 (1-362) x HS133N16S (1-411)

Qy          89 AsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGln 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         411 AATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAA 352

Qy         109 GluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsn 128
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         351 GAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAAT 292

Qy         129 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 148
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         291 CTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTAAGTATTATCACTTCTA 232

Qy         149 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 162
               |||  |||  |||  |||  |||
Db         231 TTTAATAGTAATGTG---------TCTTATGTATATACAAAG 199

RESULT 43
BX957279
LOCUS       BX957279               198752 bp    DNA     linear   VRT 29-APR-2005
DEFINITION  Zebrafish DNA sequence from clone DKEY-218N20 in linkage group 19,
            complete sequence.
ACCESSION   BX957279
VERSION     BX957279.13  GI:62989971
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 198752)
  AUTHORS   Clark,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:61673710.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-218N20
            is from a Zebrafish BAC library
            VECTOR: pIndigoBAC-5.
FEATURES             Location/Qualifiers
     source          1. .198752
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-218N20"
                     /clone_lib="DanioKey"
ORIGIN

Alignment Scores:
Pred. No.:                5.38e-13      Length:          198752
Score:                    277.50        Matches:         73
Percent Similarity:       43.3%         Conservative:    40
Best Local Similarity:    28.0%         Mismatches:      81
Query Match:              14.6%         Indels:          67
DB:                       5             Gaps:            5

US-10-757-745-2 (1-362) x BX957279 (1-198752)

Qy         159 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178
               |||:::|||  |||:::  :::  :::  :::  :::  :::  |||
Db      147680 AGTTTTTTAGATAAACAAGAATGTTCACTTGCCATGTTTCTTAAATATCTGCAAACATAT 147739

Qy         179 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198
               |||||||||  |||||||||                          :::
Db      147740 TAATGAATTATTATGCTTTAAAAGAGTCGAAAA--------------------CTTACA 147778

Qy         199 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal-------------- 213
               :::  :::  |||::::::  :::||||||  :::::::::
Db      147779 TACAGCACCTTTAAGGTCATTATTAGCCTCCTTAAGCTATATATTTTTTCGATAATCCAC 147838

Qy         214 ------AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
                     |||                 |||||||:::|||                |||
Db      147839 AGAACAAACCACTGTTATACAATTACTTGCCTAATTACCCCAACTTGCCTAGTTACCCTA 147898

Qy         232 GlyHisAlaAlaGluArgMetAsn---------------------------------- 239
                  :::  :::  :::|||
Db      147899 ATAAACATAGTTAAGCCTTTAAATGTAACTTTAAACTGAATACTAGTATCTTGAAAAATA 147958

Qy         239 ---------------------------------------------------------- 239

Db      147959 TCTAGTCAAATATTATTTATTGTCATCATGGTAAAATAAAAGATAAAATAAATCAGTTAT 148018

Qy         240 ----------------------GlnLeuLysMetValLeuLysLysMetGlnGluAla 251
                                    |||:::  |||||||||  :::
Db      148019 TAGAAATGAGTTATTATCACTATTATGTTAAGAAATGTGTTGAAAAATCTCTCCATTAAA 148078
```

```
Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsn---------LeuArgAspArg 268
              :::             |||:::  |||      |||         |||  ::::::
Db     148079 CAGAAATTGGGGAATAATATATACAGGGGGGCTAATAATTGACTTCAACTGTATAATAAA 148138

Qy        269 GluValThrArgCys----------------------------GlyGlyLeuProAsn 278
              :::      |||                              ||||||||||||
Db     148139 GGAATGTTCGTGTGTGTATGTGTGTGTGTAGGTGGCCAATGTTGGAGGTTTGCCTGCA 148198

Qy        279 AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAsp 298
              :::  ||||||||||||  |||||||||  :::||||||:::||||||||||||
Db     148199 GGTGTGTGTGATGTGTGGGAACAGCTGGGCAAACAAGAACACTGCAGGTACACCTGGGAC 148258

Qy        299 ThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 318
              |||:::  |||||||||      :::          :::  ||||||||||||||||
Db     148259 ACCAAAGCTAACAGCAATAAGACTGTGCCTTACGTCAGCAGATGTCGCTTCGACCGCATA 148318

Qy        319 PhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGly 338
              |||  |||:::|||          :::  |||      :::  |||:::|||
Db     148319 TTCCTGAGGTCTGCTAAAACCGCTCCGCCGGTCACTCCTGACCACATGGCCTTGATTGGA 148378

Qy        339 LeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358
              :::||||||||||||||||||:::  ||||||||||||:::  |||
Db     148379 ATGGAGAAACTGGACTGCGGTCGTTATACCAGTGATCACTGGGGTATTTACTGCACTTTC 148438

Qy        359 Asp 359
              :::
Db     148439 AAC 148441
```

```
RESULT 44
BX511258
LOCUS       BX511258              200076 bp    DNA     linear   VRT 02-APR-2004
DEFINITION  Zebrafish DNA sequence from clone CH211-81E5 in linkage group 19,
            complete sequence.
ACCESSION   BX511258
VERSION     BX511258.6  GI:42820869
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 200076)
  AUTHORS   Heath,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 25, 2004 this sequence version replaced gi:42538803.
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-81E5
            is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .200076
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-81E5"
                     /clone_lib="CHORI-211"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              6.69e-13        Length:         200076
Score:                  276.50          Matches:        54
Percent Similarity:     56.8%           Conservative:   21
Best Local Similarity:  40.9%           Mismatches:     44
Query Match:            14.5%           Indels:         13
DB:                     5               Gaps:           2

US-10-757-745-2 (1-362) x BX511258 (1-200076)

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              |||:::  |||||||||  :::          :::          |||:::
Db      90172 TTAAGAAATGTGTTGAAAAATCTCTCCATTAAACAGAAATTGGGGAATAATATATACAGG 90231

Qy        261 GlyAspThrAsn---------LeuArgAspArgGluValThrArgCys------------ 273
              |||      |||         |||   ::::::  :::      |||
Db      90232 GGGGCTAATAATTGACTTCAACTGTATAATAAAGGAATGTTCGTGTGTGTATGTGTGTGT 90291

Qy        274 -----------------GlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
                              ||||||||||||      :::  ||||||||||||  |||
Db      90292 GTGTAGGTGGCCAATGTTGGAGGTTTGCCTGCAGGTGTGTGTGATGTGTGGGAACAGCTG 90351

Qy        288 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 307
              ||||||  :::||||||:::||||||||||||||:::  |||||||||      :::
Db      90352 GGCAAACAAGAACACTGCAGGTACACCTGGGACACCAAAGCTAACAGCAATAAGACTGTG 90411

Qy        308 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 327
                    :::  ||||||||||||||||||  |||:::|||
Db      90412 CCTTACGTCAGCAGATGTCGCTTCGACCGCATATTCCTGAGGTCTGCTAAAACCCGTCCA 90471

Qy        328 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 347
              :::  |||      :::  |||:::|||:::|||||||||||||||||||||:::
Db      90472 GCGGTCACTCCTGACCACATGGCCTTGATTGGAATGGAGAAACTGGACTGCGGTCGTTAT 90531

Qy        348 ProSerAspHisTrpGlyLeuLeuCysAsnLeuAsp 359
              ||||||||||||||:::  |||      :::
Db      90532 ACCAGTGATCACTGGGGTATTTACTGCACTTTCAAC 90567
```

```
RESULT 45
AP006486_2/c
WPCOMMENT
  Sequence split into 6 fragments  LOCUS AP006486 Accession AP006486
     Fragment Name      Begin     End
     AP006486_0             1     110000
     AP006486_1        100001     210000
     AP006486_2        200001     310000
     AP006486_3        300001     410000
     AP006486_4        400001     510000
     AP006486_5        500001     513790
Continuation (3 of 6) of AP006486 from base 200001 (AP006486 Cyanidioschyzon merolae str

Alignment Scores:
Pred. No.:              3.31e-12        Length:         110000
```

```
Score:                 265.50        Matches:       79
Percent Similarity:    43.9%         Conservative:  51
Best Local Similarity: 26.7%         Mismatches:    111
Query Match:           14.0%         Indels:        55
DB:                    14            Gaps:          10

US-10-757-745-2 (1-362) x AP006486_2 (1-110000)

Qy        104 GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGly 123
              |||     |||      :::   |||        |||:::||||||||||:::|||
Db      35554 GAGCATAATCAGCTTCTTTCCATGTCGTATATACGGCTGGTAACATGGAACGTTGAC--- 35498

Qy        124 LeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyr 143
              |||                |||||      :::      :::
Db      35497 TTCGACTCTCCGTTTAACTACGAGCGGTTTACAGCTATTGCCGAAGAGGTCCTGGGTGCA 35438

Qy        144 SerProAspValIlePheLeuGlnGluValIlePro---ProTyrTyrSerTyr------ 160
              |||   |||:::      |||||||||        |||:::   |||:::
Db      35437 AAACCCGGCGTGGTGCTTTTTCAAGAAGTCTCTGATTTACCGTTCTCCTCTTGGCAGGGC 35378

Qy        160 ------------------------------------------------------------ 160

Db      35377 GAGAAAGAGCAGTCGACGATTACGTCCGCTCGTCAGCATGATTCTTGTTACAAGGTCCTC 35318

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGly-----------HisGluGlu 176
              |||::::::      :::   ||||||   |||                   |||:::
Db      35317 TTGCGTCGTCTTGAGGCTCTATATGAGCCAATTGTTGATCTCGGCTTACCCTCAGAGCAG 35258

Qy        177 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 196
              ||||||              :::   :::|||:::   ::::::          |||
Db      35257 GCATACTTCGTGATGGCGTTCGTTGACCGAAGCAAGTATCGAGTTCACGGATACGAATAC 35198

Qy        197 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 216
                |||   ::::::        |||  ||||||          :::  :::
Db      35197 ACTGAATTTGAAAACTCTTCACAGGGAAGAGGTCTGCTCGCA------ATACAGCTCGAG 35144

Qy        217 GlyAsnGluLeuCys-----------------LeuMetThrSerHisLeuGluSerThr 230
              ::::::|||                    ::::::|||||||||||||||||
Db      35143 CCCAGAGATATGTGCGGAGATTTCAAGTTTCGAGTGCTTACATCACACCTCGAATCCGGA 35084

Qy        231 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 250
                  :::     |||:::   |||::::::   :::      :::|||   |||
Db      35083 ACTGCCTGTTCGGAAGTCCGAGTTCAGCAGATGAGGCAAATCGCGTATGAGATGCTCGAG 35024

Qy        251 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu--- 269
                  |||      |||||||||||||||||||    |||::::::|||
Db      35023 GGTGATACCCCGGCA------ATATTTGCAGGAGATACCAATTTTCGAGAGAAAGAAAAA 34970

Qy        270 --------------------ValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 282
                                   |||                 :::   |||
Db      34969 GCAGGATTTGAGAAGGAAATCGCTAAAGCGATCGGCTCAGAAGAGATGCGACTTTTTGAC 34910

Qy        283 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrThrTrpAspThrGlnMetAsn 302
              ::::::     :::|||:::|||   |||   ::::::|||||||||||:::   |||
Db      34909 CTATTTGTTGAGGTAGGGGAGCCTTCACATTTGCGTTTCACCTGGGACACACGCACAAAC 34850

Qy        303 SerAsn------------LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 318
              :::|||           |||              :::   |||||||||::::::
Db      34849 ACGAACAAAGTCAAGGAGTTGCCACCTCGGTTCAGGCCGCGGGCTCGTTTCGACCAAGTG 34790

Qy        319 PhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGly 338
              |||             :::                 ::::::   |||||||||
Db      34789 CTTTTCACTAGGCCTTTTCAA--------------GCGAACGCCATGCGGCTTTTGGGT 34745

Qy        339 LeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeu 354
              |||:::|||              |||||||||||||:::|||:::
Db      34744 CGAGAGCGACTCGGGCATGCTGTATTTCCATCTGACCATTACGGGATT 34697

RESULT 46
BT000599
LOCUS       BT000599               1326 bp    mRNA    linear   PLN 25-SEP-2002
DEFINITION  Arabidopsis thaliana unknown protein  (At1g11800/F25C20_3) mRNA,
            complete cds.

ACCESSION   BT000599
VERSION     BT000599.1  GI:23308396
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1326)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hayashizaki,Y.,
            Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
            Nguyen,M., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
            Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1326)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hayashizaki,Y.,
            Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
            Nguyen,M., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
            Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-SEP-2002) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
            Chen,H., Kim,C.J., Shinn,P., Banh,J. Bowser,L., Chan,M.M.,
            Chang,E., Dale,J.M., Deng, J.M., Goldsmith,A.D., Jones,T.,
            Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamada,K., Yamamura, Y., Yu,G., Davis,R.W.,
            Theologis,A., and Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES                Location/Qualifiers
     source             1. .1326
                        /organism="Arabidopsis thaliana"
                        /mol_type="mRNA"
                        /db_xref="taxon:3702"
                        /chromosome="1"
                        /clone="U22176"
                        /ecotype="Columbia"
                        /note="This clone is in pUNI 51"
     CDS                1. .1326
                        /note="unknown protein"
                        /codon_start=1
                        /product="At1g11800/F25C20_3"
                        /protein_id="AAN18168.1"
                        /db_xref="GI:23308397"
                        /translation="MSTTIRFLRIVTSRAMSSSWSCNKCTFLNSASQKLNCMICLAPV
                        SLPSLSPPPPSLSISANDEAKWACKACTFLNTYKNSICDVCGTRSPTSSLLGFQDLTD
                        SGLESNDADSSVGSVFFPLRRCIKRKAMDDDVVEVDGASVVCSESQGVMKKNKEIETK
                        GVASDSGTPLTCLKILSYNVWFREDLELNLRMRAIGHLIQLHSPHLICFQEVTPEIYD
                        IFRKSNWWKAYSCSVSVDVAVSRGYYCMLLSKLGVKSFSSKSFGNSIMGRELSIAEVE
                        VPGRKPLVFATSHLESPCPGPPKWDQMFSRERVEQAKEAIEILRPNANVIFGGDMNWC
                        DKLDGKFPLPDKWVDVWEVLKPGDLGFTYDTKANPMLSGNRALQKRLDRILCRLDDYK
```

LGGIEMVGKEAIPGLSYVKEKKVRGDIKKLELPVLPSDHFGLLVTLSPK"
ORIGIN

Alignment Scores:

Pred. No.:	2.52e-11	Length:	1326
Score:	231.00	Matches:	84
Percent Similarity:	44.2%	Conservative:	39
Best Local Similarity:	30.2%	Mismatches:	109
Query Match:	12.1%	Indels:	46
DB:	15	Gaps:	11

US-10-757-745-2 (1-362) x BT000599 (1-1326)

```
Qy        113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132
              :::   |||          |||       :::   |||:::||||||          |||
Db        523 ATCTTGAGCTACAATGTTTGGTTTCGTGAAGACCTTGAGTTGAAT---------CTAAGG 573

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
                  |||   :::          :::   |||:::||||||   :::|||      ||||||
Db        574 ATGAGAGCGATTGGTCATCTTATTCAGCTTCATTCTCCTCATCTCATATGTTTTCAGGAA 633

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArg-----------------SerSer 166
              |||   |||      |||          :::|||                    |||
Db        634 GTCACTCCTGAGATATATGATATTTTCCGCAAGTCGAATTGGTGGAAAGCGTATTCTTGT 693

Qy        167 AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186
              :::      :::                 ||||||:::      ::::::
Db        694 TCTGTCTCAGTTGATGTGGCAGTTTCGAGAGGCTACTACTGCATGCTGTTG--------- 744

Qy        187 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 206
              |||::::::   :::||||||          |||   ::::::   |||   |||
Db        745 AGCAAGTTGGGAGTGAAATCTTTCAGCAGCAAAAGCTTTGGGAACTCGATAATGGGAAGA 804

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGlu---LeuCysLeuMetThrSer 225
                 |||          |||   |||   |||   :::   |||          ||||||
Db        805 GAACTATCTATCGCAGAAGTTGAAGTTCCAGGGAGAAAGCCACTTGTTTTTGCAACTAGC 864

Qy        226 HisLeuGluSer-----------------------ThrArgGlyHisAlaAlaGluArg 237
              ||||||||||||                      :::     :::   ||||||
Db        865 CATCTTGAGAGTCCCTGTCCAGGTCCTCCAAAATGGGACCAAATGTTCAGTCGAGAACGC 924

Qy        238 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
              :::   |||   |||      ::::::  ::::::           :::|||   |||
Db        925 GTTGAACAAGCCAAGGAGGCAATTGAGATTCTCAGA---------CCTAATGCCAATGTG 975

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly------ 275
              ||||||   ||||||   |||      |||          :::   |||
Db        976 ATATTCGGGGGAGACATGAACTGGTGCGAC-----------AAGCTAGATGGGAAGTTT 1023

Qy        276 ---LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294
                 ||||||:::      ||||||||||||||   |||   ||||||
Db       1024 CCTTTGCCAGACAAATGGGTGGATGTTTGGGAGGTTTTG---AAACCAGGTGATTTGGGT 1080

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314
              :::|||:::||||||:::   |||      |||          |||   :::   |||
Db       1081 TTTACATATGACACAAAAGCAAACCCAATGTTGTCCGGTAACCGGGCATTGCAGAAACGG 1140

Qy        315 PheAspArgIlePhePheArg------------------------------------- 321
                 |||||||||      |||
Db       1141 TTAGACCGGATTTTGTGTCGTCTGGATGATTACAAGCTTGGTGGAATCGAGATGGTTGGG 1200

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGly---LeuGlu 340
                   |||      |||          :::          :::   |||   ::::::
Db       1201 AAAGAGGCCATTCCTGGTTTATCATATGTAAAAGAGAAGAAAGTGAGGGGTGACATTAAG 1260

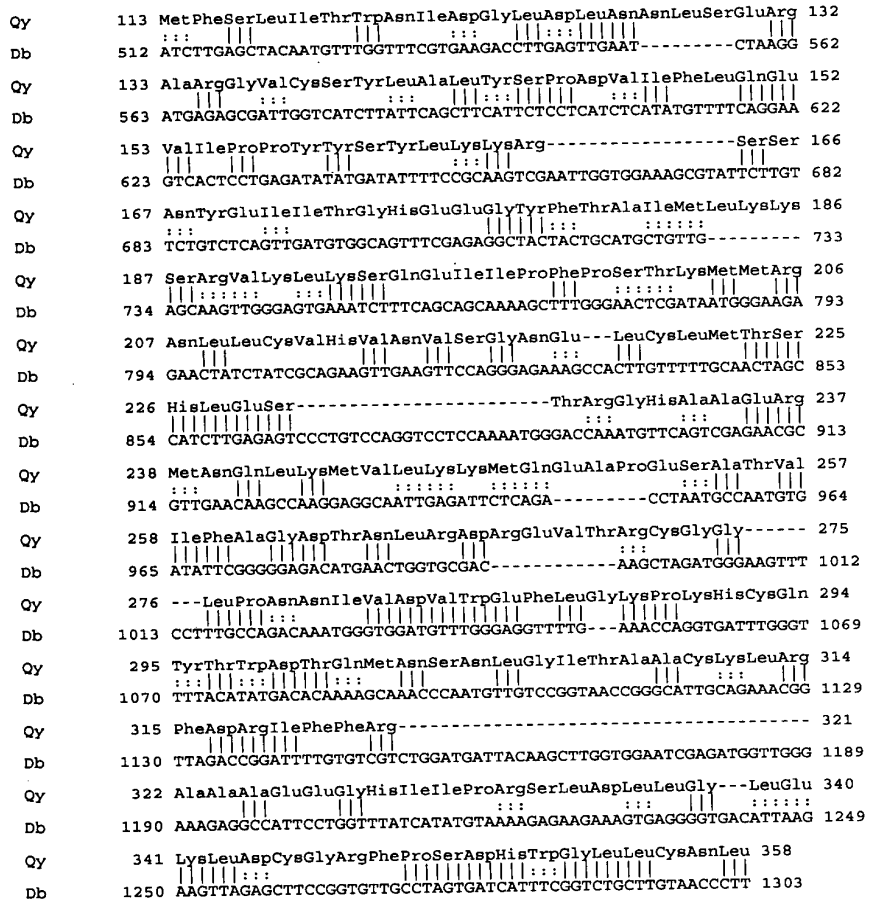
Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358
              ||||||:::   |||||||||||:::||||||||||      |||
Db       1261 AAGTTAGAGCTTCCGGTGTTGCCTAGTGATCATTTCGGTCTGCTTGTAACCCTT 1314
```

RESULT 47
AY085643
LOCUS AY085643 1557 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 16491 mRNA, complete sequence.
ACCESSION AY085643
VERSION AY085643.1 GI:21404353
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1557)
AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
TITLE Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
PUBMED 12093376
REFERENCE 2 (bases 1 to 1557)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Full-Length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1557)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or LAer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.
FEATURES Location/Qualifiers
source 1. .1557
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="16491"
CDS 35. .1315
/codon_start=1
/product="unknown"
/protein_id="AAM62864.1"
/db_xref="GI:21553771"
/translation="MSSSWSCNKCTFLNSASQKLNCMICLAPVSLPSLSPPPPSLSIS
ANDEAKWACKACTFLNTYKNSICDVCGTRSPTSSLLGFQDLTDSGLESNDADSSVGSV
FFPLRRCIKRKAMDDDVVEVDGASVVCSASQGVMKKNKEIETKGVASDSGTPLTCLKI
LSYNVWFREDLELNLRMRAIGHLIQLHSPHLICFQEVTPEIYDIFRKSNWWKAYSCSV
SVDVAVSRGYYCMLLSKLGVKSFSSKSFGNSIMGRELSIAEVEVPGRKPLVFATSHLE
SPCPGPPKWDQMFSRERVEQAKEAIEILRPNANVIFGGDMNWCDKLDGKFPLPDKWVD
VWEVLKPGDLGFTYDTKANPMLSGNRALQKRLDRILCRLDDYKLGGIEMVGKEAIPGL
SYVKEKKVRGDIKKLELPVLPSDHFGLLVTLSPK"
ORIGIN

Alignment Scores:

Pred. No.:	3.04e-11	Length:	1557
Score:	231.00	Matches:	84
Percent Similarity:	44.2%	Conservative:	39
Best Local Similarity:	30.2%	Mismatches:	109
Query Match:	12.1%	Indels:	46
DB:	15	Gaps:	11

US-10-757-745-2 (1-362) x AY085643 (1-1557)



RESULT 48
AY128286

```
LOCUS       AY128286               1582 bp    mRNA    linear   PLN 07-AUG-2002
DEFINITION  Arabidopsis thaliana At1g11800/F25C20_3 mRNA, complete cds.
ACCESSION   AY128286
VERSION     AY128286.1  GI:22135815
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1582)
  AUTHORS   Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Banh,J., Bowser,L.,
            Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
            Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
            and Ecker,J.R.
  TITLE     Arabidopsis cDNA clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1582)
  AUTHORS   Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Banh,J., Bowser,L.,
            Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
            Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
            and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUN-2002) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

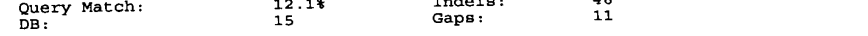
            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Shinn,P., Chen,H.,
            Cheuk,R., Kim,C.J.,  Banh,J. Bowser,L., Chang,E., Dale,J.M.,
            Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M.,
            Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,
            Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,
            Y., Yu,G., Yu,S., Davis,R.W.,  Theologis,A., and Ecker,J.R.

            Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES              Location/Qualifiers
     source           1. .1582
                      /organism="Arabidopsis thaliana"
                      /mol_type="mRNA"
                      /db_xref="taxon:3702"
                      /chromosome="1"
                      /clone="RAFL09-10-B22(R22176)"
                      /ecotype="Columbia"
     CDS              28. .1353
                      /note="unknown protein"
                      /codon_start=1
                      /product="At1g11800/F25C20_3"
                      /protein_id="AAM91094.1"
                      /db_xref="GI:22135816"
                      /translation="MSTTIRFLRIVTSRAMSSSWSCNKCTFLNSASQKLNCMICLAPV
                      SLPSLSPPPPSLSISANDEAKWACKACTFLNTYKNSICDVCGTRSPTSSLLGFQDLTD
                      SGLESNDADSSVGSVFFPLRRCIKRKAMDDDVVEVDGASVVCSESQGVMKKNKEIETK
                      GVASDSGTPLTCLKILSYNVWFREDLELNLRMRAIGHLIQLHSPHLICFQEVTPEIYD
                      IFRKSNWWKAYSCSVSVDVAVSRGYYCMLLSKLGVKSFSSKSFGNSIMGRELSIAEVE
                      VPGRKPLVFATSHLESPCPGPPKWDQMFSRERVEQAKEAIEILRPNANVIFGGDMNWC
                      DKLDGKFPLPDKWVDVWEVLKPGDLGFTYDTKANPMLSGNRALQKRLDRILCRLDDYK
                      LGGIEMVGKEAIPGLSYVKEKKVRGDIKKLELPVLPSDHFGLLVTLSPK"
ORIGIN
```

Alignment Scores:

Pred. No.:	3.1e-11	Length:	1582
Score:	231.00	Matches:	84
Percent Similarity:	44.2%	Conservative:	39
Best Local Similarity:	30.2%	Mismatches:	109
Query Match:	12.1%	Indels:	46
DB:	15	Gaps:	11

US-10-757-745-2 (1-362) x AY128286 (1-1582)



```
Db        550 ATCTTGAGCTACAATGTTTGGTTTCGTGAAGACCTTGAGTTGAAT---------CTAAGG 600

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152

Db        601 ATGAGAGCGATTGGTCATCTTATTCAGCTTCATTCTCCTCATCTCATATGTTTTCAGGAA 660

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArg-----------------SerSer 166

Db        661 GTCACTCCTGAGATATATGATATTTTCCGCAAGTCGAATTGGTGGAAAGCGTATTCTTGT 720

Qy        167 AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186

Db        721 TCTGTCTCAGTTGATGTGGCAGTTTCGAGAGGCTACTACTGCATGCTGTTG---------- 771

Qy        187 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 206

Db        772 AGCAAGTTGGGAGTGAAATCTTTCAGCAGCAAAAGCTTTGGGAACTCGATAATGGGAAGA 831

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGlu---LeuCysLeuMetThrSer 225

Db        832 GAACTATCTATCGCAGAAGTTGAAGTTCCAGGGAGAAAGCCACTTGTTTTTGCAACTAGC 891

Qy        226 HisLeuGluSer-----------------------ThrArgGlyHisAlaAlaGluArg 237

Db        892 CATCTTGAGAGTCCCTGTCCAGGTCCTCCAAAATGGGACCAAATGTTCAGTCGAGAACGC 951

Qy        238 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257

Db        952 GTTGAACAAGCCAAGGAGGCAATTGAGATTCTCAGA---------CCTAATGCCAATGTG 1002

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly------ 275

Db       1003 ATATTCGGGGGAGACATGAACTGGTGCGAC-----------AAGCTAGATGGGAAGTTT 1050

Qy        276 ---LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294

Db       1051 CCTTTGCCAGACAAATGGGTGGATGTTTGGGAGGTTTTG---AAACCAGGTGATTTGGGT 1107

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314

Db       1108 TTTACATATGACACAAAAGCAAACCCAATGTTGTCCGGTAACCGGGCATTGCAGAAACGG 1167

Qy        315 PheAspArgIlePhePheArg---------------------------------------- 321

Db       1168 TTAGACCGGATTTTGTGTCGTCTGGATGATTACAAGCTTGGTGGAATCGAGATGGTTGGG 1227

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGly---LeuGlu 340

Db       1228 AAAGAGGCCATTCCTGGTTTATCATATGTAAAAGAGAAGAAAGTGAGGGGTGACATTAAG 1287

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358

Db       1288 AAGTTAGAGCTTCCGGTGTTGCCTAGTGATCATTTCGGTCTGCTTGTAACCCTT 1341

RESULT 49
CR954207_1/c
WPCOMMENT
  Sequence split into 8 fragments  LOCUS CR954207 Accession CR954207
     Fragment Name      Begin     End
     CR954207_0             1    110000
     CR954207_1        100001    210000
     CR954207_2        200001    310000
     CR954207_3        300001    410000
     CR954207_4        400001    510000
     CR954207_5        500001    610000
     CR954207_6        600001    710000
     CR954207_7        700001    739027
Continuation (2 of 8) of CR954207 from base 100001 (CR954207 Ostreococcus tauri strain C

Alignment Scores:
Pred. No.:              2.16e-08        Length:         110000
Score:                  223.50          Matches:        96
Percent Similarity:     40.2%           Conservative:   61

Best Local Similarity:  24.6%           Mismatches:     159
Query Match:            11.8%           Indels:         75
DB:                     14              Gaps:           17

US-10-757-745-2 (1-362) x CR954207_1 (1-110000)

Qy         20 ProGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAsp 39

Db      16762 CCGAACGCCAAACGCGCGCGC-----------------ACGGATGTCAAGGAACCATCG 16721

Qy         40 AlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrp-----------GluMetGlu 55

Db      16720 AGTTCAAAAGCTCCAGCTGACAGCTCCTCGAACGGTTGGCTCGCCGCTTTACACGCCGAG 16661

Qy         56 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg------ 73

Db      16660 CGATGCGCGCGCAGGGGTATAGATGCGGGCGTGAGCGTCCAAGAGGCGGAGCGGAACCGC 16601

Qy         74 ---------ArgProGluThrIle----------------------------SerGlu 80

Db      16600 GCGAACGAGCGCGCGCGGGCGATCGCGACGACGGTGTTGGGTGGAGGCGGGCGGGCTGAG 16541

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db      16540 ACGCGACGAGCGACGGATGTCGCGAGGCAGCGGCGGACGAATTCGACGAGCGCAAAGACG 16481

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db      16480 AGCGATGAGGCGGAGACGCGCGAGCACGATGCGGGTGATTTACGTTTGTTGACATACAAC 16421

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db      16420 GTGTGGTTCGGAGATGTCGCTCAA---GCGGCGAGATTGGAGGCGCTTGGAAAGGTCATC 16364

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db      16363 GCGGAGTGCGACCCACACGTGTTGTGTCTGCAGGAGGTGACGCCGCAGTCTCTGTTCGCA 16304

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu---------GluGly 177

Db      16303 CTTCGCGCGCACGAGTGGTGGAGCGAGTACGTCGTTGGTCCGAAACCGCAATTACAGGAG 16244

Qy        178 TyrPheThrAlaIleMetLeuLysLys---------SerArgValLysLeuLysSerGln 194

Db      16243 TACTTCAGCGTCGTGCTGTTCAAGAGGGAGATCGGGAGCGAGCTAGTGCGGCGAGACCAA 16184

Qy        195 GluIleIleProPheProSerThrLysMetMetArgAsnLeu---------LeuCysVal 211

Db      16183 CAT------CCTTTTGCGAACTCGAAGATGGGGCGGTGCCTGGACATCGTGAGTTACGTA 16130

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu--------- 228

Db      16129 CGGTGTCCACTCAGTGGGACGAACTTCGCGGTCGGGAGCTCGCATCTCGAGTCCTACATC 16070

Qy        229 ---SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 247

Db      16069 AGTTCCAATCGTACCAGCGCTCAGGAGCGACGAGTGCAGATAGAAGAGTGTTTCTCGCGG 16010

Qy        248 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 267

Db      16009 CTGAAC---------AAGCACGCGAACGCGATTTTCATGGGCGACACGAATTGGGACGAT 15959

Qy        268 ArgGluValThrArgCysGlyGly------LeuProAsnAsnIleValAspValTrpGlu 285

Db      15958 AAGGAT------------GGAGACGTACCTCTGCCGGAGGGATGGCGAGACGCGTGGCTC 15911

Qy        286 PheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeu 305

Db      15910 GAGCTT---CGCCCGGATGAACCTGGTTTCACGTACGACGCGCGCAAAAACGACATGCTA 15854

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 325

Db      15853 ------CGCGGCTACCTACAGAAGCGGCTGGATCGAGCGTTCGTTAAGCTCAAG------ 15806
```

```
Qy        326 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeu------------------ 339
                  |||          ::::::::::::::::::|||
Db      15805 ------CACTTTGACGTGAAAAACATCGAGATGGTGGGCACGCAACCGATTCCCGGTGTG 15752

Qy        340 ----------------------------------GluLysLeuAspCysGlyArgPhe 347
                                                 |||   ::::::
Db      15751 ACTTACCACAAGACTGTGAGTAATCGCGGGAAAATCTGAGATCATCGAACTCCCAGTATTG 15692

Qy        348 ProSerAspHisTrpGlyLeuLeuCysAsnLeu 358
                |||||||||||:::|||||||||     |||
Db      15691 CCGAGCGATCATTTCGGGCTATTACTGACACTT 15659
```

RESULT 50
AR240556/c

LOCUS	AR240556	268 bp	DNA	linear	PAT 20-DEC-2002

DEFINITION Sequence 312 from patent US 6468758.
ACCESSION AR240556
VERSION AR240556.1 GI:27285636
KEYWORDS .
SOURCE Unknown.
  ORGANISM Unknown.
  Unclassified.
REFERENCE 1 (bases 1 to 268)
  AUTHORS Benson,D.R., Lodes,M.J., Mitcham,J.L. and King,G.E.
  TITLE Compositions and methods for ovarian cancer therapy and diagnosis
  JOURNAL Patent: US 6468758-A 312 22-OCT-2002;
  Corixa Corporation; Seattle, WA;
  WOX;

FEATURES Location/Qualifiers
  source 1. .268
  /organism="unknown"
  /mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.:	3.12e-11	Length:	268
Score:	221.00	Matches:	41
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	11.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2 (1-362) x AR240556 (1-268)

```
Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAA 208

Qy        342 LeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 CTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATA 148

Qy        362 Leu 362
              |||
Db        147 TTG 145
```

RESULT 51
F25C20

LOCUS	F25C20	91566 bp	DNA	linear	PLN 15-MAY-1999

DEFINITION Arabidopsis thaliana chromosome 1 BAC F25C20 sequence, complete sequence.
ACCESSION AC007296
VERSION AC007296.2 GI:4760411
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
  ORGANISM Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 91566)
  AUTHORS Vysotskaia,V.S., Schwartz,J.R., Yu,G., Toriumi,M., Lenz,C., Liu,S., Li,J., Lee,J.M., Kremenetskaia,I., Luros,J., Ngan,I., Liu,A., Gonzalez,A., Altafi,H., Araujo,R., Chao,Q., Conn,L., Conway,A.B., Dunn,P., Hansen,N., Huizar,L., Kim,C., Palm,C., Rowley,D., Shinn,P., Walker,M., Davis,R.W., Ecker,J.R., Federspiel,N.A. and Theologis,A.
  TITLE Arabidopsis thaliana chromosome 1 BAC F25C20 sequence
  JOURNAL Unpublished (1999)
  REMARK This sequence is of BAC F25C20 from Arabidopsis thaliana chromosome 1. The sequence does not represent the sequence of the entire insert of this clone. It is shorter by 1656 bp because we submit only the unique sequence of the clone. However, in order to facilitate the joining of overlapping clones in the future for creation of larger contigs, we provide a small overlap (200 bp) between overlapping submitted clones. The 5' end of this sequence overlaps by 200 bp the 3' end of the sequence of the BAC F12F1.
REFERENCE 2 (bases 1 to 91566)
  AUTHORS Theologis,A.
  TITLE Direct Submission
  JOURNAL Submitted (13-APR-1999) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
REFERENCE 3 (bases 1 to 91566)
  AUTHORS Theologis,A.
  TITLE Direct Submission
  JOURNAL Submitted (08-MAY-1999) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
REFERENCE 4 (bases 1 to 91566)
  AUTHORS Theologis,A.
  TITLE Direct Submission
  JOURNAL Submitted (15-MAY-1999) Plant Gene Expression Center, 800 Buchanan St., Albany, CA 94710, USA
COMMENT On May 8, 1999 this sequence version replaced gi:4582412.
  The sequence of BAC F25C20 from Arabidopsis thaliana chromosome 1.

FEATURES Location/Qualifiers
  source 1. .91566
  /organism="Arabidopsis thaliana"
  /mol_type="genomic DNA"
  /db_xref="taxon:3702"
  /chromosome="1"
  /clone="F25C20"
  /ecotype="Columbia"
  gene <4. .294
  /gene="F25C20.1"
  CDS join(<4. .50,162. .294)
  /gene="F25C20.1"
  /note="EST gb|AA605362 comes from this gene."
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAD30264.1"
  /db_xref="GI:4835798"
  /translation="GDAYVSGHGSCIFPGSKKVGNRTQTVVNSTEVAAGEATSRSLSR
  GFCVTIMILVTFSIL"
  gene complement(542. .1675)
  /gene="F25C20.2"
  CDS complement(542. .1675)
  /gene="F25C20.2"
  /note="EST gb|AA605362 comes from this gene."
  /codon_start=1
  /evidence=not_experimental
  /protein_id="AAD30240.1"
  /db_xref="GI:4835774"
  /translation="MTTTMSTLPVVLVDEILARVPITSLRSLRSTCKKWEASSKTNLV
  GGKATARKSSHVGFILIGDKICSMKLDLNGGDDFVDTSVNQVSAFDGFAISQLFHCDG
  LLFCISNNHYSNYTLMVCNMYLGETRLIQNRSLFESYQNFCSYAFGYDSSKNRNHKIL
  RNNSVSGGYEIYSLKSDSWKDLNVDLEKSIHLWRLGSVSLKGNAYFRVIKVIEEGVWE
  YNLLCFDFTRESFGKLLSLPFESLGDEGEGIMVISCVRDDHLAVLYQRDTLGIWISTE
  IEPNKVSWREFLQVDLATLDGFPDVFIAGRFIVDEEKQVVVVFGQETELDLNHGNAFI
  FGRDGYFTSFTVGDAPPEDFTSYVPSLVSLQIDKTGKRKARDD"
  gene 4175. .5918
  /gene="F25C20.3"
  CDS join(4175. .4655,4887. .5035,5129. .5239,5334. .5918)
  /gene="F25C20.3"
  /note="EST gb|F14156 comes from this gene."
  /codon_start=1
  /evidence=not_experimental

/protein_id="AAD30241.1"
/db_xref="GI:4835775"
/translation="MSTTIRFLRIVTSRAMSSSWSCNKCTFLNSASQKLNCMICLAPV
SLPSLSPPPPSLSISANDEAKWACKACTFLNTYKNSICDVCGTRSPTSSLLGFQDLTD
SGLESNDADSSVGSVFFPLRRCIKRKAMDDDVVEVDGASVVCSESQGVMKKNKEIETK
GVASDSGTPLTCLKILSYNVWFREDLELNLRMRAIGHLIQLHSPHLICFQEVTPEIYD
IFRKSNWWKAYSCSVSVDVAVSRGYYCMLLSKLGVKSFSSKSFGNSIMGRELSIAEVE
VPGRKPLVFATSHLESPCPGPPKWDQMFSRERVEQAKEAIEILRPNANVIFGGDMNWC
DKLDGKFPLPDKWVDVWEVLKPGDLGFTYDTKANPMLSGNRALQKRLDRILCRLDDYK
LGGIEMVGKEAIPGLSYVKEKKVRGDIKKLELPVLPSDHFGLLVTLSPK"

gene complement(6476. .9962)
/gene="F25C20.4"

CDS complement(join(6476. .6532,6615. .6696,6776. .6843,
6930. .7019,7103. .7180,7360. .7455,7533. .7613,7697. .7897,
7988. .8080,8167. .8234,9698. .9962))
/gene="F25C20.4"
/note="Similar to gi|2392772 T32N15.11 putative
chloroplast prephenate dehydratase from Arabidopsis
thaliana BAC gb|AC002534 and is a member of the PF|00800
Prephenate dehydratase family. ESTs gb|T21562 and
gb|T21062 come from this gene."
/codon_start=1
/evidence=not_experimental
/protein_id="AAD30242.1"
/db_xref="GI:4835776"
/translation="MALRCFPIWVCPQTTHHRSPLMGLAEFDADKRRRFCLWECSSSA
SQRAVTAIEGEIPFSRELKKSSDELGLTQETQSLSFHRDLSMLPKPLTANSLYSSDGD
DSKVRISFQGIPGAYSETAALKAFPNCETVPCEQFEAAFQAVELWLVDKAVLPIENSV
GGSIHRNYDLLLRHRLHIVQEVHLPVNHCLLGVPGVKKEDIKCVLSHPQALDQCVNSL
NNLGIQRISAKDTATAAQTVSSSGKIDVGAIASVRAANIYGLDILAENIQDDVNNVTR
FLILAREPMIPRTDRPYKTSIVFSLEEGPGVLFKALAVFALRSINLSKIESRPQRRRP
LRVVDGSNNGSAKYFDYLFYIDFEASMADTRAQHALGHLQEFASFIRILGCYPMDLVR
"

gene 11062. .11290
/gene="F25C20.5"

CDS join(11062. .11142,11180. .11290)
/gene="F25C20.5"
/codon_start=1
/evidence=not_experimental
/product="F25C20.5"
/protein_id="AAD30243.1"
/db_xref="GI:4835777"
/translation="MRGERDETSNGREEEVIKVVVVCMVCMWSCVNCGRNNGDGGRDD
SVMVVSMVVTIAVVMVMMV"

gene 12261. .13824
/gene="F25C20.6"

CDS join(12261. .12367,12447. .12595,12676. .12989,13077. .13215,
13302. .13368,13453. .13601,13685. .13824)
/gene="F25C20.6"
/codon_start=1
/evidence=not_experimental
/product="F25C20.6"
/protein_id="AAD30244.1"
/db_xref="GI:4835778"
/translation="MYESANVSDDADRTAFRRAEKKYKLYYEQDSKFSRKKKLPKPID
LSELLDFNLISQNFNNDGVLPDGIRVSKVDSSPVFCIDNRPGFYFIPDALSLKEQCKW
IKESLTSFPQPPNRTNHNAIYGPIDDLFDSAKENKVLVQDDLTNNKWKFYEEVDIEKA
TRSSCKSVSASVLLRKLRWSTLGLQFDWSKRNYDVSLPHNNIPDALCQLAKTHAAIAM
PDGEEFRPEGAIVNYFGIGDTLGGHLDDMEADWSKPIVSMSLGCKAIFLLGGKSKDDP
PHAMYLRSGDVVLMAGEARECFHGNLLHFQLDGIPRIFTGEENADIGALESELSHESG
HFFAEYIKTSRININIRQVF"

gene complement(14149. .15759)
/gene="F25C20.7"

CDS complement(14149. .15759)
/gene="F25C20.7"
/note="Strong similarity to gb|AF049347 berberine bridge
enzyme from Berberis stolonifera."
/codon_start=1
/evidence=not_experimental
/protein_id="AAD30246.1"
/db_xref="GI:4835780"
/translation="MKIFCLILFLISSFISTSLAVEPPPETIYQNFLQCFTNQTKAPP
NSLADVVLPKTAAAFTPVLRAYIRNARFNTTATPKPAIVIAARSESHVQAAVICTKSL
NIQLKTRSGGHDYEGVSYISHVPFFVLDMSNLRNITVDPATESAWVGAGATLGEVYYR
IWEKTKSHGFPAGVCPTVGAGGHISGGGYGNMIRKYGLSVDYVTDAKIVDVNGQVLDR
KGMGEDMFWAINGGGGASFGVILAFKIKLVPVPPTVTVFRVEKNLVENATEMVHKWQF
VAPKTDPGLFMRLLLQPVTRNKMQTVRASVVALFLGDQNTVMSMLTKEFPELGLKKEN
CTEMTWIQSVMWWANNDNATQIKPEILLDRNPDMATFGKRKSDFVEKEITKDGLDFLF
KKMIEVGKIGLVFNPYGGIMSTVATTKTPFPHRKKLYKIQHSMNWKDPGTEAETSFLQ
KAKSFYSYMAPFVTKNPRHTYINYRDLDIGVNTPGPNSYRVAEVFGRMYFGENFDRLV
KVKTAVDPQNFFRDEQSIPTLPGKPARR"

gene complement(17116. .17529)
/gene="F25C20.8"

CDS complement(17116. .17529)
/gene="F25C20.8"
/codon_start=1
/evidence=not_experimental
/product="F25C20.8"
/protein_id="AAD30245.1"
/db_xref="GI:4835779"
/translation="MKNTSIFLVVFALCMISNAYGHSNPFKGKKTSLLFRNSLSHNKW
LKVRCKSGDNDVRERYMKPGEDWGFSFHDDVMGETLFWCTMYKGADYKVQKKFDAYVQ
DKRMPHGGSYNYVAQEDGIYHSTLIHRMRKMYDWN"

gene 18142. .21450
/gene="F25C20.9"

CDS join(18142. .18195,18614. .19044,19945. .19996,20154. .20424,
20494. .20555,20686. .20881,20962. .21058,21147. .21229,
21443. .21450)
/gene="F25C20.9"
/codon_start=1
/evidence=not_experimental
/product="F25C20.9"
/protein_id="AAD30247.1"
/db_xref="GI:4835781"
/translation="MKKKKKKESDTNRTLRSSPGRDQMDNIVDSLNKAYEKFVLASAGV
LESKESAGGQKALLTDTALENFKEKWELFRVACDQAEEFVESVKQRIGSECLVDEATG
LTTTAAGGQAPAAVTGAATSLPPISAVRLEQMSRAVRWLVLELQRGSGVAPGSVHSSS
TGYMDSNQSMRLLSAWIGQIGDLGLNLLWRFIHIVVSLWYIVSGIFEAIESYAITLGL
NKKYGSIDLEKLRCLAVVVDIEAAQDVANVVELLQWLTTIGVKQVGLFDSQGLLKKSK
DLILETVPGSMLLEEIEKDVAPDGKRIALEFISSSDNKEAVMKAANILLQRYLKSSHP
EDDKGEDFFTESHLNDALRVVGENVHVPDLLLVYGPIRSHLGFPAWRLRYTEIVHMGT
LKYMRYGSLLKAIHKFTGVHQNYGT"

gene complement(21902. .23828)

Alignment Scores:

Pred. No.:	3.99e-05	Length:	91566
Score:	186.50	Matches:	92
Percent Similarity:	36.5%	Conservative:	42
Best Local Similarity:	25.1%	Mismatches:	117
Query Match:	9.8%	Indels:	116
DB:	15	Gaps:	14

US-10-757-745-2 (1-362) x F25C20 (1-91566)

```
Qy        84 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSer 103
             :::|||   :::   ::::::      |||      |||      ||||||
Db      4880 TTTGTAGGGGTAGCTTCTGATTCAGGAACACCTTTAACTTGCTTGAAGATC--------- 4930

Qy       104 GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGly 123
                                              |||       |||      :::
Db      4931 ----------------------------TTGAGCTACAATGTTTGGTTTCGTGAAGAC 4960

Qy       124 LeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyr 143
             |||:::||||||           |||  |||  :::      :::   |||:::
Db      4961 CTTGAGTTGAAT---------CTAAGGATGAGAGCGATTGGTCATCTTATTCAGCTTCAT 5011

Qy       144 SerProAspValIlePheLeuGln---------------------------------- 151
             ||||||   :::|||     |||
Db      5012 TCTCCTCATCTCATATGTTTTCAGGTCACTTTTACATTTTGCTATAGAATGATTTCTATC 5071

Qy       152 ----------------------------------------------------------Glu 152
                                                                       |||
Db      5072 TGCAACATTTAGAAATAGCTCAGCTTTCTTGACAAGTGGTTTTTGTGTGGTTTGTAGGAA 5131

Qy       153 ValIleProProTyrTyrSerTyrLeuLysLysArg-----------------SerSer 166
             |||   |||      |||         :::|||                    |||
```

```
Db       5132 GTCACTCCTGAGATATATGATATTTTCCGCAAGTCGAATTGGTGGAAAGCGTATTCTTGT 5191

Qy        167 AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe---------------------- 179

Db       5192 TCTGTCTCAGTTGATGTGGCAGTTTCGAGAGGCTACTACTGCATGCTGGTAAAGATTGTG 5251

Qy        180 ------------------------------------------------------------Thr 180

Db       5252 CTTGCTTGGCTTAGGCTCTTGCCGTCTGTATCTGAATAATTCCCTGTGGTTCTTGTAACT 5311

Qy        181 -AlaIleMet------LeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePr 198

Db       5312 GACCCTTATGTGTAATTTGCAGTTGAGCAAGTTGGGAGTGAAATCTTTCAGCAGCAAAAG 5371

Qy        198 oPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAs 218

Db       5372 CTTTGGGAACTCGATAATGGGAAGAGAACTATCTATCGCAGAAGTTGAAGTTCCAGGGAG 5431

Qy        218 nGlu---LeuCysLeuMetThrSerHisLeuGluSer----------------------- 229

Db       5432 AAAGCCACTTGTTTTTGCAACTAGCCATCTTGAGAGTCCCTGTCCAGGTCCTCCAAAATG 5491

Qy        230 -ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGl 249

Db       5492 GGACCAAATGTTCAGTCGAGAACGCGTTGAACAAGCCAAGGAGGCAATTGAGATTCTCAG 5551

Qy        249 nGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGl 269

Db       5552 A---------CCTAATGCCAATGTGATATTCGGGGGAGACATGAACTGGTGCGAC----- 5597

Qy        269 uValThrArgCysGlyGly---------LeuProAsnAsnIleValAspValTrpGluPh 286

Db       5598 -------AAGCTAGATGGGAAGTTTCCTTTGCCAGACAAATGGGTGGATGTTTGGGAGGT 5650

Qy        286 eLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGl 306

Db       5651 TTTG---AAACCAGGTGATTTGGGTTTTACATATGACACAAAAGCAAACCCAATGTTGTC 5707

Qy        306 yIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg--------------- 321

Db       5708 CGGTAACCGGGCATTGCAGAAACGGTTAGACCGGATTTTGTGTCGTCTGGATGATTACAA 5767

Qy        322 ------------------------AlaAlaAlaGluGluGlyHisIleIleProArgSe 333

Db       5768 GCTTGGTGGAATCGAGATGGTTGGGAAAGAGGCCATTCCTGGTTTATCATATGTAAAAGA 5827

Qy        333 rLeuAspLeuLeuGly---LeuGluLysLeuAspCysGlyArgPheProSerAspHisTr 352

Db       5828 GAAGAAAGTGAGGGGTGACATTAAGAAGTTAGAGCTTCCGGTGTTGCCTAGTGATCATTT 5887

Qy        352 pGlyLeuLeuCysAsnLeu 358

Db       5888 CGGTCTGCTTGTAACCCTT 5906
```

RESULT 52
CQ069444

LOCUS	CQ069444 444 bp DNA linear PAT 20-JAN-2004
DEFINITION	Sequence 5244 from Patent WO0157278.
ACCESSION	CQ069444
VERSION	CQ069444.1 GI:41039313
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE	Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human hela cells or other human cervical epithelialcells
JOURNAL	Patent: WO 0157278-A 5244 09-AUG-2001; Aeomica, Inc. (US)
FEATURES	Location/Qualifiers
source	1. .444 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="MAP TO AL031775.1 EXPRESSED IN HELA, SIGNAL = 1.9"
ORIGIN	

Alignment Scores:

Pred. No.:	1.77e-07	Length:	444
Score:	182.50	Matches:	37
Percent Similarity:	74.5%	Conservative:	4
Best Local Similarity:	67.3%	Mismatches:	11
Query Match:	9.6%	Indels:	3
DB:	6	Gaps:	1

US-10-757-745-2 (1-362) x CQ069444 (1-444)

```
Qy        135 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 151

Db        221 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 280

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171

Db        281 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 340

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186

Db        341 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 385
```

RESULT 53
CQ078617

LOCUS	CQ078617 691 bp DNA linear PAT 20-JAN-2004
DEFINITION	Sequence 14417 from Patent WO0157278.
ACCESSION	CQ078617
VERSION	CQ078617.1 GI:41048486
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
TITLE	Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human hela cells or other human cervical epithelialcells
JOURNAL	Patent: WO 0157278-A 14417 09-AUG-2001; Aeomica, Inc. (US)
FEATURES	Location/Qualifiers
source	1. .691 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="MAP TO AL031775.1 EXPRESSED IN HELA, SIGNAL = 1.9 SWISSPROT HIT: P18004, EVALUE 7.00e-01 NT HIT: AF201687.1, EVALUE 3.00e-45 EST_HUMAN HIT: AV717253.1, EVALUE 5.00e-45"
ORIGIN	

Alignment Scores:

Pred. No.:	2.98e-07	Length:	691
Score:	182.50	Matches:	37
Percent Similarity:	74.5%	Conservative:	4
Best Local Similarity:	67.3%	Mismatches:	11
Query Match:	9.6%	Indels:	3
DB:	6	Gaps:	1

US-10-757-745-2 (1-362) x CQ078617 (1-691)

```
Qy        135 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 151
```

```
            |||:::|||      |||                 |||||||||||||||||||||||||||
Db         83 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 142

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 202

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186
            ||||||:::          |||   |||   :::       :::|||
Db        203 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 247

RESULT 54
CEY63D3A/c
LOCUS       CEY63D3A               33899 bp    DNA     linear   INV 09-AUG-2005
DEFINITION  Caenorhabditis elegans YAC Y63D3A, complete sequence.
ACCESSION   AL032652
VERSION     AL032652.1  GI:3810723
KEYWORDS    HTG.
SOURCE      Caenorhabditis elegans
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 33899)
  AUTHORS   .
  CONSRTM   C. elegans Sequencing Consortium
  TITLE     Genome sequence of the nematode C. elegans: a platform for
            investigating biology
  JOURNAL   Science 282 (5396), 2012-2018 (1998)
   PUBMED   9851916
REFERENCE   2  (bases 1 to 33899)
  AUTHORS   White,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-OCT-1998) Nematode Sequencing Project, Sanger
            Institute, Hinxton, Cambridge CB10 1SA, England and Department of
            Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
            worm@sanger.ac.uk
COMMENT     On May 14, 2001 this sequence version replaced gi:3646980.
            Coding sequences below are predicted from computer analysis, using
            predictions from Genefinder (P. Green, U. Washington), and other
            available information.
            Current sequence finishing criteria for the C. elegans genome
            sequencing consortium are that all bases are either sequenced
            unambiguously on both strands, or on a single strand with both a
            dye primer and dye terminator reaction, from distinct subclones.
            Exceptions are indicated by an explicit note.
            IMPORTANT:  This sequence is NOT necessarily the entire insert of
            the specified clone.  It may be shorter because we only sequence
            overlapping sections once, or longer because we arrange for a small
            overlap between neighbouring submissions.
            For a graphical representation of this sequence and its analysis
            see:- http://www.wormbase.org/perl/ace/elegans/seq/sequence?
            name=Y63D3A;class=Sequence
            IMPORTANT: This sequence is not the entire insert of clone Y63D3A.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we arrange for a small overlap between
            neighbouring submissions.
            The true left end of clone K03D10 is at 33796 in this sequence. The
            true right end of clone F47G4 is at 103 in this sequence. The true
            right end of clone Y71A12 is at 113 in this sequence. The start of
            this sequence (1. .103) overlaps with the end of sequence Z99171.
            The end of this sequence (33796. .33899) overlaps with the start of
            sequence Z81561.
FEATURES              Location/Qualifiers
     source           1. .33899
                      /organism="Caenorhabditis elegans"
                      /mol_type="genomic DNA"
                      /strain="Bristol N2"
                      /db_xref="taxon:6239"
                      /chromosome="I"
                      /clone="Y63D3A"
     gene             complement(join(49. .345,612. .1065,1117. .1199))
                      /locus_tag="Y63D3A.9"
     CDS              complement(join(49. .345,612. .1065,1117. .1199))
                      /locus_tag="Y63D3A.9"
                      /standard_name="Y63D3A.9"
                      /note="contains similarity to Pfam domain PF00646 (F-box
                      domain)"
                      /codon_start=1
                      /product="Hypothetical protein Y63D3A.9"
                      /protein_id="CAA21711.1"
                      /db_xref="GI:3881237"
                      /db_xref="InterPro:IPR001810"
                      /db_xref="UniProt/TrEMBL:Q9XWG0"
                      /translation="MSSPSFPILRLPSKALRTTLQNFDFIDLLSFSFVSPTPTKSLNM
                      KVTDFIIIKNYGIYIVARCCGIALRPITSQIFVDGQPQERQWEIQSITVKKCFENFLA
                      AFNHFEVEYICVRGEDDISYNLEEIKVVKNLEIFSADCGIFGQLYRVANKVTLSQRSL
                      LHEKPRDFFTENMDSLQLEKAVIYRPFTLSDLNLFLQHWRKGSNPRLEEIYIGLKEDN
                      LVDFKKTLLKGIEYWKYGIFRVYNNSPSIEGLEIERTDGIRAIIFVNPMHRSIGMGVL
                      Q"
     gene             join(1524. .1606,1651. .2086,4038. .4352)
                      /locus_tag="Y63D3A.10"
     CDS              join(1524. .1606,1651. .2086,4038. .4352)
                      /locus_tag="Y63D3A.10"
                      /standard_name="Y63D3A.10"
                      /note="contains similarity to Pfam domain PF00646 (F-box
                      domain)"
                      /codon_start=1
                      /product="Hypothetical protein Y63D3A.10"
                      /protein_id="CAA21706.1"
                      /db_xref="GI:3881232"
                      /db_xref="InterPro:IPR001810"
                      /db_xref="UniProt/TrEMBL:Q9XWG4"
                      /translation="MAHVAFNLLHLPAKTLRNVFQLMSHSELFIFSFASICAKRLVKS
                      MNFEMLELDLVYSRPCKLKISFGEHEEFIIEFDEFGTLSKQEPSDAFVFLIAYKLDIR
                      DCINHVLEVYKISKLHTLKVIGAINNPQNLLKIIGKLDLKVLWIFFNFSENSTTISVY
                      REILNHLGARCDVRSLRRGFYDKEPQKSFFLLRLSLSYALSQEPHSRPTATLAARRLK
                      RRQDTQPLDATTSVLMLRSCRTVGRGTSLPTHPPFSFVIAPHYYSCTCSDIQGEPDKL
                      Q"
     gene             complement(join(2689. .3289,3365. .3484,3604. .3686))
                      /locus_tag="Y63D3A.11"
     CDS              complement(join(2689. .3289,3365. .3484,3604. .3686))
                      /locus_tag="Y63D3A.11"
                      /standard_name="Y63D3A.11"
                      /note="contains similarity to Pfam domain PF00646 (F-box
                      domain)"
                      /codon_start=1
                      /product="Hypothetical protein Y63D3A.11"
                      /protein_id="CAB76748.1"
                      /db_xref="GI:7160769"
                      /db_xref="InterPro:IPR001810"
                      /db_xref="UniProt/TrEMBL:Q9NA68"
                      /translation="MSSPSLPILRLPSKALRTTLQNFEYLDLLSFSLVSPTPVKSLNM
                      KVTDFILISNDEIVIFGRATMFFRKSFENFLIAFNHFEVEHVYVNIENDIPYNLEEIK
                      VVKDLFVRSADRGIFSQFYRTANKISLSNRSLHNDKSRNFFTENIDSLWLDSSFHLEF
                      NDLLAINAKKAVIYRPFTLSDLNLFLQHWRKGSNPRLEEIYIGLEEDNLVDFNETLLK
                      GIVHWKSCLSKVYNNSPSIEGLEIERTDGTRAIIFVNPMKRSFEMGVLQ"
     gene             join(5095. .5177,5974. .6859)
                      /locus_tag="Y63D3A.3"
     CDS              join(5095. .5177,5974. .6859)
                      /locus_tag="Y63D3A.3"
                      /standard_name="Y63D3A.3"
                      /note="contains similarity to Pfam domain PF00646 (F-box
                      domain)"
                      /codon_start=1
                      /product="Hypothetical protein Y63D3A.3"
                      /protein_id="CAB63397.2"
                      /db_xref="GI:7160770"
                      /db_xref="InterPro:IPR001810"
                      /db_xref="UniProt/TrEMBL:Q9NA67"
                      /translation="MPNYPFPILWLPPKPLGHLLRNMDFRNLIIFSLTSKPAKWLTEL
                      LNMDAPMISVHVDHQINIIVSSVKSVTYHASNGSLRESRGKIDRQINFPGSLKQFVEH
                      LLHIFKEHKIDDLEIMCNSVYQYKPDIQALKGIKMRTLKLRYQESILKMFGTSADRIE
                      FSTIKSAELENSKDILIQNFDYLKFDFEGSLTADNLIGINGKVVNIWAADIEMKDVNI
                      FLKHWLKGSNPRLEKLILKTYRLDTDGNSVEDSVADLFRDLDCESIAESQFEILRDDE
                      SIAKISILEHGYWYNKIHCYFSLNEAMNLYTTYFWDIVHLQEKILH"
     gene             complement(join(7747. .8617,13260. .13342))
```

/locus_tag="Y63D3A.2"
CDS complement(join(7747. .8617,13260. .13342))
/locus_tag="Y63D3A.2"
/standard_name="Y63D3A.2"
/note="contains similarity to Pfam domain PF00646 (F-box domain)"
/codon_start=1
/product="Hypothetical protein Y63D3A.2"
/protein_id="CAA21709.1"
/db_xref="GI:3881235"
/db_xref="InterPro:IPR001810"
/db_xref="UniProt/TrEMBL:Q9XWG1"
/translation="MPALPFPILSLPPKSRRIALKCMDTLDLINISFTHPLAKRLAIE
LSVGIEVGSLRLLVRNGFKVFFSFYAPPHNIIADISYLSADGETFDLIKDEDVLWINS
KIKLRDLLVDLLDVLHLRKLGEVSVHTSWDLILDELRGIEIHQLIVGESITRQFKLLK
LQAEELELNVPLTQADVICHNLHRIALKSPGNIQCEDLIGMNMMHIEVFMMFDHSSLT
HREINLFLLHWVKGSNPRMRYYRSPLQEPLQFEEQFVAQILKGIPHRAITPGELEKLR
VQFRGKTFQFRGYHIERNEGDAKAMISVDPWSPIVEVFVWF"
gene complement(join(14704. .14833,15019. .15302,16493. .16724,16858. .16952,17625. .17845,17895. .17969,18033. .18084))
/locus_tag="Y63D3A.4"
CDS complement(join(14704. .14833,15019. .15302,16493. .16724,16858. .16952,17625. .17845,17895. .17969,18033. .18084))
/locus_tag="Y63D3A.4"
/standard_name="Y63D3A.4"
/note="contains similarity to Pfam domain PF03372 (Endonuclease/Exonuclease/phosphatase family)"
/codon_start=1
/product="Hypothetical protein Y63D3A.4"
/protein_id="CAA21707.1"
/db_xref="GI:3881233"
/db_xref="InterPro:IPR005135"
/db_xref="UniProt/TrEMBL:Q9XWG3"
/translation="MSNSDDEIQEIEAKRQKMSQEDSEVEIEILDEPEQGKLKNSSMS
DEQKLHEFAIITATDEAFAQSILQDVDWDLKKALDVFYGSEAFAEARSAAVMGASSSM
ASSGAAVMTAEDLKGFEVSVMSWNIDGLDGRSLLTRMKAVAHIVKNVNPDILFLQEVV
DRDLAPIDKLQSLYKIYYSNKGCQYYTAILVSKMPDVEKHDVIHFQNSGMYRTLQILE
GSIGGLKVFLLNTHLESTREHRPQRCAQFGPCMDKVREIIAQNPGALVFFGGDLNLRD
EEVSRVPDGVKDAWEAAGSDNKTKFTWDTFKNDNKQGFHGAKMRFDRLYWSGPLDKVK
FTLEGRQRIRSCLCFPSDHWAINATFFA"
gene join(19048. .19228,19281. .19426,20077. .21210)
/gene="tfg-1"
/locus_tag="Y63D3A.5"
CDS join(19048. .19228,19281. .19426,20077. .21210)
/gene="tfg-1"
/locus_tag="Y63D3A.5"
/standard_name="Y63D3A.5"
/note="C. elegans TFG-1 protein; contains similarity to Pfam domain PF00564 (PB1 domain)"
/codon_start=1
/product="Hypothetical protein Y63D3A.5"
/protein_id="CAB63398.1"
/db_xref="GI:6580329"
/db_xref="InterPro:IPR000270"
/db_xref="UniProt/TrEMBL:Q9U1W1"
/translation="MVHSNGAITSTILKARHADVVRKTSLHHANDLTLIDLVLNVQRL
LALPSDANFVLKYKDEEGDLVTLAEDSDLLLALHTSGATLDVTVVVDSRAREAVHDVQ
KQVEQIKLDVGKLLGALSALDVAQIAEQSNTSVANLSAPKQSHHDNIVFQKSFEAAPP
SPVPSEKAELPATIQPSVHEQFNHRPAHVEEEIPLENHYAPPPHQQIPDDLNTSFSSQ
PPPPIEQFGAIPPPNATIPSFPTSNAASPPVQEFAPPPPQQQQQQFQAPPPPMASHSS
ISSTPVQQQQGFAPPQQFGGPPPSGPPSEYGGYAPPQQQQQQFGAPPPQGAPQQGFGAP
PQGPPQGGPPQGSFGAPPPQQFHAPSPQSFGGPPPPVSSAPGNFAPPPQSGPPGAFAP
PPSAFGAPQGPGGPGGYGPPPPGGPGAPGSYGPPQGGPGGFGPPPPGGPGAYGPPPTG
FPPVGAPPPGAAGAPGGNPFARGPSATGYRQSPYQQ"
gene complement(join(21795. .21920,22061. .22165,23337. .23672,24863. .25354,26233. .26446,27066. .27309,27804. .27912,27973. .28209,29577. .29732,29793. .29908,29969. .30092))

Alignment Scores:
Pred. No.: 3.92e-05 Length: 33899
Score: 181.00 Matches: 74
Percent Similarity: 35.2% Conservative: 29
Best Local Similarity: 25.3% Mismatches: 81

Query Match: 9.5% Indels: 110
DB: 2 Gaps: 11

US-10-757-745-2 (1-362) x CEY63D3A (1-33899)

```
Qy        160 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 179
              :::|||||||||          |||:::|||                      ||||||
Db      15565 TTTTTGAAGAAAAATGAGGGAAATTTCGAG----------------------TATTTT 15530

Qy        180 ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer--------------Gln 194
              |||   |||:::|||       |||:::|||                        |||
Db      15529 CATCAGATTTCGCTTCGGAAATGTTTGGTGAGGCTAAGCCGAGTTTGTAGTTTGTAACCT 15470

Qy        195 GluIleIlePro-PheProSerThrLysMetMetArgAsnLeuLeuCysValHis----- 212
              |||       ||| ||||||      |||:::          |||:::   :::
Db      15469 GAGCCACAACCACTTTCCCGAAAAAAAACTGAAGAATTTCCTCGTTTTAATTGATTTTTT 15410

Qy        213 ------------------------------ValAsnValSerGlyAsn----------- 218
                                             :::     |||   |||
Db      15409 TGAAAGAAAACTGAACATTTCTATTAAAATCATTCGCTTTTCGGCAAATTTAACAAATTT 15350

Qy        219 ----------------------------------------GluLeuCysLeuMetT 224
                                                     ::::::   ||:::
Db      15349 TTGTAGTTTGTAGCCTGAGCCTCAACCAATTTTCTCGTTTTTTGCAGGTATTCCTACTTA 15290

Qy        224 hrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetV 244
               :::|||||||||||||||||   |||     :::|||          |||
Db      15289 ACACCCATCTCGAATCAACACGTGAGCACCGCCCGCAACGTTGTGCTCAATTCGGCTTCT 15230

Qy        244 alLeuLysLysMetGlnGlu---------AlaProGluSerAlaThrValIlePheAlaG 261
               :::  |||:::::|||                                |||  |||  |||  |
Db      15229 GCATGGACAAAGTCCGTGAAATCATCGCTCAAAACCCGGGCGCCCTCGTCTTTTCGGCG 15170

Qy        261 lyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleV 281
              |||||    |||||||||||||   |||||||:::|||          :::|||:::   :::
Db      15169 GCGATTTGAACTTACGAGACGAGGAGGTCAGCCGT---------GTGCCTGACGGTGTTA 15119

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnM 301
               |||  |||||||          |||                ::::::|||||||||||||
Db      15118 AGGATGCCTGGGAGGCTGCCGGCAGTGATAATAAGACAAAATTCACGTGGGACACCTTCA 15059

Qy        301 etAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg------------------- 314
               |||  |||                        |||:::|||
Db      15058 AAAATGACAATAAACAGGGATTCCATGGCGCTAAGATGAGGTAAAAGCGAAGAGACAAAT 14999

Qy        314 ------------------------------------------------------------ 314

Db      14998 AAAAAAGAAATAAAAAATATTTTGCCACACCGGGGGCTCGAACCCCGGTCGTGTGATTGG 14939

Qy        315 --------------------------PheAspArg-----------------------I 318
                                         ||||||                       |
Db      14938 TAAACGGCGAAAAGCGCACTGAGCCATTCGATCATTTTTTAAAAAATAAACAAAAATCGA 14879

Qy        318 lePhePheArg----------------------------------AlaAlaAlaGluG 326
               |||||||||                                  :::  |||
Db      14878 TTTTTTTCAATTTTTTAAAAAATCTAAAAAGCTTCAAATTTTCAGATTCGACCGCCTATA 14819

Qy        326 luGly------HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu---A 343
               |||      |||          :::     |||  |||  :::::::::
Db      14818 CTGGTCTGGACCACTGGACAAAGTCAAATTCACACTTGAA-GGCCGCCAACGCATCCGTT 14760

Qy        343 spCysGlyArgPheProSerAspHisTrpGlyLeu 354
               |||      ||||||||||||||||||   :::
Db      14759 CCTGTTTGTGCTTCCCATCTGATCATTGGGCTATC 14725
```

RESULT 55
AK113716
LOCUS AK113716 2036 bp mRNA linear INV 30-NOV-2002
DEFINITION Ciona intestinalis cDNA, clone:cicl005m10, full insert sequence.
ACCESSION AK113716
VERSION AK113716.1 GI:23584533
KEYWORDS FLI_CDNA.
SOURCE Ciona intestinalis

```
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1
  AUTHORS   Satou,Y., Yamada,L., Mochizuki,Y., Takatori,N., Kawashima,T.,
            Sasaki,A., Hamaguchi,M., Awazu,S., Yagi,K., Sasakura,Y.,
            Nakayama,A., Ishikawa,H., Inaba,K. and Satoh,N.
  TITLE     A cDNA resource from the basal chordate Ciona intestinalis
  JOURNAL   Genesis 33 (4), 153-154 (2002)
   PUBMED   12203911
REFERENCE   2  (bases 1 to 2036)
  AUTHORS   Satou,Y. and Satoh,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2002) Nori Satoh, Kyoto University, Department of
            Zoology; Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            (E-mail:satoh@ascidian.zool.kyoto-u.ac.jp, Tel:81-75-753-4095,
            Fax:81-75-705-1113)
COMMENT     Ciona intestinalis cDNA Project (URL:
            http://ghost.zool.kyoto-u.ac.jp/indexr1.html).
FEATURES             Location/Qualifiers
     source          1. .2036
                     /organism="Ciona intestinalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:7719"
                     /clone="cicl005m10"
ORIGIN

Alignment Scores:
Pred. No.:                4.11e-05        Length:          2036
Score:                    165.00          Matches:         56
Percent Similarity:       40.9%           Conservative:    29
Best Local Similarity:    26.9%           Mismatches:      95
Query Match:              8.7%            Indels:          28
DB:                       2               Gaps:            5

US-10-757-745-2 (1-362) x AK113716 (1-2036)

Qy        116 LeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn--------------AsnLeuSer 130
Db        142 ATACTTAGTTGGAATTTGCATTGTCGACCTCCCCATACCCGCTTGCATAAAAACAGAACT 201

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db        202 GAAAGAATAGAAAATGTTTGTCAAACAATACAAAGGGAGAATCCAACATTGGTTAGTCTT 261

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db        262 CAAGAGGTT------GGGGTTTATGAATATAATAAAATTCATAATCATTTCGTAAATATC 315

Qy        171 IleThrGlyHisGlu--------------------------------GluGlyTyrPhe 179
Db        316 TATGGTTTCCACTCAGTGAAAAGTTCCATTGGATTCGTTGAAGACTTGGATTTCTGTATC 375

Qy        180 ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 199
Db        376 ACAGTCGTAATGTTCAAGAAAGATGTTGTTTTTCATTATTCTGGAAATCATTATAAATTC 435

Qy        200 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 219
Db        436 CACAGCCTTCGAATGTGTGAATATTTGCATGTGTTTCAAGGAAAAGTTAAAGGAAATGAA 495

Qy        220 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
Db        496 TTTGTATTGTTTACATCTCAACTTGAAAAAAGAAAATCAAGTTCTAAAGAAAGACAAGAC 555

Qy        240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 259
Db        556 CAATTGGGCCAGGTTATTCATTTCCTTAACAAAACTCCAGCCACATCAACAGCTATATTT 615

Qy        260 AlaGlyAspThrAsnLeuArgAspArgGluValThrArg------CysGlyGlyLeuPro 277
Db        616 GCTGGTAGTTTATACATTACAGATTTTGAGCTAGAGAAGGAAATGTCGCAACGAGATGAG 675

Qy        278 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db        676 AGAAATGTTTGTGAGTTGTGGAAAGAATTAAAACTCAAAAAA---------------- 717

Qy        298 AspThrGlnMetAsnSerAsnLeu 305
Db        718 ------TCTTTCAATGAAAATCTT 735

RESULT 56
AP008218_128
WPCOMMENT
  Sequence split into 276 fragments  LOCUS AP008218 Accession AP008218
     Fragment Name       Begin      End
     AP008218_000            1   110000
     AP008218_001       100001   210000
     AP008218_002       200001   310000
     AP008218_003       300001   410000
     AP008218_004       400001   510000
     AP008218_005       500001   610000
     AP008218_006       600001   710000
     AP008218_007       700001   810000
     AP008218_008       800001   910000
     AP008218_009       900001  1010000
     AP008218_010      1000001  1110000
     AP008218_011      1100001  1210000
     AP008218_012      1200001  1310000
     AP008218_013      1300001  1410000
     AP008218_014      1400001  1510000
     AP008218_015      1500001  1610000
     AP008218_016      1600001  1710000
     AP008218_017      1700001  1810000
     AP008218_018      1800001  1910000
     AP008218_019      1900001  2010000
     AP008218_020      2000001  2110000
     AP008218_021      2100001  2210000
     AP008218_022      2200001  2310000
     AP008218_023      2300001  2410000
     AP008218_024      2400001  2510000
     AP008218_025      2500001  2610000
     AP008218_026      2600001  2710000
     AP008218_027      2700001  2810000
     AP008218_028      2800001  2910000
     AP008218_029      2900001  3010000
     AP008218_030      3000001  3110000
     AP008218_031      3100001  3210000
     AP008218_032      3200001  3310000
     AP008218_033      3300001  3410000
     AP008218_034      3400001  3510000
     AP008218_035      3500001  3610000
     AP008218_036      3600001  3710000
     AP008218_037      3700001  3810000
     AP008218_038      3800001  3910000
     AP008218_039      3900001  4010000
     AP008218_040      4000001  4110000
     AP008218_041      4100001  4210000
     AP008218_042      4200001  4310000
     AP008218_043      4300001  4410000
     AP008218_044      4400001  4510000
     AP008218_045      4500001  4610000
     AP008218_046      4600001  4710000
     AP008218_047      4700001  4810000
     AP008218_048      4800001  4910000
     AP008218_049      4900001  5010000
     AP008218_050      5000001  5110000
     AP008218_051      5100001  5210000
     AP008218_052      5200001  5310000
     AP008218_053      5300001  5410000
     AP008218_054      5400001  5510000
     AP008218_055      5500001  5610000
     AP008218_056      5600001  5710000
     AP008218_057      5700001  5810000
     AP008218_058      5800001  5910000
     AP008218_059      5900001  6010000
```

AP008218_060	6000001	6110000
AP008218_061	6100001	6210000
AP008218_062	6200001	6310000
AP008218_063	6300001	6410000
AP008218_064	6400001	6510000
AP008218_065	6500001	6610000
AP008218_066	6600001	6710000
AP008218_067	6700001	6810000
AP008218_068	6800001	6910000
AP008218_069	6900001	7010000
AP008218_070	7000001	7110000
AP008218_071	7100001	7210000
AP008218_072	7200001	7310000
AP008218_073	7300001	7410000
AP008218_074	7400001	7510000
AP008218_075	7500001	7610000
AP008218_076	7600001	7710000
AP008218_077	7700001	7810000
AP008218_078	7800001	7910000
AP008218_079	7900001	8010000
AP008218_080	8000001	8110000
AP008218_081	8100001	8210000
AP008218_082	8200001	8310000
AP008218_083	8300001	8410000
AP008218_084	8400001	8510000
AP008218_085	8500001	8610000
AP008218_086	8600001	8710000
AP008218_087	8700001	8810000
AP008218_088	8800001	8910000
AP008218_089	8900001	9010000
AP008218_090	9000001	9110000
AP008218_091	9100001	9210000
AP008218_092	9200001	9310000
AP008218_093	9300001	9410000
AP008218_094	9400001	9510000
AP008218_095	9500001	9610000
AP008218_096	9600001	9710000
AP008218_097	9700001	9810000
AP008218_098	9800001	9910000
AP008218_099	9900001	10010000
AP008218_100	10000001	10110000
AP008218_101	10100001	10210000
AP008218_102	10200001	10310000
AP008218_103	10300001	10410000
AP008218_104	10400001	10510000
AP008218_105	10500001	10610000
AP008218_106	10600001	10710000
AP008218_107	10700001	10810000
AP008218_108	10800001	10910000
AP008218_109	10900001	11010000
AP008218_110	11000001	11110000
AP008218_111	11100001	11210000
AP008218_112	11200001	11310000
AP008218_113	11300001	11410000
AP008218_114	11400001	11510000
AP008218_115	11500001	11610000
AP008218_116	11600001	11710000
AP008218_117	11700001	11810000
AP008218_118	11800001	11910000
AP008218_119	11900001	12010000
AP008218_120	12000001	12110000
AP008218_121	12100001	12210000
AP008218_122	12200001	12310000
AP008218_123	12300001	12410000
AP008218_124	12400001	12510000
AP008218_125	12500001	12610000
AP008218_126	12600001	12710000
AP008218_127	12700001	12810000
AP008218_128	12800001	12910000
AP008218_129	12900001	13010000
AP008218_130	13000001	13110000
AP008218_131	13100001	13210000
AP008218_132	13200001	13310000
AP008218_133	13300001	13410000
AP008218_134	13400001	13510000
AP008218_135	13500001	13610000
AP008218_136	13600001	13710000
AP008218_137	13700001	13810000
AP008218_138	13800001	13910000
AP008218_139	13900001	14010000
AP008218_140	14000001	14110000
AP008218_141	14100001	14210000
AP008218_142	14200001	14310000
AP008218_143	14300001	14410000
AP008218_144	14400001	14510000
AP008218_145	14500001	14610000
AP008218_146	14600001	14710000
AP008218_147	14700001	14810000
AP008218_148	14800001	14910000
AP008218_149	14900001	15010000
AP008218_150	15000001	15110000
AP008218_151	15100001	15210000
AP008218_152	15200001	15310000
AP008218_153	15300001	15410000
AP008218_154	15400001	15510000
AP008218_155	15500001	15610000
AP008218_156	15600001	15710000
AP008218_157	15700001	15810000
AP008218_158	15800001	15910000
AP008218_159	15900001	16010000
AP008218_160	16000001	16110000
AP008218_161	16100001	16210000
AP008218_162	16200001	16310000
AP008218_163	16300001	16410000
AP008218_164	16400001	16510000
AP008218_165	16500001	16610000
AP008218_166	16600001	16710000
AP008218_167	16700001	16810000
AP008218_168	16800001	16910000
AP008218_169	16900001	17010000
AP008218_170	17000001	17110000
AP008218_171	17100001	17210000
AP008218_172	17200001	17310000
AP008218_173	17300001	17410000
AP008218_174	17400001	17510000
AP008218_175	17500001	17610000
AP008218_176	17600001	17710000
AP008218_177	17700001	17810000
AP008218_178	17800001	17910000
AP008218_179	17900001	18010000
AP008218_180	18000001	18110000
AP008218_181	18100001	18210000
AP008218_182	18200001	18310000
AP008218_183	18300001	18410000
AP008218_184	18400001	18510000
AP008218_185	18500001	18610000
AP008218_186	18600001	18710000
AP008218_187	18700001	18810000
AP008218_188	18800001	18910000
AP008218_189	18900001	19010000
AP008218_190	19000001	19110000
AP008218_191	19100001	19210000
AP008218_192	19200001	19310000
AP008218_193	19300001	19410000
AP008218_194	19400001	19510000
AP008218_195	19500001	19610000
AP008218_196	19600001	19710000

Alignment Scores:

Pred. No.:	0.0192	Length:	110000
Score:	158.00	Matches:	57
Percent Similarity:	42.8%	Conservative:	32
Best Local Similarity:	27.4%	Mismatches:	75
Query Match:	8.3%	Indels:	44
DB:	15	Gaps:	8

```
US-10-757-745-2 (1-362) x AP008218_128 (1-110000)

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 203
              ::::::   |||::::::   ::::::|||   :::      ||||||   ::::::
Db      27848 ATGCAGATGAGCAAATTGCCTGTTGAGTCTTTTGACTGCACTCCATTTTCCAACTCAATA 27907

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn---GluLeuCysLeu 222
              |||   |||   |||          ||||||          |||       :::|||   |||
Db      27908 ATGGGAAGAGAACTGTGTGTAGCACATGTCAAAACTGGAGGTGCAGTTAAGCTAGTGCTG 27967

Qy        223 MetThrSerHisLeuGluSer----------------------ThrArgGlyHisAla 234
              ||||||||||||||||||                        :::      ::::::
Db      27968 GCCACTAGCCACCTTGAGAGTCCAATGCCTGGGCCTCCTACTTGGGATCAGATGTACAGC 28027

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
              ||||||:::   |||          ||||||   :::
Db      28028 ACGGAACGAGTTGCTCAGGCAAACAAATCTCTCAAAATTTTG---------GGATCCTTC 28078

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
              |||||||||   ||||||   |||      ||:::          |||
Db      28079 CGCAATGTAATATTTTGTGGCGATATGAACTGGGATGACAAA--------------GGA 28123

Qy        275 Gly------------LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 290
                          ||||||          |||   |||       |||   ||||||
Db      28124 GATGGGCCCTTCCCCCTACCAGCTGGCTGGACCGATGCTTGGATTGAGTTG---AAGCCG 28180

Qy        291 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 310
                    :::|||:::||||||:::   ||||||   |||
Db      28181 GGTGAAGATGGCTGGACTTATGACACAAAAGCTAACAGTATGCTATCAGCCAACCGTAAG 28240

Qy        311 CysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 330
              :::   |||   ||||||          :::   |||
Db      28241 TTACAGAAGAGATTGGACCGATTCGTGTGCAAGCTGGCA------------GATTTCAAG 28288

Qy        331 ProArgSerLeuAspLeuLeuGly---------------------------------- 338
              ::::::   ::::::|||
Db      28289 ATTAACAACATCCAGATGATCGGGAAGAATGCGATACCTGGTCTCTCATATGTCAAGGAG 28348

Qy        339 ------------------LeuGluLysLeuAspCysGlyArgPheProSerAspHisTrp 352
                              :::   ||||||:::              |||||||||||:::
Db      28349 AAGAAGGTCCGCAAGGAAGTTCGCAAACTGGAACTTCCTGTCTTACCCAGTGACCACTTT 28408

Qy        353 GlyLeuLeuCysAsnLeuAspIle 360
              ||||||:::   ::::::   :::
Db      28409 GGACTTGTTTTGAGCATTACTCTT 28432

RESULT 57
AC149575/c
LOCUS       AC149575              111678 bp    DNA     linear   PLN 09-JUN-2004
DEFINITION  Populus trichocarpa clone Pop1-41G18, complete sequence.
ACCESSION   AC149575
VERSION     AC149575.1  GI:48475260
KEYWORDS    HTG.
SOURCE      Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
  ORGANISM  Populus trichocarpa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 111678)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 111678)
  AUTHORS   Stanford Human Genome Center.
  CONSRTM   DOE Joint Genome Institute
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUN-2004) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;

            Estimated Total Number of Errors is 0.1.
FEATURES             Location/Qualifiers
     source          1. .111678
                     /organism="Populus trichocarpa"
                     /mol_type="genomic DNA"
                     /sub_species="trichocarpa"
                     /db_xref="taxon:3694"
                     /clone="Pop1-41G18"
ORIGIN

Alignment Scores:
Pred. No.:              0.0195          Length:         111678
Score:                  158.00          Matches:        57
Percent Similarity:     43.8%           Conservative:   35
Best Local Similarity:  27.1%           Mismatches:     66
Query Match:            8.3%            Indels:         52
DB:                     15              Gaps:           9

US-10-757-745-2 (1-362) x AC149575 (1-111678)

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 203
              |||:::   |||::::::   :::|||||||             ||||||   ::::::
Db      61185 TTGCAGTTAAGCAAACTGCCAGTAAAATCCTTCAGCACTAAACCCTTTATGAATTCAATA 61126

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu---LeuCysLeu 222
              |||   |||   |||          :::   |||   |||   :::   |||   :::
Db      61125 ATGGGGAGAGAACTTTGCATTGCTGAGCTTGAAGTTCCAGGAAAAAAGTCATTAGTTGTT 61066

Qy        223 MetThrSerHisLeuGluSer----------------------ThrArgGlyHisAla 234
              ||||||||||||||||||                        :::      :::
Db      61065 GCCACTAGCCATCTAGAGAGTCCATGTCCAGCACCTCCAAAATGGGATCAGATGTTCAGC 61006

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
              ||||||:::::|||   |||      :::      ::::::          ::::::
Db      61005 AAGGAACGTGTGGATCAGGCAAAGGAGGCCATCAACCTTCTCAAG---------AAGAAC 60955

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
              :::   |||||||||   ||||||   |||      |||             :::
Db      60954 TCAAATGTGATTTTTTGTGGTGACATGAACTGGGATGAT------------AAGTTGGAT 60907

Qy        275 Gly---------LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
              |||             |||:::      ||||||   |||      |||   ||||||
Db      60906 GGTCAGTTTCCTTTCCCTGATGGATGGGTTGATGCCTGGGTAGAGTTG---AAGCCAGGG 60850

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
                  :::|||:::||||||:::   |||      |||                |||
Db      60849 GACAATGGATGGACATATGATACCAAATCCAATCAAATGTTGTCTGGTAATCGTGCATTG 60790

Qy        312 LysLeuArgPheAspArgIlePhe--------------PheArgAlaAlaAlaGluGlu 326
              :::   |||   ||||||                  |||:::   :::
Db      60789 CAAAAGAGACTGGATCGATTCATATGCAGTCTGTGTGATTTCAAGATCAGT--------- 60739

Qy        327 GlyHisIleIleProArgSerLeuAspLeuLeuGly---------------------- 338
                              :::|||::::::|||
Db      60738 ------------------AAAATTGACATGATTGGTAAGGATGCAATACCTGGCTTATCA 60697

Qy        339 ----------------------------LeuGluLysLeuAspCysGlyArgPhePro 348
                                         ::::::   |||:::          |||
Db      60696 TATATGAAGGAGAAGAAAGTGAGAAAAGAGGTAAAGATGTTGGAGTTACCTGTCTTGCCT 60637

Qy        349 SerAspHisTrpGlyLeuLeuCysAsnLeu 358
              |||||||||:::||||||||||      :::
Db      60636 AGTGACCATTATGGTTTGCTTTTAACAATA 60607

RESULT 58
CNS08CB2/c
LOCUS       CNS08CB2              156724 bp    DNA     linear   PLN 21-NOV-2003
DEFINITION  Oryza sativa chromosome 12, . BAC OSJNBa0036H15 of library OSJNBa
            from chromosome 12 of cultivar Nipponbare of ssp. japonica of Oryza
            sativa (rice), complete sequence.
ACCESSION   AL844879
VERSION     AL844879.5  GI:28625295
KEYWORDS    HTG.
```

```
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 156724)
  AUTHORS   Choisne,N., Orjeda,G., Cattolico,L., Demange,N., Wincker,P.,
            Segurens,B., Pelletier,E., Scarpelli,C., Salanoubat,M.,
            Weissenbach,J. and Quetier,F.
  TITLE     Oryza sativa chromosome 12 sequencing
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 156724)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2003) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Feb 28, 2003 this sequence version replaced gi:28460625.
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            ---------------
            The following sequence is oriented from the T7 to the SP6 end. The
            nucleotide sequence of this BAC clone was generated by combining
            Syngenta and Genoscope sequencing data.
            Upstream BAC (overlapping the T7 end) : OSJNBa0017G24 (AC=AL928777)
            Downstream BAC (overlapping the SP6 end) : OSJNBa0032G05
            (AC=AL831800) ---------- Finishing boundaries
            FINISHED SEGMENT STARTS AT BASE 5830
            FINISHED SEGMENT ENDS AT BASE 100667
            ---------------.
FEATURES             Location/Qualifiers
     source          1. .156724
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /sub_species="japonica"
                     /db_xref="taxon:39947"
                     /chromosome="12"
                     /clone="OSJNBa0036H15"
                     /clone_lib="OSJNBa"
ORIGIN

Alignment Scores:
Pred. No.:                  0.029          Length:          156724
Score:                      158.00         Matches:         57
Percent Similarity:         42.8%          Conservative:    32
Best Local Similarity:      27.4%          Mismatches:      75
Query Match:                8.3%           Indels:          44
DB:                         15             Gaps:            8

US-10-757-745-2 (1-362) x CNS08CB2 (1-156724)

Qy      184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 203
            ::::::   |||::::::   ::::::|||   :::       ::::::   ::::::
Db    91379 ATGCAGATGAGCAAATTGCCTGTTGAGTCTTTGACTGCACTCCATTTTCCAACTCAATA 91320

Qy      204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn---GluLeuCysLeu 222
            |||   |||  |||       ||||||         |||      :::|||   |||
Db    91319 ATGGGAAGAGAACTGTGTGTAGCACATGTCAAAACTGGAGGTGCAGTTAAGCTAGTGCTG 91260

Qy      223 MetThrSerHisLeuGluSer----------------------ThrArgGlyHisAla 234
              ||||||||||||||||||                    :::   ::::::
Db    91259 GCCACTAGCCACCTTGAGAGTCCAATGCCTGGGCCTCCTACTTGGGATCAGATGTACAGC 91200

Qy      235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
            ::::::   |||       ::::::   :::
Db    91199 ACGGAACGAGTTGCTCAGGCAAACAAATCTCTCAAAATTTTG---------GGATCCTTC 91149

Qy      255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
              |||||||||   ::::::   |||      |||:::                 |||
Db    91148 CGCAATGTAATATTTTGTGGCGATATGAACTGGGATGACAAA--------------GGA 91104

Qy      275 Gly-----------LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 290
                        ::::::           |||   |||       |||   ::::::
Db    91103 GATGGGCCCTTCCCCCTACCAGCTGGCTGGACCGATGCTTGGATTGAGTTG---AAGCCG 91047

Qy      291 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 310
                  :::|||:::||::::::   ::::::   |||
Db    91046 GGTGAAGATGGCTGGACTTATGACACAAAAGCTAACAGTATGCTATCAGCCAACCGTAAG 90987

Qy      311 CysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 330
              :::   |||   ::::::         :::   |||
Db    90986 TTACAGAAGAGATTGGACCGATTCGTGTGCAAGCTGGCA-----------GATTTCAAG 90939

Qy      331 ProArgSerLeuAspLeuLeuGly---------------------------------- 338
                    ::::::   ::::::|||
Db    90938 ATTAACAACATCCAGATGATCGGGAAGAATGCGATACCTGGTCTCTCATATGTCAAGGAG 90879

Qy      339 -----------------LeuGluLysLeuAspCysGlyArgPheProSerAspHisTrp 352
                            :::   ::::::|::   ::::::::::::::::::
Db    90878 AAGAAGGTCCGCAAGGAAGTTCGCAAACTGGAACTTCCTGTCTTACCCAGTGACCACTTT 90819

Qy      353 GlyLeuLeuCysAsnLeuAspIle 360
            ::::::|::   ::::::   :::
Db    90818 GGACTTGTTTTGAGCATTACTCTT 90795

RESULT 59
AC151708
LOCUS       AC151708               86100 bp    DNA     linear   PLN 16-DEC-2004
DEFINITION  Medicago truncatula clone mth2-15h5, complete sequence.
ACCESSION   AC151708
VERSION     AC151708.6  GI:56461838
KEYWORDS    HTG.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 86100)
  AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
  TITLE     Medicago truncatula BAC Clone mth2-15h5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 86100)
  AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-SEP-2004) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   3  (bases 1 to 86100)
  AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2004) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   4  (bases 1 to 86100)
  AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-2004) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   5  (bases 1 to 86100)
  AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-2004) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Dec 9, 2004 this sequence version replaced gi:54306117.
            ---------- Genome Center
```

```
           Center: Department Of Chemistry And Biochemistry
           The University Of Oklahoma
           Center code:UOKNOR
           ----------------------.
FEATURES             Location/Qualifiers
     source          1. .86100
                     /organism="Medicago truncatula"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3880"
                     /clone="mth2-15h5"
                     /clone_lib="Medicago truncatula BAC library H2"
ORIGIN

Alignment Scores:
Pred. No.:              0.0332          Length:         86100
Score:                  154.00          Matches:        56
Percent Similarity:     41.6%           Conservative:   35
Best Local Similarity:  25.6%           Mismatches:     82
Query Match:            8.1%            Indels:         46
DB:                     15              Gaps:           9

US-10-757-745-2 (1-362) x AC151708 (1-86100)

Qy        175 GluGluGlyTyrPheThrAlaIleMet----------------LeuLysLysSerArg 188
              |||   |||   ::::::   ::::::                 :::   |||:::
Db      64637 GAAAGAGGGTGTTATTCACATGTACTGAAATATCTGTCAAATTAATTTCAGTTAAGCAAA 64696

Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
              :::   :::||||||                |||   ::::::   |||   |||   |||
Db      64697 CTGCCGGTGAAATCTTTCTCCGCCAAATCATTTAGCAACTCTATAATGGGAAGAGAACTT 64756

Qy        209 LeuCysValHisVal---AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
                  |||   :::|||   |||             :::   ||||||||||||
Db      64757 TGCATTGCTGAAGTGGAGGATGTGGGCGGCAAGTCGTTTGTTGTGGCCACTAGCCATCTC 64816

Qy        228 GluSer-----------------------ThrArgGlyHisAlaAlaGluArgMetAsn 239
              ||||||                      :::     :::   ||||||:::
Db      64817 GAGAGTCCTTGTCCGGCCCCTCCAAAATGGGATCAGATGTTCAGCAAGGAACGAGTGGAG 64876

Qy        240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 259
              |||           |||     ::::::     |||           |||:::|||
Db      64877 CAGGCTAATGAGGCTCTGAACATTCTCAAAAGGCACCCT---------AATGTTGTTTTT 64927

Qy        260 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro------ 277
              ||||||   |||     |||:::     :::   |||     |||
Db      64928 GGGGGTGATATGAACTGGGATGACAAA---------AAGGATGGTCAATATCCTTTGCAA 64978

Qy        278 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
              :::     :::|||   |||     |||   :::|||           :::|||:::
Db      64979 GACGGGTGGCTTGATGCCTGGTCTGTACTA---AGACCAAATGAAGCTGGTTGGACATAT 65035

Qy        298 AspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArg 317
              ||||||:::   |||     |||             :::   |||   ||||||
Db      65036 GATACCAAGTCAAACCAAATGTTGACAGGCAATCGTACTCTCCAAAAGCGATTGGATCGT 65095

Qy        318 IlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu 337
                  |||                           ::::::|||::::::
Db      65096 TTTGTTTGTCGTTTACGT-----------GATTTTAAGATAAGCAATATTGACATGATT 65143

Qy        338 GlyLeuGlu---------------------------------------------------- 340
              |||::::::
Db      65144 GGGATGGATGAAATACCTGGTGTTTCATATAACAAAGAGAAGAAAGTAAGAGGCGAGATC 65203

Qy        341 ---LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358
                 :::|||   |||           ||||||||||||:::|||||||||     |||
Db      65204 AAACAACTGGTATGCCCTGTTTTGCCCAGCGATCACTATGGTCTCCTTTTGACACTT 65260

RESULT 60
AC134967/c
LOCUS       AC134967               133028 bp    DNA     linear   PLN 15-AUG-2003
DEFINITION  Medicago truncatula clone mth2-26b8, complete sequence.
ACCESSION   AC134967
VERSION     AC134967.19  GI:33589882
KEYWORDS    HTG.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 133028)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Medicago truncatula BAC Clone mth2-26b8
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 133028)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-2002) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   3  (bases 1 to 133028)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-AUG-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   4  (bases 1 to 133028)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Aug 12, 2003 this sequence version replaced gi:33342303.
            ------------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            ----------------------.
FEATURES             Location/Qualifiers
     source          1. .133028
                     /organism="Medicago truncatula"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3880"
                     /clone="mth2-26b8"
                     /clone_lib="Medicago truncatula BAC library H2"
ORIGIN

Alignment Scores:
Pred. No.:              0.0553          Length:         133028
Score:                  154.00          Matches:        56
Percent Similarity:     41.6%           Conservative:   35
Best Local Similarity:  25.6%           Mismatches:     82
Query Match:            8.1%            Indels:         46
DB:                     15              Gaps:           9

US-10-757-745-2 (1-362) x AC134967 (1-133028)

Qy        175 GluGluGlyTyrPheThrAlaIleMet----------------LeuLysLysSerArg 188
              |||   |||   ::::::   ::::::                 :::   |||:::
Db      68387 GAAAGAGGGTGTTATTCACATGTACTGAAATATCTGTCAAATTAATTTCAGTTAAGCAAA 68328

Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
              :::   :::||||||                |||   ::::::   |||   |||   |||
Db      68327 CTGCCGGTGAAATCTTTCTCCGCCAAATCATTTAGCAACTCTATAATGGGAAGAGAACTT 68268

Qy        209 LeuCysValHisVal---AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
                  |||   :::|||   |||             :::   ||||||||||||
Db      68267 TGCATTGCTGAAGTGGAGGATGTGGGCGGCAAGTCGTTTGTTGTGGCCACTAGCCATCTC 68208

Qy        228 GluSer-----------------------ThrArgGlyHisAlaAlaGluArgMetAsn 239
              ||||||                      :::     :::   ||||||:::
Db      68207 GAGAGTCCTTGTCCGGCCCCTCCAAAATGGGATCAGATGTTCAGCAAGGAACGAGTGGAG 68148
```

```
Qy        240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 259
Db      68147 CAGGCTAATGAGGCTCTGAACATTCTCAAAAGGCACCCT---------AATGTTGTTTTT 68097

Qy        260 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro------ 277
Db      68096 GGGGGTGATATGAACTGGGATGACAAA---------AAGGATGGTCAATATCCTTTGCAA 68046

Qy        278 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db      68045 GACGGGTGGCTTGATGCCTGGTCTGTACTA---AGACCAAATGAAGCTGGTTGGACATAT 67989

Qy        298 AspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArg 317
Db      67988 GATACCAAGTCAAACCAAATGTTGACAGGCAATCGTACTCTCCAAAAGCGATTGGATCGT 67929

Qy        318 IlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu 337
Db      67928 TTTGTTTGTCGTTTACGT------------GATTTTAAGATAAGCAATATTGACATGATT 67881

Qy        338 GlyLeuGlu---------------------------------------------------- 340
Db      67880 GGGATGGATGAAATACCTGGTGTTTCATATAACAAAGAGAAGAAAGTAAGAGGCGAGATC 67821

Qy        341 ---LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358
Db      67820 AAACAACTGGTATGCCCTGTTTTGCCCAGCGATCACTATGGTCTCCTTTTGACACTT 67764
```

```
RESULT 61
AY228171
LOCUS       AY228171               2709 bp    DNA     linear   INV 27-MAR-2003
DEFINITION  Trypanosoma brucei RNA editing complex protein MP100 (MP100) gene,
            complete cds; nuclear gene for mitochondrial product.
ACCESSION   AY228171
VERSION     AY228171.1  GI:29293901
KEYWORDS    .
SOURCE      Trypanosoma brucei
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 2709)
  AUTHORS   Panigrahi,A.K., Schnaufer,A., Ernst,N.L., Wang,B., Carmean,N.,
            Salavati,R. and Stuart,K.
  TITLE     Identification of novel components of Trypanosoma brucei editosomes
  JOURNAL   RNA 9 (4), 484-492 (2003)
   PUBMED   12649499
REFERENCE   2  (bases 1 to 2709)
  AUTHORS   Panigrahi,A.K., Ernst,N.L. and Stuart,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2003) Seattle Biomedical Research Institute, 4
            Nickerson St, Seattle, WA 98109, USA
FEATURES             Location/Qualifiers
     source          1. .2709
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5691"
     gene            <1. .>2709
                     /gene="MP100"
     mRNA            <1. .>2709
                     /gene="MP100"
                     /product="RNA editing complex protein MP100"
     CDS             1. .2709
                     /gene="MP100"
                     /note="protein identified by mass spectrometry;
                     endonuclease/exonuclease/phosphatase protein family;
                     similar to Trypanosoma brucei MP99; putative exonuclease"
                     /codon_start=1
                     /product="RNA editing complex protein MP100"
                     /protein_id="AAO63565.1"
                     /db_xref="GI:29293902"
                     /translation="MALAQSCRHFAATTYRQSKITDRFQLLITKKEIECDYVAVDATA
                     LVATGLRLAKNVTTEQRRHKEVARHVVQSIQQLLKKVRCKKFLLIAMDGAENLLKADR
```

```
                     TRGSSLTRKVESRLMRLPGTPLMQAVEERIVRMMPERQILPGEVVFAGTCVQGCVEQK
                     MSAWALDLASRDTFNGSSDSLNLIGASELYLNVLALSPFYNVSSVVQNNADLRHIRLQ
                     DILEWLELDKKAKEGESVTIAKMRTDILFLFIIANGASATELNPIPAVGFHEIVDRYM
                     KVLAENGATPPSQSATASAPTASANLPAAPPFLFEDSPGNTLQLNLRMLCRIMQLVAR
                     KEAPARVDGLSEAYLEHALQTHAMLCTGQAPQCTYLPAAGTGQLGSTAPSAVQLAGHL
                     AALCASTAQRNARCTSLQPSLVTPEKGSCEEDSPATKLPPTPPTSAAAPVEDKPLTAA
                     EYTILCQSLPAAVEGLIHQYVGIAPKPDVGKMITTATTEEAYRLVREVLSYANPLRPH
                     KCLCLSPSYCWLQNEKTQLWRFEYVDIGVKSHELGTRRQLNAMKGVTLEVNMSRDGPS
                     HFDPLSGTWEPIANFPCGAECEYDKHTEECTSSPSAPQTAKEVEGGPVKASLKLLTWN
                     VMFDRYSGKPTPLGMPGIDWCSPKRYPVIAKIIEEEDADVVGMQEVEPVFWEFLSKRA
                     LIRQRYYFSCGHMSPAIAPWGVLMLIHRRRLPVQSINYLNVPAWTNHVSLMPVVGLKM
                     VHGTVHIAAAHLLAPYTKSHENARTSQDTALRHHMTKTLGGGDVVTMGDFNDWPTNEF
                     IMPHESQYVDCWPVLHPNNPGKTMDETNTFCKLKVEEMFFGRSDKVFLRSRRLVPVEA
                     HLVGTRSVNDENNNADAPAYLFPSDHYGVSITFSLKVPEVQIIGAL"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                0.00201        Length:          2709
Score:                    148.00         Matches:         66
Percent Similarity:       39.0%          Conservative:    56
Best Local Similarity:    21.1%          Mismatches:      127
Query Match:              7.8%           Indels:          64
DB:                       2              Gaps:            14

US-10-757-745-2 (1-362) x AY228171 (1-2709)

Qy         89 AsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGln 108
Db       1768 GACAAACATACTGAAGAATGCACATCTTCTCCTTCAGCTCCACAAACTGCAAAAGAGGTT 1827

Qy        109 GluAsnGly---------SerMetPheSerLeuIleThrTrpAsnIle------------ 121
Db       1828 GAAGGAGGACCAGTTAAGGCATCTTTAAAATTGCTGACATGGAATGTCATGTTTGATCGT 1887

Qy        122 ------------------------------AspGlyLeuAspLeuAsnAsnLeuSerGlu 131
Db       1888 TACAGCGGCAAACCAACTCCTTTAGGTATGCCTGGTATAGATTGGTGCTCC---CCCAAG 1944

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
Db       1945 AGGTATCCCGTCATTGCCAAGATCATAGAGGAGGAGGATGCGGATGTTGTTGGCATGCAA 2004

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSer---------SerAsnTyr 168
Db       2005 GAGGTGGAGCCTGTATTTTGGGAGTTCCTTTCCAAGCGGGCACTTATACGCCAAAGGTAC 2064

Qy        169 GluIleIleThrGlyHisGluGlu------------GlyTyrPheThrAlaIleMetLeu 184
Db       2065 TACTTTTCGTGCGGTCATATGAGTCCAGCTATTGCCCCGTGGGGGGTGCTGATGTTGATA 2124

Qy        185 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer---ThrLys 203
Db       2125 CACCGTCGGCGTCTTCCAGTACAGAGCATCAACTATCTTAACGTCCCTGCGTGGACTAAT 2184

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMet 223
Db       2185 CATGTTTCTCTGATGCCGGTTGTGGGTCTGAAGATGGTGCACGGTACTGTACACATAGCT 2244

Qy        224 ThrSerHisLeuGluSer------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 241
Db       2245 GCCGCTCACTTACTCGCGCCATACACGAAGTCCCACGAGAACGCTCGTACTTCCCAGGAT 2304

Qy        242 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 261
Db       2305 ACCGCGTTACGCCATCACATGACGAAGACGTTGGGTGGCGGGGATGTGGTAACTATGGGA 2364

Qy        262 AspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleVal 281
Db       2365 GATTTCAAC----------------------------GACTGGCCTACAAAT------ 2388

Qy        282 AspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr------TrpAspThr 299
Db       2389 ---------GAATTTATT---ATGCCACATGAATCACAGTACGTGGACTGTTGGCCCGTT 2436
```

```
Qy        300 GlnMetAsnSerAsnLeuGlyIleThr--------------AlaAlaCysLysLeuArg 314
Db       2437 TTGCACCCCAACAACCCTGGAAAGACTATGGATGAAACAAACACCTTTTGTAAACTGAAA 2496

Qy        315 PheAspArgIlePhePhe------------ArgAlaAlaAlaGluGluGlyHisIleIle 330
Db       2497 GTGGAGGAAATGTTTTTTGGCCGTTCCGACAAGGTGTTTCTGCGAAGCAGGCGGCTTGTG 2556

Qy        331 ProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAsp---------------------- 343
Db       2557 CCAGTCGAAGCTCATCTCGTCGGAACACGGAGTGTAAACGATGAGAACAACAACGCCGAC 2616

Qy        344 ------CysGlyArgPheProSerAspHisTrpGlyLeu 354
Db       2617 GCACCCGCATATTTGTTCCCATCCGATCACTACGGCGTA 2655

RESULT 62
AP004494/c
LOCUS       AP004494               103129 bp    DNA     linear   PLN 22-JUL-2003
DEFINITION  Lotus corniculatus var. japonicus genomic DNA, chromosome 4,
            clone:LjT09I17, TM0026, complete sequence.
ACCESSION   AP004494
VERSION     AP004494.1  GI:17736861
KEYWORDS    HTG.
SOURCE      Lotus corniculatus var. japonicus (Lotus japonicus)
  ORGANISM  Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1
  AUTHORS   Sato,S., Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T. and Tabata,S.
  TITLE     Structural Analysis of a Lotus japonicus Genome. I. Sequence
            Features and Mapping of Fifty-six TAC clones which cover the 5.4 Mb
            Regions of the Genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 103129)
  AUTHORS   Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-DEC-2001) Yasukazu Nakamura, Kazusa DNA Research
            Institute, Department of Plant Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
            URL:http://www.kazusa.or.jp, Tel:81-438-52-3935,
            Fax:81-438-52-3934)
FEATURES             Location/Qualifiers
     source          1. .103129
                     /organism="Lotus corniculatus var. japonicus"
                     /mol_type="genomic DNA"
                     /variety="japonicus"
                     /db_xref="taxon:34305"
                     /chromosome="4"
                     /clone="LjT09I17"
                     /clone_lib="LjT library"
                     /note="TAC clone:TM0026
                     synonym: Lotus japonicus"
ORIGIN

Alignment Scores:
Pred. No.:                0.144          Length:          103129
Score:                    148.00         Matches:         53
Percent Similarity:       43.1%          Conservative:    35
Best Local Similarity:    26.0%          Mismatches:      76
Query Match:              7.8%           Indels:          40
DB:                       15             Gaps:            8

US-10-757-745-2 (1-362) x AP004494 (1-103129)

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 203
Db      32779 TTGCAGTTAAGCAAACTACCTGTGAAATCTTTCAGCAACAAGTCCTTTAGCAATTCTATA 32720

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsn---ValSerGlyAsnGluLeuCysLeu 222
Db      32719 ATGGGAAGGGAACTTTGCATTGCTGAGGTTGAAGCTGCAAATGGCAAGTCATTGGTTATT 32660

Qy        223 MetThrSerHisLeuGluSer----------------------ThrArgGlyHisAla 234
Db      32659 GCCACTAGCCATCTTGAGAGTCCCTGTCCAGGCCCTCCGAAATGGGATCAGATGTACAGC 32600

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
Db      32599 AAGGAACGAGTAGAGCAGGCGAACGAGGCTGTAAACCTGCTCAAGAAGCACCCT------ 32546

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
Db      32545 ---AATGTTGTTTTTGGAGGTGACATGAACTGGGATGACAAA---------CTGGATGGT 32498

Qy        275 GlyLeuPro------AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db      32497 CAATATCCTTCACAAGATGGGTGGATCGATGCCTGGTCTCAGCTA---AGGCCAGATCAA 32441

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
Db      32440 AGTGGGTGGACGTATGATACTAAGTCGAACCAGATGTTGTCAGGCAACCGTACTCTCCAA 32381

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
Db      32380 AAGCGATTAGATCGTTTCATTTGCCGTTTAAGT------------GATCTTAAGATAAGC 32333

Qy        333 SerLeuAspLeuLeuGlyLeuGlu---------------------------------- 340
Db      32332 AGCATTGACATGATTGGGATGGATGCAATACCTGGACTGTCATACATCAAAGAGAAGAAA 32273

Qy        341 ------------------LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeu 354
Db      32272 GTAAGAAAAGAGACAAAACAACTGGTACTCCCTGTTCTGCCTAGTGATCATTATGGTCTG 32213

Qy        355 LeuCysAsnLeu 358
Db      32212 CTTTTGACAATT 32201

RESULT 63
AC156673/c
LOCUS       AC156673               231985 bp    DNA     linear   HTG 01-JUL-2005
DEFINITION  Bos taurus clone CH240-59M13, *** SEQUENCING IN PROGRESS ***, 21
            unordered pieces.
ACCESSION   AC156673
VERSION     AC156673.2  GI:68266455
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 231985)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
```

```
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 231985)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA

REFERENCE   3  (bases 1 to 231985)
  AUTHORS   Cow Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA

COMMENT     On Jun 28, 2005 this sequence version replaced gi:58531393.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: FDJG
                Center clone name: CH240-59M13
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 224220 bases at least Q40
                Consensus quality: 226361 bases at least Q30
                Consensus quality: 227836 bases at least Q20
                Estimated insert size: 228962; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 21 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     3298: contig of 3298 bp in length
            *    3299     3348: gap of 50 bp
            *    3349     5753: contig of 2405 bp in length
            *    5754     6081: gap of 328 bp
            *    6082     7706: contig of 1625 bp in length
            *    7707     7806: gap of unknown length
            *    7807    24881: contig of 17075 bp in length
            *   24882    24931: gap of 50 bp
            *   24932    53019: contig of 28088 bp in length
            *   53020    53069: gap of 50 bp
            *   53070    56864: contig of 3795 bp in length
            *   56865    56914: gap of 50 bp
            *   56915    64212: contig of 7298 bp in length
            *   64213    64400: gap of 188 bp
            *   64401    74558: contig of 10158 bp in length
            *   74559    74608: gap of 50 bp
            *   74609    98289: contig of 23681 bp in length
            *   98290    98389: gap of unknown length
            *   98390   100834: contig of 2445 bp in length
            *  100835   100884: gap of 50 bp
            *  100885   132970: contig of 32086 bp in length
            *  132971   133113: gap of 143 bp
            *  133114   134462: contig of 1349 bp in length
            *  134463   134562: gap of unknown length
            *  134563   141930: contig of 7368 bp in length
            *  141931   141980: gap of 50 bp
            *  141981   186048: contig of 44068 bp in length
            *  186049   186098: gap of 50 bp
            *  186099   203328: contig of 17230 bp in length
            *  203329   203378: gap of 50 bp
            *  203379   205841: contig of 2463 bp in length
            *  205842   205941: gap of unknown length
            *  205942   207514: contig of 1573 bp in length
            *  207515   207564: gap of 50 bp
            *  207565   212129: contig of 4565 bp in length
            *  212130   212979: gap of 850 bp
            *  212980   214003: contig of 1024 bp in length
            *  214004   214103: gap of unknown length
            *  214104   215330: contig of 1227 bp in length
            *  215331   215430: gap of unknown length
            *  215431   231985: contig of 16555 bp in length.
FEATURES             Location/Qualifiers
     source          1. .231985
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9913"
                     /clone="CH240-59M13"
     gap             3299. .3348
                     /estimated_length=50
     gap             5754. .6081
                     /estimated_length=328
     gap             7707. .7806
                     /estimated_length=unknown
     gap             24882. .24931
                     /estimated_length=50
     gap             53020. .53069
                     /estimated_length=50
     gap             56865. .56914
                     /estimated_length=50
     gap             64213. .64400
                     /estimated_length=188
     gap             74559. .74608
                     /estimated_length=50
     gap             98290. .98389
```

```
                     /estimated_length=unknown
     gap             100835. .100884
                     /estimated_length=50
     gap             132971. .133113
                     /estimated_length=143
     gap             134463. .134562
                     /estimated_length=unknown
     gap             141931. .141980
                     /estimated_length=50
     gap             186049. .186098
                     /estimated_length=50
     gap             203329. .203378
                     /estimated_length=50
     gap             205842. .205941
                     /estimated_length=unknown
     gap             207515. .207564
                     /estimated_length=50
     gap             212130. .212979
                     /estimated_length=850
     gap             214004. .214103
                     /estimated_length=unknown
     gap             215331. .215430

Alignment Scores:
Pred. No.:                0.372          Length:          231985
Score:                    148.00         Matches:         28
Percent Similarity:       88.2%          Conservative:    2
Best Local Similarity:    82.4%          Mismatches:      4
Query Match:              7.8%           Indels:          0
DB:                       14             Gaps:            0

US-10-757-745-2 (1-362) x AC156673 (1-231985)

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348
              :::||||||   ||||||||||||||||||||||||||||||||||   ||||||||
Db       1492 GTTATTCCCTGAAGTTTGGATCTGCTAGGGCTGGAAAAACTGGACTGTAGTAGATTTCCC 1433

Qy        349 SerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||   ||||||   |||||||||:::||||||
Db       1432 AGTGATCACTGGGATCTTCTGTGAAACTTAGATGTAATACTG 1391

RESULT 64
MMU70139
LOCUS       MMU70139                 9737 bp    mRNA    linear   ROD 27-APR-1999
DEFINITION  Mus musculus probable nocturnin protein mRNA, partial cds.
ACCESSION   U70139
VERSION     U70139.1  GI:2251233
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 9737)
  AUTHORS   Puech,A., Dupressoir,A., Loireau,M.P., Mattei,M.G. and Heidmann,T.
  TITLE     Characterization of two age-induced intracisternal
            A-particle-related transcripts in the mouse liver. Transcriptional
            read-through into an open reading frame with similarities to the
            yeast ccr4 transcription factor
  JOURNAL   J. Biol. Chem. 272 (9), 5995-6003 (1997)
   PUBMED   9038221
REFERENCE   2  (bases 1 to 9737)
  AUTHORS   Puech,A. and Dupressoir,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-SEP-1996) Cell Biology, Albert Einstein College of
            Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA
FEATURES             Location/Qualifiers
     source          1. .9737
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="DBA2/J"
                     /db_xref="taxon:10090"
                     /chromosome="III"
                     /map="III B-D"
                     /tissue_type="liver"
                     /dev_stage="18-month-old"
                     /note="intracisternal A particle chimeric transcript"
     CDS             <7661. .8632
                     /note="probable ortholog of Xenopus laevis nocturnin gene;
                     similar to yeast transcription factor CCR4;
                     transcriptional readthrough occurs with transcription
                     being initiated at the IAP and continues into an exon"
                     /codon_start=1
                     /product="probable nocturnin protein"
                     /protein_id="AAB62717.1"
                     /db_xref="GI:2251234"
                     /translation="GTSGGMQGCSAHSATPLPADFVDLRTDCSSSHSPIRVMQWNILA
                     QALGEGKDNFVQCPVEALKWEERKCLILEEILAYQPDILCLQEVDHYFDTFQPLLSRL
                     GYQGTFFPKPWSPCLDVEHNNGPDGCALFFLQNRFKLISSTNIRLTAMTLKTNQVAIA
                     QTLECKESGRQFCIAVTHLKARTGWERFRSAQGCDLLQNLQNITQGAKIPLIVCGDFN
                     AEPTEEVYKHFASSSLNLNSAYKLLSPDGQSEPPYTTWKIRTSGECRHTLDYIWYSRH
                     ALSVTSALDLLTEEQIGPNRLPSFHYPSDHLSLVCDFSFNEEPHELF"
ORIGIN

Alignment Scores:
Pred. No.:                0.0111         Length:          9737
Score:                    147.00         Matches:         86
Percent Similarity:       33.7%          Conservative:    53
Best Local Similarity:    20.8%          Mismatches:      161
Query Match:              7.7%           Indels:          113
DB:                       9              Gaps:            16

US-10-757-745-2 (1-362) x MMU70139 (1-9737)

Qy         35 ValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp--------- 51
              :::|||   |||      ||||||           :::   |||   |||
Db       7429 ATGGCGTGATGCCCACAGGCAGTGGGGCCCTGGCTCATACCGGAGAGGGACAGTGAAGGC 7488

Qy         52 ----------------------------------------------TrpGluMetGlu 55
                                                           ||||||
Db       7489 CCTGACTGACAGCTGGCTATGGGACTGTTTCTCCTGCAGTGAGTTCCATGGGAA------ 7542

Qy         56 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 75
                                        ||||||:::   :::   |||      ||||||
Db       7543 --------------------ACGGCACCAGTCGACTCTACAGTGCTCTCGCCAAGACCG 7581

Qy         76 GluThr------------IleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91
              |||            :::|||         |||:::      ||||||
Db       7582 TCAACAGCAGTGCAGCTGCTCAGCACCCGGAGTACTTGGTGTCAACTGACCCCGAACATC 7641

Qy         92 ThrThrAspSer------------------------------------------------ 95
              :::  |||
Db       7642 TGGAGCCCATCGATCCTAAGGAACTTCTGGAGGAATGCAGGGCTGTTCTGCACACTCGGC 7701

Qy         96 ------ThrThrSerLysIleSerProSerGluAspThrGln-GlnGluAsnGlySerMe 113
                    ||||||                ||||||                :::   |||
Db       7702 CACCCCGCTACCAGCGGATTTTGTGGACCTGAGGACAGATTGCTCCAGCAGCCACTCTCC 7761

Qy        113 tPheSerLeuIleThrTrpAsnIle-----------------AspGlyLeuAsp----- 125
              ::::::   |||||||||                   :::|||   |||
Db       7762 CATTCGCGTCATGCAGTGGAACATCCTCGCCCAAGCTCTCGGAGAAGGCAAAGACAACTT 7821

Qy        126 ------------------LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
                               ||||||      :::
Db       7822 TGTGCAGTGCCCCGTGGAAGCGCTCAAATGGGAAGAGAGGAAGTGCCTGATCCTGGAGGA 7881

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
              :::      |||   ||||||::::::   ||||||||||||         |||:::
Db       7882 GATCCTGGCTTACCAGCCAGACATACTGTGCCTCCAGGAAGTG---GACCACTACTTTGA 7938

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
              :::      |||                  |||::::::     :::||
Db       7939 CACCTTCCAGCCACTCCTCAGTAGA-----------CTGGGCTACCAAGGCACGTTTTT 7986

Qy        179 e----------------------------------------------------ThrAl 181
              |                                                      ||
```

```
Db        7987 CCCCAAGCCCTGGTCACCATGTCTAGATGTGGAACACAACAACGGTCCAGATGGCTGTGC 8046

Qy         181 aIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSe 201

Db        8047 CTTATTTTTTCTCCAAAACCGATTCAAGCTTATCAGCAGCACCAATATTAGGCTGACAGC 8106

Qy         201 rThrLysMetMetArgAsnLeuLeuCysVal-----------HisValAsnValSerGl 217

Db        8107 CATGACCCTAAAAACCAATCAGGTGGCCATCGCACAGACCCTGGAGTGCAAGGAGTCTGG 8166

Qy         217 yAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluAr 237

Db        8167 GCGACAGTTCTGCATTGCTGTCACCCACTTAAAAGCCCGCACTGGCTGGGAGCGGTTCCG 8226

Qy         237 gMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVa 257

Db        8227 GTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTGCAGAACATCACCCAGGGAGCAAAGAT 8286

Qy         257 l------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu------ValThrArgCy 273

Db        8287 CCCCCTGATCGTCTGCGGGGACTTCAACGCAGAGCCAACCGAAGAGGTCTACAAACACTT 8346

Qy         273 sGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------GlyLysPr 290

Db        8347 TGCGTCCTCCAGCCTCAACCTCAACAGCGCCTACAAGCTGCTGAGTCCCGATGGACAGTC 8406

Qy         290 oLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAl 310

Db        8407 GGAGCCTCCCTACACAACCTGG-----------------AAGATCCGGACCTCAGGCGA 8448

Qy         310 aCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIl 330

Db        8449 GTGTCGCCACACGCTGGACTATATCTGGTACTCCAGACATGCTCTGAGT---------GT 8499

Qy         330 eProArgSerLeuAspLeuLeu---------------GlyLeuGluLysLeuAspCysGl 345

Db        8500 CACGTCCGCCCTGGATCTACTCACTGAAGAACAGATTGGGCCCAACCGGCTCCCATCCTT 8559

Qy         345 yArgPheProSerAspHisTrpGlyLeuLeuCysAsn 357

Db        8560 CCATTACCCCTCGGACCACCTGTCCCTGGTGTGTGAC 8596
```

RESULT 65
AC159426
LOCUS AC159426 156973 bp DNA linear INV 07-APR-2005
DEFINITION Trypanosoma brucei chromosome 7 clone RPCI93-27M11, complete sequence.
ACCESSION AC159426
VERSION AC159426.1 GI:62175798
KEYWORDS HTG.
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 156973)
AUTHORS Ghedin,E., Blandin,G., Bartholomeu,D., Caler,E., Haas,B., Hannick,L., Shallom,J., Hou,L., Djikeng,A., Feldblyum,T., Hostetler,J., Johnson,J., Jones,K., Koo,H.L., Larkin,C., Pai,G., Peterson,J., Khalak,H.G., Salzberg,S., Simpson,A.J., Tallon,L., Van Aken,S., Wanless,D., White,O., Wortman,J., Fraser,C.M. and El-Sayed,N.M.A.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 156973)
AUTHORS Haas,B., Blandin,G. and El-Sayed,N.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2005) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, nelsayed@tigr.org
REFERENCE 3 (bases 1 to 156973)
AUTHORS Haas,B., Blandin,G. and El-Sayed,N.
TITLE Direct Submission
JOURNAL Submitted (07-APR-2005) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, nelsayed@tigr.org
COMMENT BAC clone RPCI93-27M11 from Trypanosoma brucei chromosome 7.

Putative protein coding regions (>30 codons) were predicted by combining the output of the gene finding algorithm GLIMMER with similarity data and manual annotation of open reading frames. Further refinement of gene models (additions, deletions and alterations to boundary coordinates) were done manually based on detailed comparisons with the T. cruzi and L. major genomes. Genes on BACs were assigned a systematic name based on the chromosome and the BAC from which they originated (e.g. Tb03.27F10.410). Gene products were assigned a putative function when they shared significant similarity with experimentally characterized gene products or when they contained functionally known protein domains. Gene products were labelled either 'hypothetical protein', or 'hypothetical protein, conserved' in the case of products showing significant similarity with proteins or domains of unknown function from other organisms. Gene products of unknown function with predicted orthologs in Trypanosoma cruzi and Leishmania major genomes were automatically classified as conserved proteins. Short predicted coding sequences (less than 150 codons) with no additional evidence of coding potential were labelled 'hypothetical protein, unlikely'.

FEATURES Location/Qualifiers
source 1. .156973
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/isolate="GUTat10.1"
/db_xref="taxon:5691"
/chromosome="7"
/clone="RPCI93-27M11"
misc_feature 1. .20096
/note="annotation for this sequence region can be found on overlapping clone 29K4."
misc_feature 19986. .20025
/note="AT_rich"
gene <20389. .>23073
/locus_tag="Tb927.7.1070"
/note="synonym: Tb07.27M11.120; RNA-editing complex protein MP100"
mRNA <20389. .>23073
/locus_tag="Tb927.7.1070"
/product="RNA-editing complex protein MP100"
CDS 20389. .23073
/locus_tag="Tb927.7.1070"
/note="99% similar to GB:AAO63565.1: RNA editing complex protein MP100; putative exonuclease {Trypanosoma brucei}(PMID:12649499) Protein identified by mass spectrometry; go_component: mRNA editing complex [goid 0045293]; go_process: RNA editing [goid 0016547]"
/codon_start=1
/product="RNA-editing complex protein MP100"
/protein_id="AAX69956.1"
/db_xref="GI:62175829"
/translation="MALAQSCRHFAATTYRQSKITDRFQLLITKKEIECDYVAVDATA
LVATGLRLAKNVTTEQRRHKEVARHVVQSIQQLLKKVRCKKSLLIAMDGAENLLKADR
TRGSSLTRKVESRLMRLPGTPLMQAVEERIVRMMPERQILPGEVVFAGTCVQGCVEQK
MSAWALDLASRDTFNGSSDSLNLIGASELYLNVLALSPFYNVSSVVQNNADLRHIRLQ
DILEWLELDKKAKEGESVTIAKMRTDILFLFIIANGASATELNPIPAVGFHEIVDRYM
KVLAENGATPPSQSATASAPTASANVPAAPAFLFEDSPGNTLQLNLRMLCRIMQLVAR
KEAPARVDGLSEAYLEHALQTHAMLCTGQAPQCTYLPAAGTGQLGSTAPSAVQLAGHL
AALCASTAQRNARCTSLQPSLVTPEKGSCEEDSSATKLPPTPPTSAVAPVEDKPLTAA
EYTILCQSLPAAVEGLIHQYVGVAPKPDVGKMITTATTEEAYRLVREVLSYANPLRPH
KCLCLSPSYCWLQNEKTQLWRFEYVDIGVKSHELGTRRQLNAMKGVTLEVNMSRDGPS
HFDDPLSGTWEPIANFPCGAECEYDKHTEECTSSPSAPQTAKEVEGGPVKASLKLLTWN
VMFDRYSGKPTPLGMPGIDWCSPKRYPVIAKIIEEEDADVVGMQEVEPVFWEFLSKRP
LIRQRYYFSCGHMSPAIAPWGVLMLIHRRRLPVQSINYLNVPAWTNHVSLMPVVGLKM
VHGTVHIAAAHLLAPYTKSHENARTSQDTALRHHMTKTLGGGDVVTMGDFNDWPTNEF
IMPHESQYVDCWPVLHPNNPGKTMDETNTFCKLKVEEMFFGRSDKVFLRSRRLVPVEA
HLVGTRSVNDENNNADAPAYLFPSDHYGVSITFSLKLP"
gene <23560. .>24642
/locus_tag="Tb927.7.1080"
/note="synonym: Tb07.27M11.140; hypothetical protein, conserved"
mRNA <23560. .>24642

```
                     /locus_tag="Tb927.7.1080"
                     /product="hypothetical protein, conserved"
     CDS             23560. .24642
                     /locus_tag="Tb927.7.1080"
                     /codon_start=1
                     /product="hypothetical protein, conserved"
                     /protein_id="AAX69957.1"
                     /db_xref="GI:62175830"
                     /translation="MMQRCGTSLSRLCFRRLLRTPLLVYSIPPIRDVPSGIAHCPLSC
                     SMRMVTSSNDDEFVFDPTLSIQKDAAIHTAKKSFETIVLEYVPAHAPEEARQKVKSYL
                     TQHPIDILITQPKVQITHLEDAESGAETKVSLSPCDLPEALQQARERGMNLVQMGARG
                     DVAYCRIRRESARILSLIHTELEALREQEEKQQGKGRGGVQAAAKMGELIDHTFRDAV
                     DAHFVGWRSKKIVEDIRRRHPVKLTIKEFQSPECAIGKLREMCQAMQHYAQEKVIYHH
                     FTSIVANDREASITFAPALPMAKSDSWKHIKYPGEKEWTNALRRMEDACRKSGRYGTY
                     AKSNKLKLRSLGQTSYRVDKYGRKMD"
     repeat_region   24695. .24718
                     /note="microsatellite"
                     /rpt_type=tandem
                     /rpt_unit="atggtg"
     gene            <25900. .>30525
                     /locus_tag="Tb927.7.1090"
                     /note="synonym: Tb07.27M11.170; hypothetical protein,
                     conserved"
     mRNA            <25900. .>30525
                     /locus_tag="Tb927.7.1090"
                     /product="hypothetical protein, conserved"
     CDS             25900. .30525
                     /locus_tag="Tb927.7.1090"
                     /codon_start=1
                     /product="hypothetical protein, conserved"
                     /protein_id="AAX69958.1"
                     /db_xref="GI:62175831"
                     /translation="MVGNGNDTVLTDGDDDECIAEQEDKVTKALQARLRCGMQRFATS
                     ATNQGGLVKPSGPPSSVAAAAQRSTTTKRDNLTNSRRQEIFVEHVKHFMRYTVIADKA
                     RKLVSLEERPKEDPDVGNDRVELLRIHVSLPCVWRYILHEWERVEEASKKSSSLAASR
                     DASVSRLVSLFMKCLNSLDGIHALLFIACAACRALDLTGGVTSVLLHLLLETGRVEED
                     RLLKSDSGELVYKAIEEIDKEMHSRFGGSDCCYFQTHSYLLLLIWYTQRDVSDNGGPL
                     KMSHQVVLSQAFVKATSPYQTARTRRGSKAAASVAPSEESGKRKPRGRRCAGSSEDGA
                     ALSGFDRHATLVTEVQEHETRWAEDSSKFLFSLYALQNASGSASYGMKKLALREVMGA
                     VFLFHRVLSESGVNDFLLESKESRIRELHEWSLSKGFVTPPLTHGGESSNNGTTVNSF
                     ASTDVGEDGLNVMSMAGRTYRECTDVERSFPRVANARAMLQHVTQIAAANYRGNTRAV
                     GEGFPHEYIRARLHPHQVESLAYMLERETRGLAQYVELRGFFEATSGLPSEVADRSAG
                     RGSREFSLFYSSMSSEWFIVRSDSLSARFARRALHCGFLCDEMGLGKTLVVLSLCATA
                     KKRAIELEETEVPTVTTLGTVKGSNVGAVNKAFAESDAGTAGAGEDEPIYLWRHIRVR
                     KTGAPWSDMLEVLYGSLPQPVYLPTPKPLPKTTLVVVPLSLLIQWVSEIDRFYPLARY
                     ILFHGPSRQRYTYEDFCSADFVITTYETVGAHLRVETDAAFRLLTMTLPYMEHKDASL
                     AQWDTFCRGCNERRARAASASQGVNDCDGSADVFDSILSLFESARRQVVLEAEGAPPG
                     PTLDFVSQWDKTFREFLRSDVLYLTDFVFERIILDESQKCSKTSRLQYLQARRRWVVS
                     GTPINNNDVNSLQPLFKFLWAELSPSYSVSPLSAFYCSSYDKQRHMISALNAFRSSRK
                     MARYAINSFYVGRHSKSSPSDTRHVLTQPAIRKPLCMCRLCNPFTKQTVYPKSPVLTL
                     KDALYRRVTSAASEEETEGFISDRGGLFRSWRFSVPLTHLMVRHEKTEELQRALQLSP
                     VTEAHVEVPLLDGERFLYDYAGDVVARGVALLQRQGLVGNCSAKLLSWVCLLGRMALH
                     PSTAFADVRNSDQALYVQSHQKNEQVASNFLSSVVSVTPNDALAYARKLIANRLNTTG
                     TRKGRGETNVLVPPETTVETLRKLTLDPPEVPECAICLDTMLHPTLLSCFHLFCKECL
                     FATIQVARPTLSNETTARCPHCRNPKSMREKMMVIQVENTQVSTLGALQDEGMAPPNG
                     ENVSHEELKRRVDAVGRGSRIPRLIQLLEDIWDEAPDDNVLVFSKVPAILQMGEAALK
                     EAGYQNVHVIDNHVTLVQRKKIFAQLHANAATTASATGSASGSTPSGNNRGHVLFLSS
                     RVACSGLNLVFANRLIFMEPNLNPAQHQQAVGRIDRCGQLKRTYVYVMYAPRTIEERI
                     MNRTSALLQEASQGTFPDVATGRGAPAGNTSSMGQSELITLLQAQQA"
     misc_feature    30633. .30663
                     /note="AT_rich"
     misc_feature    30860. .30882
                     /note="AT_rich"
     repeat_region   31003. .31039
                     /note="microsatellite"
                     /rpt_type=tandem
                     /rpt_unit="ta"
     misc_feature    31588. .31651
                     /note="A-rich"
     misc_feature    32647. .32674
                     /note="AT_rich"
     repeat_region   33412. .33434
                     /note="microsatellite"
                     /rpt_type=tandem
                     /rpt_unit="ta"
     gene            <33543. .>35756
                     /locus_tag="Tb927.7.1100"
                     /note="synonym: Tb07.27M11.240; hypothetical protein,
                     conserved"
     mRNA            <33543. .>35756
                     /locus_tag="Tb927.7.1100"
                     /product="hypothetical protein, conserved"
     CDS             33543. .35756
                     /locus_tag="Tb927.7.1100"
                     /codon_start=1
                     /product="hypothetical protein, conserved"
                     /protein_id="AAX69959.1"
                     /db_xref="GI:62175832"
                     /translation="MNTTGEPAKAAAPPAVRRRGRGTFNRPLRAEEPTAPNATASHAG
                     DNVSVITTVTTAVLPASTGPSTRVPEEDLAAEYGGQKRDIPLNIRENFGVHEEEQVVD
                     TVRQLLFRGEKRRRYCVIVGNLQNPDRPITLKLFGHKGLKPFKLQNKLIRLVNISVSR
                     PRLIKQTDASKEEQEVWSGGGSDDEVIADAQADTVTKTENDFDNQGSCDHGQQQKWRK
                     EEQEGTPKRDYESEIQRAHERRNPGGDGGVKDGEESQPTIRMSGSEVEKNGSGDMDVK
                     ANERPPSQSGEGRDQGWEEKTMDTETRDVDDYANEDVVLPLTVEIRDSTATNIGQPSS
                     IANSPVLTAARFDTLYGYFFREFQHLLQLEDVRMHQVSINLSLGAVYCCLAREGSYIT
                     PLKYATFSEFVQRMNEESMQLYFLKDAPICVSRAVDRECGALQSEAAFSLVKIYFFSN
                     ERQSRTVARAMWDAHENRFVLLDMENMGVTFTWTVFSVDETDIQNNRKKHQGEGVSND
                     SGIPDGDSATETDKPMRKKTVAFPFELELRAYRRWKQHTEHPVAPIAEAILDKLSLAE

Alignment Scores:
Pred. No.:              0.358          Length:          156973
Score:                  146.00         Matches:         66
Percent Similarity:     38.7%          Conservative:    55
Best Local Similarity:  21.1%          Mismatches:      128
Query Match:            7.7%           Indels:          64
DB:                     2              Gaps:            14

US-10-757-745-2 (1-362) x AC159426 (1-156973)

Qy         89 AsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGln 108
               ::::::   |||   :::   |||:::|||       :::|||           :::
Db      22156 GACAAACATACTGAAGAATGCACATCTTCTCCTTCAGCTCCACAAACTGCAAAAGAGGTT 22215

Qy        109 GluAsnGly---------SerMetPheSerLeuIleThrTrpAsnIle------------ 121
               |||   |||         :::         |||:::||||||||:::
Db      22216 GAAGGAGGACCAGTTAAGGCATCTTTAAAATTGCTGACATGGAATGTCATGTTTGATCGT 22275

Qy        122 ----------------------------AspGlyLeuAspLeuAsnAsnLeuSerGlu 131
                                          |||:::|||       :::      :::
Db      22276 TACAGCGGCAAACCAACTCCTTTAGGTATGCCTGGTATAGATTGGTGCTCC---CCCAAG 22332

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
               |||         :::         :::                |||||:::   :::|||
Db      22333 AGGTATCCCGTCATTGCCAAGATCATAGAGGAGGAGGATGCGGATGTTGTTGGCATGCAA 22392

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArg---------SerSerAsnTyr 168
               ||||||   |||   ::::::   :::|||   ||||||                  |||
Db      22393 GAGGTGGAGCCTGTATTTTGGGAGTTCCTTTCCAAGCGGCCACTTATACGCCAAAGGTAC 22452

Qy        169 GluIleIleThrGlyHisGluGlu-----------GlyTyrPheThrAlaIleMetLeu 184
                         ||||||                 :::          :::::::::
Db      22453 TACTTTTCGTGCGGTCATATGAGTCCAGCTATTGCCCCGTGGGGGGTGCTGATGTTGATA 22512

Qy        185 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer---ThrLys 203
               :::   |||:::   ::::::|||         :::      |||:::   |||
Db      22513 CACCGTCGGCGTCTTCCAGTACAGAGCATCAACTATCTTAACGTCCCTGCGTGGACTAAT 22572

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMet 223
               :::      :::       |||   :::   :::             :::   :::
Db      22573 CATGTTTCTCTCATGCCGGTTGTGGGTCTGAAGATGGTGCACGGTACTGTACACATAGCT 22632

Qy        224 ThrSerHisLeuGluSer------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 241
               :::||||||   :::       |||:::   |||         |||   :::|||
Db      22633 GCCGCACACTTACTCGCGCCATACACGAAGTCCCACGAGAACGCTCGTACTTCCCAGGAT 22692

Qy        242 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 261
```

```
Db      22693 ACCGCGTTACGCCATCACATGACGAAGACGTTGGGTGGCGGGGATGTGGTAACTATGGGA 22752

Qy        262 AspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleVal 281

Db      22753 GATTTCAAC------------------------------GACTGGCCTACAAAT------ 22776

Qy        282 AspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr------TrpAspThr 299

Db      22777 ---------GAATTTATT---ATGCCACATGAATCACAGTACGTGGACTGTTGGCCCGTT 22824

Qy        300 GlnMetAsnSerAsnLeuGlyIleThr--------------AlaAlaCysLysLeuArg 314

Db      22825 TTGCACCCCAACAACCCTGGAAAGACTATGGATGAAACAAACACCTTTTGTAAACTGAAA 22884

Qy        315 PheAspArgIlePhePhe------------ArgAlaAlaAlaGluGluGlyHisIleIle 330

Db      22885 GTGGAGGAAATGTTTTTTGGCCGTTCCGACAAGGTGTTTCTGCGAAGCAGGCGGCTTGTG 22944

Qy        331 ProArgSerLeuAspLeuGluGlyLeuGluLysLeuAsp---------------------- 343

Db      22945 CCAGTCGAAGCTCATCTCGTCGGAACACGGAGTGTAAACGATGAGAACAACAACGCCGAC 23004

Qy        344 ------CysGlyArgPheProSerAspHisTrpGlyLeu 354

Db      23005 GCACCCGCATATTTGTTCCCATCCGATCACTACGGCGTA 23043
```

```
RESULT 66
BD112459
LOCUS       BD112459                379 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  EST and encoded human protein.
ACCESSION   BD112459
VERSION     BD112459.1  GI:23207277
KEYWORDS    JP 2002010789-A/4536.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
  TITLE     EST and encoded human protein
  JOURNAL   Patent: JP 2002010789-A 4536 15-JAN-2002;
            GENSET CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002010789-A/4536
            PD   15-JAN-2002
            PF   07-AUG-2000 JP 2000280989
            PR   05-AUG-1999 US    60/147499
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE  PI
             GIORDANO
            PC   C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
             C12N1/21,
            PC   C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
            C12N15/00
            CC   EST and encoded human protein
            FH   Key             Location/Qualifiers
            FT   source          1. .379
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                0.000374        Length:          379
Score:                    145.00          Matches:         28
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    96.6%           Mismatches:      0
Query Match:              7.6%            Indels:          0
DB:                       6               Gaps:            0
```

US-10-757-745-2 (1-362) x BD112459 (1-379)

```
Qy          55 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 74

Db          19 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 78

Qy          75 ProGluThrIleSerGluProLysThr 83

Db          79 CCTGAAACCATCTCTGAGCCCAAGACC 105
```

```
RESULT 67
AR416906
LOCUS       AR416906                379 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 8403 from patent US 6639063.
ACCESSION   AR416906
VERSION     AR416906.1  GI:40172016
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
  TITLE     EST's and encoded human proteins
  JOURNAL   Patent: US 6639063-A 8403 28-OCT-2003;
            Genset S.A.;;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                0.000374        Length:          379
Score:                    145.00          Matches:         28
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    96.6%           Mismatches:      0
Query Match:              7.6%            Indels:          0
DB:                       6               Gaps:            0
```

US-10-757-745-2 (1-362) x AR416906 (1-379)

```
Qy          55 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 74

Db          19 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 78

Qy          75 ProGluThrIleSerGluProLysThr 83

Db          79 CCTGAAACCATCTCTGAGCCCAAGACC 105
```

```
RESULT 68
AX977600
LOCUS       AX977600                379 bp    DNA     linear   PAT 15-JAN-2004
DEFINITION  Sequence 8403 from Patent EP1104808.
ACCESSION   AX977600
VERSION     AX977600.1  GI:40983740
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
  TITLE     ESTs and encoded human proteins
  JOURNAL   Patent: EP 1104808-A 8403 06-JUN-2001;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

```
ORIGIN

Alignment Scores:
Pred. No.:                0.000374        Length:          379
Score:                    145.00          Matches:         28
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    96.6%           Mismatches:      0
Query Match:              7.6%            Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2 (1-362) x AX977600 (1-379)

Qy          55 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 74
               :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          19 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 78

Qy          75 ProGluThrIleSerGluProLysThr 83
               |||||||||||||||||||||||||||
Db          79 CCTGAAACCATCTCTGAGCCCAAGACC 105

RESULT 69
AF199491
LOCUS       AF199491                2561 bp    mRNA    linear   ROD 18-JUL-2002
DEFINITION  Mus musculus NOCTURNIN (Nocturnin) mRNA, complete cds.
ACCESSION   AF199491
VERSION     AF199491.1  GI:9885277
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2561)
  AUTHORS   Wang,Y., Osterbur,D.L., Megaw,P.L., Tosini,G., Fukuhara,C.,
            Green,C.B. and Besharse,J.C.
  TITLE     Rhythmic expression of Nocturnin mRNA in multiple tissues of the
            mouse
  JOURNAL   BMC Dev. Biol. 1 (1), 9 (2001)
   PUBMED   11394964
REFERENCE   2  (sites)
  AUTHORS   Wang,Y., Osterbur,D.L., Green,C.B. and Besharse,J.C.
  TITLE     Mammalian Homologs of Xenopus Nocturnin: Conservation of Structure
            and Circadian Regulation
  JOURNAL   Unpublished
  REMARK    Submitted
REFERENCE   3  (bases 1 to 2561)
  AUTHORS   Wang,Y., Osterbur,D.L., Green,C.B. and Besharse,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-1999) Department of Cell Biology, Neurobiology
            and Anatomy, Medical College of Wisconsin, 8701 Watertown Plank
            Road, Milwaukee, WI 53226, USA
FEATURES             Location/Qualifiers
     source          1. .2561
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="BALB/c"
                     /db_xref="ATCC:77448"
                     /db_xref="taxon:10090"
                     /clone="FN-2"
                     /tissue_type="retina"
                     /clone_lib="ATCC77448"
     gene            1. .2561
                     /gene="Nocturnin"
                     /note="circadian rhythmicity"
     CDS             151. .1440
                     /gene="Nocturnin"
                     /note="expressed in retina and liver; similar to CCR4"
                     /codon_start=1
                     /product="NOCTURNIN"
                     /protein_id="AAG01384.1"
                     /db_xref="GI:9885278"
                     /translation="MYQSPRRLCSALLLRDAPGLRRTLVPGPRRTLAPPVLGSRPKSP
                     QLQAAAASGAARSRPRTVSSMGNGTSRLYSALAKTVNSSAAAQHPEYLVSTDPEHLEP
                     IDPKELLEECRAVLHTRPPRYQRDFVDLRTDCSSSHSPIRVMQWNILAQALGEGKDNF
                     VQCPVEALKWEERKCLILEEILAYQPDILCLQEVDHYFDTFQPLLSRLGYQGTFFPKP
                     WSPCLDVEHNNGPDGCALFFLQNRFKLISSTNIRLTAMTLKTNQVAIAQTLECKESGR
                     QFCIAVTHLKARTGWERFRSAQGCDLLQNLQNITQGAKIPLIVCGDFNAEPTEEVYKH
                     FASSSLNLNSAYKLLSPDGQSEPPYTTWKIRTSGECRHTLDYIWYSRHALSVTSALDL
                     LTEEQIGPNRLPSFHYPSDHLSLVCDFSFNEEPHELF"
ORIGIN

Alignment Scores:
Pred. No.:                0.01            Length:          2561
Score:                    140.00          Matches:         89
Percent Similarity:       35.0%           Conservative:    55
Best Local Similarity:    21.7%           Mismatches:      163
Query Match:              7.4%            Indels:          104
DB:                       9               Gaps:            19

US-10-757-745-2 (1-362) x AF199491 (1-2561)

Qy          12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArg------------------ 25
               :::||||||          |||          :::   |||
Db         289 CAGGCGGCGGCGGCCTCGGGCGCCGCGAGGTCGCGTCCCCGAACAGTGAGTTCCATGGGA 348

Qy          26 ------------ArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
                        ||||||          :::|||   |||      ||||||
Db         349 AACGGCACCAGTCGACTCTACAGTGCTCTCGCCAAGACCGTCAACAGCAGTGCAGCTGCT 408

Qy          42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
                       :::|||          |||  |||                    :::
Db         409 CAGCACCCGGAGTACTTGGTGTCAACTGACCCCGAA--------------------CAT 447

Qy          62 PheGluPro------ProValGlu--------------GluSerAlaLeuGluArgArg 74
               ||||||      |||  |||                 :::   |||     |||
Db         448 CTGGAGCCCATCGATCCTAAGGAACTTCTGGAGGAATGCAGGGCTGTTCTGCACACTCGG 507

Qy          75 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 94
               |||                 :::   :::|||||||||                ||||||
Db         508 CCACCCCGCTACCAG------CGGGATTTTGTGGACCTG------------AGGACAGAT 549

Qy          95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114
               ::::::|||                                              |||
Db         550 TGCTCCAGCAGCCAC------------------------------------TCTCCCATT 573

Qy         115 SerLeuIleThrTrpAsnIle----------------AspGlyLeuAsp--------- 125
               ::::::   ||||||||              :::|||   |||
Db         574 CGCGTCATGCAGTGGAACATCCTCGCCCAAGCTCTCGGAGAAGGCAAAGACAACTTTGTG 633

Qy         126 --------------LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
                           ||||||          :::         :::
Db         634 CAGTGCCCCGTGGAAGCGCTCAAATGGGAAGAGAGGAAGTGCCTGATCCTGGAGGAGATC 693

Qy         141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
               |||   |||||::::::   ||||||||||          |||:::
Db         694 CTGGCTTACCAGCCAGACATACTGTGCCTCCAGGAAGTG---GACCACTACTTTGACACC 750

Qy         161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe--- 179
               :::          |||                |||::::::      ::|||
Db         751 TTCCAGCCACTCCTCAGTAGA-----------CTGGGCTACCAAGGCACGTTTTTCCCC 798

Qy         180 ------------------------------------------------ThrAlaIle 182
                                                               |||:::
Db         799 AAGCCCTGGTCACCATGTCTAGATGTGGAACACAACAACGGTCCAGATGGCTGTGCCTTA 858

Qy         183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 202
                      :::::::|||  ||||||  |||      |||         :::
Db         859 TTTTTTCTCCAAAACCGATTCAAGCTTATCAGCAGCACCAATATTAGGCTGACAGCCATG 918

Qy         203 LysMetMetArgAsnLeuLeuCysVal-----------HisValAsnValSerGlyAsn 218
               :::       |||  :::  :::                          ||||||
Db         919 ACCCTAAAAACCAATCAGGTGGCCATCGCACAGACCCTGGAGTGCAAGGAGTCTGGGCGA 978

Qy         219 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
               :::   |||:::      :::||||||::::::      |||              |||
Db         979 CAGTTCTGCATTGCTGTCACCCACTTAAAAGCCCGCACTGGCTGGGAGCGGTTCCGGTCA 1038
```

```
Qy        239 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal--- 257
Db       1039 GCTCAGGGCTGTGACCTCCTCCAGAACCTGCAGAACATCACCCAGGGAGCAAAGATCCCC 1098

Qy        258 ---IlePheAlaGlyAspThrAsnLeuArgAspArgGlu------ValThrArgCysGly 274
Db       1099 CTGATCGTCTGCGGGGACTTCAACGCAGAGCCAACCGAAGAGGTCTACAAACACTTTGCG 1158

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------GlyLysProLys 291
Db       1159 TCCTCCAGCCTCAACCTCAACAGCGCCTACAAGCTGCTGAGTCCCGATGGACAGTCGGAG 1218

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
Db       1219 CCTCCCTACACAACCTGG----------------AAGATCCGGACCTCAGGCGAGTGT 1260

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
Db       1261 CGCCACACGCTGGACTATATCTGGTACTCCAGACATGCTCTGAGT---------GTCACG 1311

Qy        332 ArgSerLeuAspLeuLeu--------------GlyLeuGluLysLeuAspCysGlyArg 346
Db       1312 TCCGCCCTGGATCTACTCACTGAAGAACAGATTGGGCCCAACCGGCTCCCATCCTTCCAT 1371

Qy        347 PheProSerAspHisTrpGlyLeuLeuCysAsn 357
Db       1372 TACCCCTCGGACCACCTGTCCCTGGTGTGTGAC 1404
```

RESULT 70
AF183960
LOCUS AF183960 3037 bp mRNA linear ROD 19-OCT-1999
DEFINITION Mus musculus strain BALB/c carbon catabolite repression 4 protein homolog (Ccr4) mRNA, complete cds.
ACCESSION AF183960
VERSION AF183960.1 GI:5924313
KEYWORDS .
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3037)
AUTHORS Dupressoir,A., Barbot,W., Loireau,M.P. and Heidmann,T.
TITLE Characterization of a mammalian gene related to the yeast CCR4 general transcription factor and revealed by transposon insertion
JOURNAL J. Biol. Chem. 274 (43), 31068-31075 (1999)
PUBMED 10521507
REFERENCE 2 (bases 1 to 3037)
AUTHORS Dupressoir,A., Barbot,W., Loireau,M.P. and Heidmann,T.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1999) CNRS UMR 1573, Institut Gustave Roussy, 39 rue Camille Desmoulins, Villejuif Cedex 94805, France
FEATURES Location/Qualifiers
source 1. .3037
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/chromosome="3"
/map="B-D region"
/tissue_type="brain"
gene 1. .3037
/gene="Ccr4"
CDS 644. .1933
/gene="Ccr4"
/note="conserved motif, PDILCLQEV, involved in Mg2+ or Mn2+ binding in the AP endonuclease family"
/codon_start=1
/product="carbon catabolite repression 4 protein homolog"
/protein_id="AAD56547.1"
/db_xref="GI:5924314"
/translation="MYQSPRRLCSALLLRDAPGLRRTLVPGPRRTLAPPVLGSRPKSP
QLQAAAASGAARSRPRTVSSMGNGTSRLYSALAKTVNSSAAAQHPEYLVSTDPEHLEP
IDPKELLEECRAVLHTRPPRYQRDFVDLRTDCSSSHSPIRVMQWNILAQALGEGKDNF
VQCPVEALKWEERKCLILEEILAYQPDILCLQEVDHYFDTFQPLLSRLGYQGTFFPKP
WSPCLDVEHNNGPDGCALFFLQNRFKLISSTNIRLTAMTLKTNQVAIAQTLECKESGR
QFCIAVTHLKARTGWERFRSAQGCDLLQNLQNITQGAKIPLIVCGDFNAEPTEEVYKH
FASSSLNLNSAYKLLSPDGQSEPPYTTWKIRTSGECRHTLDYIWYSRHALSVTSALDL
LTEEQIGPNRLPSFHYPSDHLSLVCDFSFNEEPHELF"
ORIGIN

Alignment Scores:

Pred. No.:	0.0122	Length:	3037
Score:	140.00	Matches:	89
Percent Similarity:	35.0%	Conservative:	55
Best Local Similarity:	21.7%	Mismatches:	163
Query Match:	7.4%	Indels:	104
DB:	9	Gaps:	19

US-10-757-745-2 (1-362) x AF183960 (1-3037)

```
Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArg---------------- 25
Db        782 CAGGCGGCGGCGGCCTCGGGCGCCGCGAGGTCGCGTCCCCGAACAGTGAGTTCCATGGGA 841

Qy         26 -----------ArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
Db        842 AACGGCACCAGTCGACTCTACAGTGCTCTCGCCAAGACCGTCAACAGCAGTGCAGCTGCT 901

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
Db        902 CAGCACCCGGAGTACTTGGTGTCAACTGACCCCGAA--------------------CAT 940

Qy         62 PheGluPro------ProValGlu--------------GluSerAlaLeuGluArgArg 74
Db        941 CTGGAGCCCATCGATCCTAAGGAACTTCTGGAGGAATGCAGGGCTGTTCTGCACACTCGG 1000

Qy         75 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 94
Db       1001 CCACCCCGCTACCAG------CGGGATTTTGTGGACCTG------------AGGACAGAT 1042

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114
Db       1043 TGCTCCAGCAGCCAC----------------------------------TCTCCCATT 1066

Qy        115 SerLeuIleThrTrpAsnIle----------------AspGlyLeuAsp--------- 125
Db       1067 CGCGTCATGCAGTGGAACATCCTCGCCCAAGCTCTCGGAGAAGGCAAAGACAACTTTGTG 1126

Qy        126 --------------LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
Db       1127 CAGTGCCCCGTGGAAGCGCTCAAATGGGAAGAGAGGAAGTGCCTGATCCTGGAGGAGATC 1186

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
Db       1187 CTGGCTTACCAGCCAGACATACTGTGCCTCCAGGAAGTG---GACCACTACTTTGACACC 1243

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe--- 179
Db       1244 TTCCAGCCACTCCTCAGTAGA------------CTGGGCTACCAAGGCACGTTTTTCCCC 1291

Qy        180 ----------------------------------------------------ThrAlaIle 182
Db       1292 AAGCCCTGGTCACCATGTCTAGATGTGGAACACAACAACGGTCCAGATGGCTGTGCCTTA 1351

Qy        183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 202
Db       1352 TTTTTTCTCCAAAACCGATTCAAGCTTATCAGCAGCACCAATATTAGGCTGACAGCCATG 1411

Qy        203 LysMetMetArgAsnLeuLeuCysVal------------HisValAsnValSerGlyAsn 218
Db       1412 ACCCTAAAAACCAATCAGGTGGCCATCGCACAGACCCTGGAGTGCAAGGAGTCTGGGCGA 1471

Qy        219 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
```

```
Db       1472 CAGTTCTGCATTGCTGTCACCCACTTAAAAGCCCGCACTGGCTGGGAGCGGTTCCGGTCA 1531

Qy        239 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal--- 257
              |||           :::|||:::   :::|||         :::   |||   :::
Db       1532 GCTCAGGGCTGTGACCTCCTCCAGAACCTGCAGAACATCACCCAGGGAGCAAAGATCCCC 1591

Qy        258 ---IlePheAlaGlyAspThrAsnLeuArgAspArgGlu------ValThrArgCysGly 274
                 |||      ||||||   |||            |||
Db       1592 CTGATCGTCTGCGGGGACTTCAACGCAGAGCCAACCGAAGAGGTCTACAAACACTTTGCG 1651

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------GlyLysProLys 291
                  |||:::         ::::::   |||         |||:::   :::
Db       1652 TCCTCCAGCCTCAACCTCAACAGCGCCTACAAGCTGCTGAGTCCCGATGGACAGTCGGAG 1711

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
                  ||||||                  :::      :::     |||
Db       1712 CCTCCCTACACAACCTGG-----------------AAGATCCGGACCTCAGGCGAGTGT 1753

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              :::         |||   |||::::::         |||                 :::
Db       1754 CGCCACACGCTGGACTATATCTGGTACTCCAGACATGCTCTGAGT---------GTCACG 1804

Qy        332 ArgSerLeuAspLeuLeu--------------GlyLeuGluLysLeuAspCysGlyArg 346
                :::||||||||||||              |||      :::|||
Db       1805 TCCGCCCTGGATCTACTCACTGAAGAACAGATTGGGCCCAACCGGCTCCCATCCTTCCAT 1864

Qy        347 PheProSerAspHisTrpGlyLeuLeuCysAsn 357
              :::||||||||||||      |||:::|||:::
Db       1865 TACCCCTCGGACCACCTGTCCCTGGTGTGTGAC 1897
```

```
RESULT 71
AY437500
LOCUS       AY437500               2745 bp    DNA     linear   INV 23-JUN-2004
DEFINITION  Trypanosoma cruzi MP99 gene, complete cds.
ACCESSION   AY437500
VERSION     AY437500.2  GI:49096009
KEYWORDS    .
SOURCE      Trypanosoma cruzi
  ORGANISM  Trypanosoma cruzi
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma; Schizotrypanum.
REFERENCE   1  (bases 1 to 2745)
  AUTHORS   Worthey,E., Schnaufer,A., Mian,I.S., Stuart,K. and Salavati,R.
  TITLE     Comparative analysis of editosome proteins in trypanosomatids
  JOURNAL   Nucleic Acids Res. (2003) In press
REFERENCE   2  (bases 1 to 2745)
  AUTHORS   Worthey,E., Schnaufer,A., Mian,I.S., Stuart,K. and Salavati,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-2003) Seattle Biomedical Research Institute, 4
            Nickerson Street, Seattle, WA 98109, USA
REFERENCE   3  (bases 1 to 2745)
  AUTHORS   Worthey,E., Schnaufer,A., Mian,I.S., Stuart,K. and Salavati,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2004) Seattle Biomedical Research Institute, 307
            Westlake Ave N, Seattle, WA 98109, USA
  REMARK    Sequence update by submitter
COMMENT     On Jun 23, 2004 this sequence version replaced gi:38073477.
FEATURES             Location/Qualifiers
     source          1. .2745
                     /organism="Trypanosoma cruzi"
                     /mol_type="genomic DNA"
                     /strain="CL Brenner"
                     /db_xref="taxon:5693"
     mRNA            <1. .>2745
                     /product="MP99"
     CDS             1. .2745
                     /note="similar to Trypanosoma brucei RNA-editing complex
                     protein MP99; putative exonuclease; TcMP99"
                     /codon_start=1
                     /product="MP99"
                     /protein_id="AAR10848.2"
                     /db_xref="GI:49096010"
                     /translation="MLQRCRILWTADFRTSPKISEEFNNVLNRSGWGFDHVAIDMGAL
                     AGAAGRLTRDLIGQDRDAETVKVLHRQVTNVILRRIQPKKSLAIFIDGSEPLWKVRHM
                     RLYPGKKFEGKFYRSAASPMVYLLEDKFRQVAVDLRTPPKEFVISGAGVPGPVEGKMS
                     AWMLDLASRALPDGAVTLSPPLLSSTPAEVTFNDTICLIGGTDLFLSALGATPFHNIT
                     TLTLQQGEMKSLSMSESLEWLAMDHLLKGEAATQGDQGTGRQELQKRLAAVRTDIVLL
                     YLMANGISATDLSGLGVNFKELMEAYMAEESAGASHEHIFEETPNGTSLLLSSAVLER
                     VLSRATRKTGGSPKPCQQSADYLEVLLQTHAMICGGGIRNFRWTPDDNDCVPEKPPKI
                     PIERFLAHLKHLTTLTREGGKGGGGGEVQGDGAEKKTTAIEKEQCRFTAGVEASVDWT
                     FALTGLETLLLTAPRAELIEQIVPVFTRGHSLPAGIADDIVSTKNVMEAAQKTRNILQ
                     VIKGNNSEEGGGAVKDHPALDHHPSHYFVRTPGCRGPPPGWSYQSINLGVRAAGFGVR
                     YRTGVGASAATSRVLDGVGVASKNTLFVYAADKNGVWSEEPCGEIHSPVLPTPTEELS
                     TLRVLTWNVQFNRHSGERTPLGRDGIDWCTPTRYIALAKTMETLDADVISMQEVEPAW
                     CKYLCEQQWVRERYALSCNEDSHAIRPWGVMLLVKRHLRVEGMHHANVAAFAGHTSVM
                     PEVTITVSKNVPVTIGSLHLLAPYNKNNVSNRTTQLENLTKRMRARPSAGGVSTGAIV
                     MGDFNDYANNYFTPPPELGFKDAWSLLNPENNKNEGYTIDGDRNPYAGQIIEREFFGR
                     ADRVLFASRHLQPIHAELVGTTSVRALGITKQVNCDKAVPEYLYPSDHFGLLVEFQVM
                     "
ORIGIN

Alignment Scores:
Pred. No.:              0.0121          Length:         2745
Score:                  139.50          Matches:        73
Percent Similarity:     37.7%           Conservative:   53
Best Local Similarity:  21.9%           Mismatches:     103
Query Match:            7.3%            Indels:         105
DB:                     2               Gaps:           17

US-10-757-745-2 (1-362) x AY437500 (1-2745)

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||          |||      |||      |||         ::::::|||||||||
Db       1837 TCTCCTGTGTTACCAACGCCAACGGAGGAGCTTTCGACGCTTCGGGTTTTAACGTGGAAT 1896

Qy        121 Ile-----------------------------------------AspGlyLeuAspLeu 126
              :::                                        ||||||:::|||
Db       1897 GTGCAATTCAATCGTCACAGTGGTGAGCGGACTCCACTGGGCCGAGACGGTATTGATTGG 1956

Qy        127 AsnAsnLeuSerGluArgAlaArgGlyValCysSer---------TyrLeuAlaLeu--- 142
                                          |||:::         |||:::||||||
Db       1957 ----------------------------TGCACACCAACACGGTACATTGCGCTTGCC 1986

Qy        143 ---------------TyrSerProAspValIlePheLeuGlnGluValIleProProTyr 157
                            |||||||||   :::|||||||||   |||   :::
Db       1987 AAGACAATGGAAACCCTAGATGCAGATGTGATCAGTATGCAGGAAGTGGAGCCGGCGTGG 2046

Qy        158 TyrSerTyrLeuLysLysArgSerSerAsnTyrGlu------IleIleThrGlyHisGlu 175
                  ||||||  ::::::          |||         ::::::   :::|||
Db       2047 TGTAAATACCTCTGTGAACAACAATGGGTAAGGGAACGGTATGCGCTTTCTTGTAATGAA 2106

Qy        176 GluGlyTyr------------PheThrAlaIleMetLeuLysLysSerArgValLysLeu 191
              :::   :::           :::      ::::::|||   |||      :::::::::
Db       2107 GATAGTCATGCCATTCGCCCTTGGGGTGTCATGCTGCTTGTAAAGCGACACCTGCGTGTG 2166

Qy        192 LysSerGln-----------------------------------GluIleIleProPhe 199
              :::                                        ::::::|||
Db       2167 GAAGGGATGCATCATGCAAATGTGGCTGCTTTTGCAGGACACACGAGTGTGATGCCAGAA 2226

Qy        200 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 219
                 :::     :::   :::|||              |||   |||:::
Db       2227 GTTACGATTACCGTTTCCAAAAAT--------------GTTCCAGTGACAATTGGTTCA 2271

Qy        220 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
              |||   |||:::      :::           :::   :::   :::   |||
Db       2272 CTTCATCTTCTTGCACCATAC---------AACAAAAACAATGTCAGCAATCGTACAACA 2322

Qy        240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal------ 257
              ||||||:::   ::   |||:::|||:::      |||   :::      |||
Db       2323 CAATTGGAAAACCTGACGAAACGGATGCGCGCACGCCCCTCTGCCGGTGGTGTGTCGACC 2382

Qy        258 ------IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
                    |||      ||||||   |||
Db       2383 GGCGCAATTGTTATGGGAGACTTTAAT---------------------------GAT 2412

Qy        276 LeuProAsnAsn--------------------------IleValAspValTrpGluPhe 286
```

```
Db      2413 TATGCGAATAATTACTTTACGTTTCCACCCGAGCTGGGCTTTAAAGACGCCTGGTCGCTG 2472

Qy       287 LeuGly------LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304

Db      2473 CTTAACCCAGAGAATAACAAGAATGAAGGGTACACCATCGACGGAGATCGCAATCCATAC 2532

Qy       305 LeuGly------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAla 322

Db      2533 GCGGGACAAATCATCGAGCGGGAGTTTTTGGTCGTGCCGATCGTGTGTTGTTTGCCTCC 2592

Qy       323 AlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGly------------ 338

Db      2593 CGC-----------CATTTGCAGCCCATCCACGCAGAACTTGTCGGCACTACCTCTGTA 2640

Qy       339 --------------LeuGluLysLeuAspCysGlyArg-----------------Phe 347

Db      2641 CGTGCGCTTGGGATCACGAAGCAGGTGAACTGCGACAAGGCTGTCCCCGAGTATTTGTAC 2700

Qy       348 ProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361

Db      2701 CCCTCTGACCACTTTGGTCTTCTTGTCGAGTTTCAGGTGATG 2742
```

RESULT 72
AY437501
LOCUS AY437501 2778 bp DNA linear INV 04-NOV-2003
DEFINITION Trypanosoma cruzi MP100 gene, complete cds.
ACCESSION AY437501
VERSION AY437501.1 GI:38073479
KEYWORDS .
SOURCE Trypanosoma cruzi
ORGANISM Trypanosoma cruzi
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma; Schizotrypanum.
REFERENCE 1 (bases 1 to 2778)
AUTHORS Worthey,E., Schnaufer,A., Mian,I.S., Stuart,K. and Salavati,R.
TITLE Comparative analysis of editosome proteins in trypanosomatids
JOURNAL Nucleic Acids Res. (2003) In press
REFERENCE 2 (bases 1 to 2778)
AUTHORS Worthey,E., Schnaufer,A., Mian,I.S., Stuart,K. and Salavati,R.
TITLE Direct Submission
JOURNAL Submitted (13-OCT-2003) Seattle Biomedical Research Institute, 4
Nickerson Street, Seattle, WA 98109, USA
FEATURES Location/Qualifiers
source 1. .2778
/organism="Trypanosoma cruzi"
/mol_type="genomic DNA"
/strain="CL Brenner"
/db_xref="taxon:5693"
mRNA <1. .>2778
/product="MP100"
CDS 1. .2778
/note="similar to Trypanosoma brucei RNA-editing complex
protein MP100; putative exonuclease; TcMP100"
/codon_start=1
/product="MP100"
/protein_id="AAR10849.1"
/db_xref="GI:38073480"
/translation="MALSRTWCRFAVTTYRQSKVSDRFQLLITKKEVECDYVAVDATA
LVAMGLRLAKNVTTEQRRHKEVARHVAQTLQQLLKKVRCKKSLLIAMDGAEALLKAER
TRTSSLTRKVEGRLMRLPGTALMQAVEERLVRMMPERQIVPGEVVPSGTCVHGCVEQK
MSAWALDLASRETFYGSVDSLCFIGASELYLNVMALSPFYNITSVVQNNADLRQMRLQ
DILEWLELDKKAKEGESVAIAKMRTDILFLFLLAHGASATELNPISGVGFQELVERYM
SSITATPTPDATPSPSGHQATSSASAEGSFLFEDVSGNALRLNVPLLCRILQMVARKE
AVPRTDAVSEMHLEHTLQTHAMLCLGGTPEYDYLPIMPSPPQAHSSTPTLVQFAGHLS
ALCGTSPYRIARCAVPRLNATGSTEKNEQTPDKRPLTAAEYTILCQSLPATVESLIHQ
YVGVAPKPDVGKTITTASTVEAHRIVREVLSYANPMKPNKCLCLSPSYCWLQSEKTQL
WRFEYVDVGVRSYQGETRRERNAIKGMTLEINMSRDGPSHFAPSSGTWEPIVNFPCGA
EGEKGATTSTIKGTEATATALQEKDLTKKDTESPPATATTTENEVKLEATAQKNASDS
VLTPHSARSSKEGGRSSNKNQTTLKLLTWNVMFDRYSGKPTPLGMPGIDWCSPKRYPV
LARVIQQEDADVVGMQEVEPVFWEFLSKQPWMRENYYFSCGYTSPSISPWGVLMLVHR
RRFPVQSITYLNVPAWTNHVSLMPVVCLQLPHGPVHVAAAHLLAPYTKSHETARTSQD
SALRHHMLKTLTGDVVTMGDFNDWPTNEFLMPAESQYVDCWPLIYHTMLENDGRDEYI
LQTECKLKVEEMFFGRSDKLFLRSRRLAPVEGHLVGTKSVNDENGNKEAPGYLFPSDH
YGVSMTFALKP"
ORIGIN

Alignment Scores:

Pred. No.:	0.0136	Length:	2778
Score:	139.00	Matches:	80
Percent Similarity:	37.7%	Conservative:	55
Best Local Similarity:	22.3%	Mismatches:	127
Query Match:	7.3%	Indels:	96
DB:	2	Gaps:	20

US-10-757-745-2 (1-362) x AY437501 (1-2778)

```
Qy        66 ValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrVal 85

Db      1759 TTACAGGAGAAGGATTTAACAAAGAAG---GATACAGAATCGCCACCAGCAACAGCAACG 1815

Qy        86 AspLeuThrAsnGluGluThrThrAspSerThrThrSerLys------------------ 99

Db      1816 ACAACGGAAAATGAAGTTAAATTGGAAGCGACGGCGCAGAAAAATGCTTCTGATTCTGTG 1875

Qy       100 IleSerPro---SerGluAspThrGlnGlnGluAsnGly---------------------- 111

Db      1876 TTAACTCCTCACTCGGCGCGTAGCAGCAAAGAAGGAGGTCGATCATCCAACAAAAACCAG 1935

Qy       112 SerMetPheSerLeuIleThrTrpAsnIle------------------------------ 121

Db      1936 ACGACATTAAAGTTACTTACGTGGAATGTTATGTTTGATCGGTATAGCGGAAAGCCAACC 1995

Qy       122 ------------AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCys 137

Db      1996 CCGTTGGGAATGCCTGGTATTGATTGGTGTTCC---CCGAAACGGTACCCCGTTCTCGCA 2052

Qy       138 SerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyr 157

Db      2053 AGGGTTATTCAGCAAGAGGATGCAGATGTTGTAGGCATGCAAGAAGTTGAACCCGTTTTC 2112

Qy       158 TyrSerTyrLeuLysLysArg---------SerSerAsnTyrGluIleIleThrGlyHis 174

Db      2113 TGGGAATTTCTTTCAAAACAGCCGTGGATGCGTGAGAATTATTACTTCTCATGCGGTTAC 2172

Qy       175 Glu------------GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db      2173 ACAAGTCCATCTATTTCACCTTGGGGTGTGCTCATGCTCGTGCACCGCCGGCGATTTCCA 2232

Qy       191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db      2233 GTTCAAAGCATTACATATCTAAATGTTCCCGCGTGGACAAATCACGTCTCTTTAATGCCC 2292

Qy       211 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis------------ 226

Db      2293 GTT-----------------------GTGTGTCTTCAGTTGCCACATGGGCCCGTGCAC 2328

Qy       227 --------------------LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239

Db      2329 GTGGCCGCTGCACATCTTTTGGCCCCGTACACGAAATCTCATGAGACGGCACGTACCTCC 2388

Qy       240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 259

Db      2389 CAA---GATTCCGCGTTGCGGCATCACATGTTAAAGACATTGACTGGTGATGTTGTGACG 2445

Qy       260 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsn 279

Db      2446 ATGGGCGACTTTAAC----------------------------GACTGGCCCACAAAT 2475

Qy       280 IleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr------Trp 297

Db      2476 --------------GAGTTTCTT---ATGCCGGCGGAATCACAATACGTTGATTGTTGG 2517

Qy       298 AspThr---------GlnMetAsnSerAsnLeuGly---------------IleThrAla 309

Db      2518 CCGCTTATTTACCACACGATGCTGGAAAACGATGGACGAGACGAATACATTTTGCAAACT 2577
```

```
Qy        310 AlaCysLysLeuArgPheAspArgIlePhePhe---------------------------- 320
              |||||||||:::   :::  :::||||||
Db       2578 GAATGCAAACTGAAGGTGGAAGAGATGTTTTTTGGACGTTCGGATAAATTGTTTTTGCGC 2637

Qy        321 ------ArgAlaAlaAlaGluGluGlyHisIleIle---ProArgSerLeu---AspLeu 336
                   |||   |||      |||||||||::::::      :::|||:::   |||
Db       2638 AGTCGTCGTCTTGCTCCGGTGGAGGGGCACCTTGTAGGCACCAAAAGTGTGAACGACGAG 2697

Qy        337 LeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeu 354
               |||   ::::::              ||||||||||||||:::|||:::
Db       2698 AATGGCAATAAGGAAGCTCCTGGCTACTTGTTTCCCTCGGATCACTACGGTGTG 2751

RESULT 73
BX927200
LOCUS       BX927200              152420 bp    DNA     linear   HTG 24-JAN-2004
DEFINITION  Danio rerio clone CH211-116A20, WORKING DRAFT SEQUENCE, 8 unordered
            pieces.
ACCESSION   BX927200
VERSION     BX927200.2  GI:41322788
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 152420)
  AUTHORS   Burton,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 25, 2004 this sequence version replaced gi:41223462.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zC116A20
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 150420 bases at least Q40
            Consensus quality: 151102 bases at least Q30
            Consensus quality: 151513 bases at least Q20
            Insert size: 151720; sum-of-contigs
            Insert size: 167766; 4.0% error; agarose-fp
            Quality coverage: 9.59x in Q20 bases; sum-of-contigs Quality
            coverage: 8.84x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    33759: contig of 33759 bp in length
            *    33760    33859: gap of 100 bp
            *    33860    43530: contig of 9671 bp in length
            *    43531    43630: gap of 100 bp
            *    43631    48335: contig of 4705 bp in length
            *    48336    48435: gap of 100 bp
            *    48436    65568: contig of 17133 bp in length
            *    65569    65668: gap of 100 bp
            *    65669    69685: contig of 4017 bp in length
            *    69686    69785: gap of 100 bp
            *    69786    76321: contig of 6536 bp in length
            *    76322    76421: gap of 100 bp
            *    76422   120851: contig of 44430 bp in length
            *   120852   120951: gap of 100 bp
            *   120952   152420: contig of 31469 bp in length.
FEATURES             Location/Qualifiers
     source          1. .152420
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-116A20"
                     /clone_lib="CHORI-211"
     misc_feature    1. .33759
                     /note="assembly_fragment:01133
                     fragment_chain:1
                     clone_end:T7
                     vector_side:left"
     misc_feature    33860. .43530
                     /note="assembly_fragment:00194
                     fragment_chain:1"
     misc_feature    43631. .48335
                     /note="assembly_fragment:00100
                     fragment_chain:1"
     misc_feature    48436. .65568
                     /note="assembly_fragment:00346
                     fragment_chain:1"
     misc_feature    65669. .69685
                     /note="assembly_fragment:00012
                     fragment_chain:1"
     misc_feature    69786. .76321
                     /note="assembly_fragment:00044
                     fragment_chain:1"
     misc_feature    76422. .120851
                     /note="assembly_fragment:01745
                     fragment_chain:1"
     misc_feature    120952. .152420
                     /note="assembly_fragment:00585
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:              1.49          Length:          152420
Score:                  139.00        Matches:         71
Percent Similarity:     38.5%         Conservative:    54
Best Local Similarity:  21.8%         Mismatches:      108
Query Match:            7.3%          Indels:          92
DB:                     14            Gaps:            15

US-10-757-745-2 (1-362) x BX927200 (1-152420)

Qy        100 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 119
              :::|||   :::::::::||||||:::   :::         ||||||   ::::::
Db     129780 CTTAGTGATACTAAAAACACTCAAAAGTGATCTTCTCAATCTTTTTCTTGTGTATCTTCA 129839

Qy        120 AsnIle------------AspGly--------------LeuAspLeuAsnAsnLeuSer 130
                 :::            ||||||              ::::::      |||   |||
Db     129840 GCTCTTGGCGAAGGCAAAGATGGTTTTGTGCGCTGTCCAATGGAGGCGTTGAACTGGTCG 129899

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ||||||       :::      :::      |||   |||||||||:::   |||
Db     129900 GAGCGCAAGTATCTCATCCTAGAAGAGATCCTCACCTACAGGCCAGACGTGGTGTGTCTG 129959

Qy        151 GlnGluVal-----------------IleProProTyrTyrSerTyrLeuLysLysArg 164
              |||||||||                 |||      |||   |||      :::
Db     129960 CAAGAGGTGGACCATTACTTCGACACGTTCCAGCCTGTGCTCTCTAGCCTGGGCTACCAG 130019

Qy        165 SerSer---------------------------------------------------- 166
              ||||||
Db     130020 AGCAGCTTCTGCCCCAAACCCTGCTCCCCCTGCCTGGACGTCCACAATAACAACGGCCCG 130079

Qy        167 ----------------------------AsnTyrGluIleIle---ThrGlyHisGlu 175
                                         ::::::::::::   |||   |||
Db     130080 GACGGCTGCGCTCTGTTCTTCAACCGCCGCCGCTTCCAGATGCTCCACACGGCGCAT--- 130136

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
```

```
Db     130137 ------------------------CTCAGACTCTCCGCCATGATGCTCAAAACGAACCAA 130172

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              :::                 ::::::       |||   |||          :::
Db     130173 GTA-----------------GCCGTCGTCGCTACGCTTCGATGT---------AAGCTC 130205

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              :::|||        |||:::      :::||||||::::::        |||     |||
Db     130206 ACAGGACGAGTTTTCTGTGTCGCCGTAACGCATCTCAAAGCACGAAGCGGATGGGAAGCA 130265

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu--------------- 250
                 |||       |||         :::|||::::::::::   |||
Db     130266 TTTCGTAGTGCGCAAGGAGCAAATCTCCTCCAGCAGCTTCACGAAATAACCTCCCAATCA 130325

Qy        251 ------------------------AlaProGluSerAlaThrValIlePheAlaGlyAsp 262
                                      |||       :::|||      ||||||
Db     130326 AACCCAGAGATGCATCAGGACGACCAAACAGAGGGTATTCCTTTGATAGTGTGTGGGGAT 130385

Qy        263 ThrAsnLeuArgAspArgGluValThr------------ArgCysGlyGlyLeuProAsn 278
                 |||             |||                |||         |||
Db     130386 TTTAATGCAGAGCCCAACGAGGAAGTGTATCGGCATTTTAGGAGTTCTTCATTA------ 130439

Qy        279 AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln------TyrThr 296
                 :::   |||::::::  |||          :::      :::      ||||||
Db     130440 GGGCTAGACAGCGTTTATAAATGTCTAAGTGACGACAGGACCACTGAGCCGCCTTATACA 130499

Qy        297 ---TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
                 |||      :::   :::                    |||
Db     130500 AGCTGGAAAATCCGGCCTAGTGGAGAG-----------------TGCTGCTCCACGCTG 130541

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg------Ser 333
              |||    |||::::::          |||   |||          :::   |||      |||
Db     130542 GACTATATTTGGTATTCGGAGAAAGCCTTCGAGGTGGACGCTGTGCTGAGGATACCGAGT 130601

Qy        334 LeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGly 353
                 :::   :::|||   ::::::|||           :::||||||||||||
Db     130602 GAAGAGCAGATCGGCCCGGACCGGCTTCCTTCATTTCATTATCCGTCAGATCATCTGTCT 130661

Qy        354 LeuLeuCysAsnLeu 358
              |||:::|||:::|||
Db     130662 CTGGTGTGTGACCTC 130676
```

```
RESULT 74
AL929007
LOCUS       AL929007              187464 bp    DNA     linear   VRT 25-JAN-2003
DEFINITION  Zebrafish DNA sequence from clone CH211-235E9 in linkage group 1,
            complete sequence.
ACCESSION   AL929007
VERSION     AL929007.6  GI:27902014
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 187464)
  AUTHORS   Corby,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 25, 2003 this sequence version replaced gi:27802813.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Repeat names
            beginning 'Dr' were identified by the Recon repeat discovery system
            (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr'
            were identified by Rick Waterman (Stephen Johnson lab, WashU). For
            further information see http://www/Projects/D_rerio/fishmask.shtml
            CH211-235E9 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .187464
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-235E9"
                     /clone_lib="CHORI-211"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                1.9          Length:          187464
Score:                    139.00       Matches:         71
Percent Similarity:       38.5%        Conservative:    54
Best Local Similarity:    21.8%        Mismatches:      108
Query Match:              7.3%         Indels:          92
DB:                       5            Gaps:            15

US-10-757-745-2 (1-362) x AL929007 (1-187464)

Qy        100 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 119
              :::|||   :::::::::||||||:::   :::          ||||||   ::::::
Db      42241 CTTAGTGATACTAAAAACACTCAAAAGTGATCTTCTCAATCTTTTTCTTGTGTATCTTCA 42300

Qy        120 AsnIle------------AspGly---------------LeuAspLeuAsnAsnLeuSer 130
                 :::            ||||||              ::::::      |||   |||
Db      42301 GCTCTTGGCGAAGGCAAAGATGGTTTTGTGCGCTGTCCAATGGAGGCGTTGAACTGGTCG 42360

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ||||||      :::     :::        |||   ||||||||||:::   |||
Db      42361 GAGCGCAAGTATCTCATCCTAGAAGAGATCCTCACCTACAGGCCAGACGTGGTGTGTCTG 42420

Qy        151 GlnGluVal------------------IleProProTyrTyrSerTyrLeuLysLysArg 164
              |||||||||                |||      |||   |||      :::
Db      42421 CAAGAGGTGGACCATTACTTCGACACGTTCCAGCCTGTGCTCTCTAGCCTGGGTTACCAG 42480

Qy        165 SerSer------------------------------------------------------ 166
              ||||||
Db      42481 AGCAGCTTCTGCCCCAAACCCTGCTCCCCCTGCCTGGACGTCCACAATAACAACGGCCCG 42540

Qy        167 ------------------------------AsnTyrGluIleIle---ThrGlyHisGlu 175
                                           ::::::::::::   |||   |||
Db      42541 GACGGCTGCGCTCTGTTCTTCAACCGCCGCCGCTTCCAGATGCTCCACACGGCGCAT--- 42597

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
                                  |||:::   |||   :::   ||||||:::   :::
Db      42598 ----------------------CTCAGACTCTCCGCCATGATGCTCAAAACCAACCAA 42633

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              :::                 ::::::       |||   |||          :::
Db      42634 GTA-----------------GCCGTCGTCGCTACGCTTCGATGT---------AAGCTC 42666

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
```

```
Db      42667 ACAGGACGAGTTTTCTGTGTCGCCGTAACGCATCTCAAAGCACGAAGCGGATGGGAAGCA 42726

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu---------------- 250

Db      42727 TTTCGTAGTGCGCAAGGAGCAAATCTCCTCCAGCAGCTTCACGAAATAACCTCCCAATCA 42786

Qy        251 ------------------------AlaProGluSerAlaThrValIlePheAlaGlyAsp 262

Db      42787 AACCCAGAGATGCATCAGGACGACCAAACAGAGGGTATTCCTTTGATAGTGTGTGGGGAT 42846

Qy        263 ThrAsnLeuArgAspArgGluValThr------------ArgCysGlyGlyLeuProAsn 278

Db      42847 TTTAATGCAGAGCCCAACGAGGAAGTGTATCGGCATTTTAGGAGTTCTTCATTA------ 42900

Qy        279 AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln------TyrThr 296

Db      42901 GGGCTAGACAGCGTTTATAAATGTCTAAGTGACGACAGGACCACTGAGCCGCCTTATACA 42960

Qy        297 ---TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315

Db      42961 AGCTGGAAAATCCGGCCTAGTGGAGAG-----------------TGCTGCTCCACGCTG 43002

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg------Ser 333

Db      43003 GACTATATTTGGTATTCGGAGAAAGCCTTCGAGGTGGATGCTGTGCTGAGGATACCGAGT 43062

Qy        334 LeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGly 353

Db      43063 GAAGAGCAGATCGGCCCGGACCGGCTTCCTTCATTTCATTATCCGTCAGATCATCTGTCT 43122

Qy        354 LeuLeuCysAsnLeu 358

Db      43123 CTGGTGTGTGACCTC 43137
```

```
RESULT 75
BX004778/c
LOCUS       BX004778               208135 bp    DNA     linear   VRT 29-MAR-2003
DEFINITION  Zebrafish DNA sequence from clone DKEY-87L3 in linkage group 1,
            complete sequence.
ACCESSION   BX004778
VERSION     BX004778.3  GI:28611072
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 208135)
  AUTHORS   Tromans,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 28, 2003 this sequence version replaced gi:28300949.
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Repeat names
            beginning 'Dr' were identified by the Recon repeat discovery system
            (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr'
            were identified by Rick Waterman (Stephen Johnson lab, WashU). For
            further information see http://www/Projects/D_rerio/fishmask.shtml
            DKEY-87L3 is from a Zebrafish BAC library
            VECTOR: pIndigoBAC-5.
FEATURES             Location/Qualifiers
     source          1. .208135
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-87L3"
                     /clone_lib="DanioKey"
ORIGIN

Alignment Scores:
Pred. No.:                2.15          Length:          208135
Score:                    139.00        Matches:         71
Percent Similarity:       38.5%         Conservative:    54
Best Local Similarity:    21.8%         Mismatches:      108
Query Match:              7.3%          Indels:          92
DB:                       5             Gaps:            15

US-10-757-745-2 (1-362) x BX004778 (1-208135)

Qy        100 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 119

Db      75505 CTTAGTGATACTAAAAACACTCAAAAGTGATCTTCTCAATCTTTTTCTTGTGTATCTTCA 75446

Qy        120 AsnIle------------AspGly--------------LeuAspLeuAsnAsnLeuSer 130

Db      75445 GCTCTTGGCGAAGGCAAAGATGGTTTTGTGCGCTGTCCAATGGAGGCGTTGAACTGGTCG 75386

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db      75385 GAGCGCAAGTATCTCATCCTAGAAGAGATCCTCACCTACAGGCCAGACGTGGTGTGTCTG 75326

Qy        151 GlnGluVal------------------IleProProTyrTyrSerTyrLeuLysLysArg 164

Db      75325 CAAGAGGTGGACCATTACTTCGACACGTTCCAGCCTGTGCTCTCTAGCCTGGGTTACCAG 75266

Qy        165 SerSer---------------------------------------------------------- 166

Db      75265 AGCAGCTTCTGCCCCAAACCCTGCTCCCCCTGCCTGGACGTCCACAATAACAACGGCCCG 75206

Qy        167 ----------------------------AsnTyrGluIleIle---ThrGlyHisGlu 175

Db      75205 GACGGCTGCGCTCTGTTCTTCAACCGCCGCCGCTTCCAGATGCTCCACACGGCGCAT--- 75149

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195

Db      75148 ----------------------CTCAGACTCTCCGCCATGATGCTCAAAACCAACCAA 75113

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215

Db      75112 GTA-----------------GCCGTCGTCGCTACGCTTCGATGT---------AAGCTC 75080

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235

Db      75079 ACAGGACGAGTTTTCTGTGTCGCCGTAACGCATCTCAAAGCACGAAGCGGATGGGAAGCA 75020

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu---------------- 250

Db      75019 TTTCGTAGTGCGCAAGGAGCAAATCTCCTCCAGCAGCTTCACGAAATAACCTCCCAATCA 74960

Qy        251 ------------------------AlaProGluSerAlaThrValIlePheAlaGlyAsp 262
```

```
Db      74959 AACCCAGAGATGCATCAGGACGACCAAACAGAGGGTATTCCTTTGATAGTGTGTGGGGAT 74900

Qy        263 ThrAsnLeuArgAspArgGluValThr----------ArgCysGlyGlyLeuProAsn 278

Db      74899 TTTAATGCAGAGCCCAACGAGGAAGTGTATCGGCATTTTAGGAGTTCTTCATTA------ 74846

Qy        279 AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln------TyrThr 296

Db      74845 GGGCTAGACAGCGTTTATAAATGTCTAAGTGACGACAGGACCACTGAGCCGCCTTATACA 74786

Qy        297 ---TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315

Db      74785 AGCTGGAAAATCCGGCCTAGTGGAGAG----------------TGCTGCTCCACGCTG 74744

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg------Ser 333

Db      74743 GACTATATTTGGTATTCGGAGAAAGCCTTCGAGGTGGATGCTGTGCTGAGGATACCGAGT 74684

Qy        334 LeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGly 353

Db      74683 GAAGAGCAGATCGGCCCGGACCGGCTTCCTTCATTTCATTATCCGTCAGATCATCTGTCT 74624

Qy        354 LeuLeuCysAsnLeu 358

Db      74623 CTGGTGTGTGACCTC 74609
```

Search completed: March 10, 2006, 23:14:17
Job time : 13380.3 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 15:40:34 ; Search time 1178.43 Seconds
(without alignments)
1746.920 Million cell updates/sec

Title: US-10-757-745-6_COPY_2_45
Perfect score: 229
Sequence: 1 GDPGSEIIESVPPAGPEASE.....ENEDDIQFVSEGPSRPVLEY 44

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/absa/ABSSWEB_spool/US10757745/runat_09032006_170329_4058/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=absa02h
-USER=US10757745_@CGN_1_1_12397_@runat_09032006_170329_4058 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	229	100.0	230	2	BG944050	BG944050 ax45h02.x
	2	229	100.0	513	7	CR746485	CR746485 CR746485
	3	229	100.0	563	2	BF103860	BF103860 601647366
	4	229	100.0	586	2	BG942992	BG942992 ax32c07.x
	5	229	100.0	645	7	CN359916	CN359916 170005977
	6	229	100.0	682	8	DN999300	DN999300 TC121030
	7	229	100.0	701	3	BI465045	BI465045 603206445
	8	229	100.0	818	5	BU945621	BU945621 AGENCOURT
	9	229	100.0	831	2	BF687541	BF687541 602066646
	10	229	100.0	858	8	CX164545	CX164545 HESC2_18_
	11	229	100.0	916	6	CA453561	CA453561 AGENCOURT
	12	229	100.0	1006	3	BQ230874	BQ230874 AGENCOURT
	13	223	97.4	325	1	AL044854	AL044854 DKFZp434D
	14	223	97.4	382	1	AW865717	AW865717 QV3-SN002
c	15	223	97.4	443	1	AI017172	AI017172 ou28b06.x
	16	223	97.4	514	3	BM749349	BM749349 K-EST0024
c	17	223	97.4	551	6	CD627106	CD627106 56066665J
c	18	223	97.4	564	6	CD627110	CD627110 56066681J
	19	223	97.4	568	6	CD627109	CD627109 56066681H
	20	223	97.4	583	3	BP229655	BP229655 BP229655
	21	223	97.4	583	3	BP248461	BP248461 BP248461
	22	223	97.4	604	7	CN481259	CN481259 hw05c03.y
	23	223	97.4	630	7	CV024153	CV024153 1345 Full
	24	223	97.4	643	8	DN991049	DN991049 TC111228
	25	223	97.4	671	2	BG722501	BG722501 602693752
	26	223	97.4	693	7	CN314844	CN314844 170004249
	27	223	97.4	694	7	CN314843	CN314843 170004249
	28	223	97.4	695	3	BI559958	BI559958 603253421
	29	223	97.4	749	2	BG701977	BG701977 602683351
	30	223	97.4	752	2	BE613457	BE613457 601504223
	31	223	97.4	760	7	CN314848	CN314848 170005318
	32	223	97.4	829	7	CT002116	CT002116 CT002116
	33	222	96.9	599	7	CR767845	CR767845 DKFZp469C
	34	222	96.9	691	2	BG777259	BG777259 602664490
	35	222	96.9	5087	4	CR858845	CR858845 Pongo pyg
	36	219	95.6	601	1	AL703808	AL703808 DKFZp686K
	37	218	95.2	362	5	BY300743	BY300743 BY300743
	38	218	95.2	366	5	BY230029	BY230029 BY230029
	39	218	95.2	378	5	BY307103	BY307103 BY307103
	40	218	95.2	388	5	BY094255	BY094255 BY094255
	41	218	95.2	392	5	BY024214	BY024214 BY024214
	42	218	95.2	398	5	BY019311	BY019311 BY019311
	43	218	95.2	620	3	BI463017	BI463017 603204540
	44	218	95.2	644	2	BB657575	BB657575 BB657575
	45	218	95.2	653	2	BB626942	BB626942 BB626942
	46	218	95.2	671	2	BB616798	BB616798 BB616798
	47	218	95.2	715	5	BY715342	BY715342 BY715342
	48	218	95.2	914	5	BQ952426	BQ952426 AGENCOURT
	49	218	95.2	929	5	BQ886508	BQ886508 AGENCOURT
	50	218	95.2	1750	4	AK015797	AK015797 Mus muscu
	51	218	95.2	3651	4	AK030088	AK030088 Mus muscu
c	52	217	94.8	460	1	AW118103	AW118103 xe03d02.x
c	53	217	94.8	538	6	CD627108	CD627108 56066673J
	54	217	94.8	567	6	CD627107	CD627107 56066673H
	55	217	94.8	627	3	BM352924	BM352924 ig69a09.y
	56	216	94.3	768	2	BF667883	BF667883 602122117
	57	216	94.3	814	2	BF030066	BF030066 601557142
	58	215	93.9	362	5	BY064671	BY064671 BY064671
	59	215	93.9	702	7	CO738423	CO738423 S1LE04c20
	60	215	93.9	796	5	BQ769546	BQ769546 UI-M-FI0-
	61	213	93.0	368	5	BY009270	BY009270 BY009270
c	62	213	93.0	547	6	CD627104	CD627104 56066565J
	63	213	93.0	553	6	CD627103	CD627103 56066565H
	64	212	92.6	353	5	BY210989	BY210989 BY210989
	65	212	92.6	378	5	BY307104	BY307104 BY307104
	66	211	92.1	1165	3	BI560244	BI560244 603253376
	67	210	91.7	307	8	DN371969	DN371969 LIB3733-0
	68	210	91.7	764	8	DN371450	DN371450 LIB3733-0
	69	208	90.8	377	5	BY028275	BY028275 BY028275
	70	208	90.8	406	6	CB808042	CB808042 AMGNNUC:N
	71	208	90.8	443	6	CB789225	CB789225 AMGNNUC:M
	72	208	90.8	578	7	CK357744	CK357744 AGENCOURT
	73	208	90.8	878	8	DN933911	DN933911 AGENCOURT
	74	203	88.6	343	2	BB854831	BB854831 BB854831
	75	200	87.3	741	2	BG286553	BG286553 602382555

ALIGNMENTS

RESULT 1

```
BG944050/c
LOCUS       BG944050                230 bp    mRNA    linear   EST 15-JAN-2003
DEFINITION  ax45h02.x1 Hembase; Erythroid Progenitor Cells (LCB:ax library)
            Homo sapiens cDNA clone ax45h02 random, mRNA sequence.
ACCESSION   BG944050
VERSION     BG944050.1  GI:14343422
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 230)
  AUTHORS   Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
  TITLE     Gene expression in proliferating human erythroid cells
  JOURNAL   Genomics 59 (2), 168-177 (1999)
  PUBMED    10409428
COMMENT     Contact: Jeffery L. Miller
            Laboratory of Chemical Biology
            National Institute of Diabetes and Digestive and Kidney Diseases
            Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
            20892, USA
            Tel: 301 402 2373
            Fax: 301 435 5148
            Email: jm7f@nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 45  row: h  column: 02
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .230
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="ax45h02"
                     /sex="unknown"
                     /tissue_type="blood"
                     /cell_type="Erythroid Cells"
                     /cell_line="Primary Culture of Peripheral Blood
                     Mononuclear Cells"
                     /dev_stage="Progenitor; EPO responsive CD71++++"
                     /lab_host="SOLR"
                     /clone_lib="Hembase; Erythroid Progenitor Cells (LCB:ax
                     library)"
                     /note="Organ: blood; Vector: Lambda ZAP II; Site_1: EcoRI;
                     Site_2: EcoRI; 65,000 proliferating erythroid cells from
                     the buffy coat of a blood donation were obtained by flow
                     cytometric separation after a 5-day culture period in the
                     presence of erythropoietin. Total RNA was purified from
                     the sorted cell population using TRIzol reagent. RNA (0.3
                     ug) was converted into double stranded cDNA using
                     Clontech's CapFinder cDNA Library Construction Kit
                     (Clontech) according to the manufacturer's protocol and
                     cloned into EcoRI digested Lambda Zap II vector
                     (Stratagene). The phage library was amplified once prior
                     to in vivo excision in SOLR cells. Individual colonies
                     were grown, and the cDNA inserts were sequenced in high
                     throughput (NIH intramural sequencing center
                     http://www.nisc.nih.gov/)."
ORIGIN

Alignment Scores:
Pred. No.:                2.33e-19        Length:          230
Score:                    229.00          Matches:         44
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       2               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BG944050 (1-230)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        194 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCCGAGGCATCTGAG 135

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        134 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 75

Qy         41 ValLeuGluTyr 44
              ||||||||||||

Db         74 GTTCTTGAATAC 63

RESULT 2
CR746485
LOCUS       CR746485                513 bp    mRNA    linear   EST 30-AUG-2004
DEFINITION  CR746485 NIH_MGC_71 Homo sapiens cDNA clone IMAGp998B1912872 ;
            IMAGE:5785266 5', mRNA sequence.
ACCESSION   CR746485
VERSION     CR746485.1  GI:51659182
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 513)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Korn,B., Neubert,P., Partsch,E.,
            Peters,M., Radelof,U. and Schneider,D.
  TITLE     I.M.A.G.E. cDNA Clone Collection
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; IMAGp998B1912872.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .513
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998B1912872 ; IMAGE:5785266"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

Alignment Scores:
Pred. No.:                5.78e-19        Length:          513
Score:                    229.00          Matches:         44
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x CR746485 (1-513)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        185 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 244

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        245 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 304

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        305 GTTCTTGAATAC 316

RESULT 3
BF103860
LOCUS       BF103860                563 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  601647366F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:3931360 5',
            mRNA sequence.
ACCESSION   BF103860
VERSION     BF103860.1  GI:10886299
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 563)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM767  row: d  column: 17
            High quality sequence stop: 516.
FEATURES             Location/Qualifiers
     source          1. .563
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3931360"
                     /tissue_type="embryonal carcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_61"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
ORIGIN

Alignment Scores:
Pred. No.:              6.43e-19        Length:         563
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     2               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BF103860 (1-563)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 89

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         90 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 149

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        150 GTTCTTGAATAC 161

RESULT 4
BG942992
LOCUS       BG942992                586 bp    mRNA    linear   EST 15-JAN-2003
DEFINITION  ax32c07.x1 Hembase; Erythroid Progenitor Cells (LCB:ax library)
            Homo sapiens cDNA clone ax32c07 random, mRNA sequence.
ACCESSION   BG942992
VERSION     BG942992.1  GI:14342364
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 586)
  AUTHORS   Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
  TITLE     Gene expression in proliferating human erythroid cells
  JOURNAL   Genomics 59 (2), 168-177 (1999)
   PUBMED   10409428
COMMENT     Contact: Jeffery L. Miller
            Laboratory of Chemical Biology
            National Institute of Diabetes and Digestive and Kidney Diseases
            Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
            20892, USA
            Tel: 301 402 2373
            Fax: 301 435 5148
            Email: jm7f@nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 32  row: c  column: 07
            Seq primer: -21M13 forward primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .586
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="ax32c07"
                     /sex="unknown"
                     /tissue_type="blood"
                     /cell_type="Erythroid Cells"
                     /cell_line="Primary Culture of Peripheral Blood
                     Mononuclear Cells"
                     /dev_stage="Progenitor; EPO responsive CD71++++"
                     /lab_host="SOLR"
                     /clone_lib="Hembase; Erythroid Progenitor Cells (LCB:ax
                     library)"
                     /note="Organ: blood; Vector: Lambda ZAP II; Site_1: EcoRI;
                     Site_2: EcoRI; 65,000 proliferating erythroid cells from
                     the buffy coat of a blood donation were obtained by flow
                     cytometric separation after a 5-day culture period in the
                     presence of erythropoietin. Total RNA was purified from
                     the sorted cell population using TRIzol reagent. RNA (0.3
                     ug) was converted into double stranded cDNA using
                     Clontech's CapFinder cDNA Library Construction Kit
                     (Clontech) according to the manufacturer's protocol and
                     cloned into EcoRI digested Lambda Zap II vector
                     (Stratagene). The phage library was amplified once prior
                     to in vivo excision in SOLR cells. Individual colonies
                     were grown, and the cDNA inserts were sequenced in high
                     throughput (NIH intramural sequencing center
                     http://www.nisc.nih.gov/)."
ORIGIN

Alignment Scores:
Pred. No.:              6.72e-19        Length:         586
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
```

```
Query Match:            100.0%          Indels:         0
DB:                     2               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BG942992 (1-586)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 109

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 169

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        170 GTTCTTGAATAC 181

RESULT 5
CN359916
LOCUS       CN359916                645 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000597714119 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN359916
VERSION     CN359916.1  GI:47359850
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
  PUBMED    15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 645   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cell, retinoic acid and
                     mitogen-treated hES cell line H7"
                     /clone_lib="GRN_PRENEU"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from hES cell line H7 (p29) maintained in feeder-free
                     conditions. Embryoid bodies were generated in the presence
                     of all-trans retinoic and mitogens."
ORIGIN

Alignment Scores:
Pred. No.:              7.5e-19         Length:         645
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CN359916 (1-645)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 85

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 145

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        146 GTTCTTGAATAC 157

RESULT 6
DN999300
LOCUS       DN999300                682 bp    mRNA    linear   EST 17-MAY-2005
DEFINITION  TC121030 Human endothelial cells, large insert, pCMV expression
            library Homo sapiens cDNA clone TC121030 5' similar to PREDICTED:
            Homo sapiens KIAA1586 (KIAA1586), mRNA sequence.
ACCESSION   DN999300
VERSION     DN999300.1  GI:66259127
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L.,
            Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M.,
            Zhang,X., Jay,G. and He,W.
  TITLE     High-throughput cloning of full-length human cDNAs directly from
            cDNA libraries optimized for large and rare transcripts
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Kovacs, KF
            High Throughput cDNA Cloning
            OriGene Technologies, Inc. ( www.origene.com )
            6 Taft Court, Suite 100, Rockville, MD 20850, USA
            Tel: 301 340 3188
            Fax: 301 340 8606
            Email: cDNA@origene.com
            This EST submission is part of an on-going human full-length
            cloning project at OriGene Technologies, Inc.
            Please contact OriGene for access.
            OriGene Technologies, Inc.
            6 Taft Ct. Suite 100
            Rockville, MD 20850
            Tel: (301) 340-3188
            http://www.origene.com
            Seq primer: pCMV6 5prime forward vector primer, OriGene
            Technologies Inc.
FEATURES             Location/Qualifiers
     source          1. .682
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TC121030"
                     /cell_type="Endothelial cells"
                     /clone_lib="Human endothelial cells, large insert, pCMV
                     expression library"
                     /note="Vector: pCMV6-XL6; Site_1: EcoR1; Site_2: Xho1/Sal1
                     compatible end ligatio; Oligo-dT primed reverse
                     transcription optimized for large and GC rich mRNA
                     transcripts, cDNA size selection, optimized ligation for
                     large inserts into mammalian expression vector, random
                     clones selected for end sequence verification of
                     full-length genes"
ORIGIN

Alignment Scores:
Pred. No.:              7.99e-19        Length:         682
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x DN999300 (1-682)
```

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 150

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 210

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        211 GTTCTTGAATAC 222
```

RESULT 7
BI465045
LOCUS BI465045 701 bp mRNA linear EST 21-AUG-2001
DEFINITION 603206445F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5272042 5', mRNA sequence.
ACCESSION BI465045
VERSION BI465045.1 GI:15255701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 701)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11686 row: j column: 11
High quality sequence stop: 700.
FEATURES Location/Qualifiers
source 1. .701
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5272042"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:

Pred. No.:	8.24e-19	Length:	701
Score:	229.00	Matches:	44
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BI465045 (1-701)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 276

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 336

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        337 GTTCTTGAATAC 348
```

RESULT 8
BU945621
LOCUS BU945621 818 bp mRNA linear EST 18-OCT-2002
DEFINITION AGENCOURT_10545564 NIH_MGC_127 Homo sapiens cDNA clone IMAGE:6736449 5', mRNA sequence.
ACCESSION BU945621
VERSION BU945621.1 GI:24134440
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 818)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM3064 row: o column: 08
High quality sequence stop: 533.
FEATURES Location/Qualifiers
source 1. .818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6736449"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_127"
/note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidnney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 1-2 kb size fraction (other fractions present in NIH_MGC_126 and NIH_MGC_128). Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:

Pred. No.:	9.82e-19	Length:	818
Score:	229.00	Matches:	44
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0

```
DB:                    5              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x BU945621 (1-818)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 119

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 179

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        180 GTTCTTGAATAC 191

RESULT 9
BF687541
LOCUS       BF687541                831 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602066646F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4065922 5',
            mRNA sequence.
ACCESSION   BF687541
VERSION     BF687541.1  GI:11972949
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 831)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM902  row: k  column: 11
            High quality sequence stop: 652.
FEATURES             Location/Qualifiers
     source          1..831
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4065922"
                     /tissue_type="glioblastoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_57"
                     /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.55 kb (range 0.9-4.0 kb).  12/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
ORIGIN

Alignment Scores:
Pred. No.:              9.99e-19        Length:          831
Score:                  229.00          Matches:         44
Percent Similarity:     100.0%          Conservative:    0
Best Local Similarity:  100.0%          Mismatches:      0
Query Match:            100.0%          Indels:          0
DB:                     2               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BF687541 (1-831)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 102

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 162

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        163 GTTCTTGAATAC 174

RESULT 10
CX164545
LOCUS       CX164545                858 bp    mRNA    linear   EST 23-DEC-2004
DEFINITION  HESC2_18_A02.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone
            IMAGE:7467918 5', mRNA sequence.
ACCESSION   CX164545
VERSION     CX164545.1  GI:56794625
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 858)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15766  row: a  column: 04
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 858.
FEATURES             Location/Qualifiers
     source          1..858
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7467918"
                     /sex="male"
                     /tissue_type="human embryonic stem cells differentiated to
                     an early endodermal cell type"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_258"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early endodermal cell type. Cell line
                     id and NIH Registry designation is BG01. Positive for
                     GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP
                     expression. Passage number 40. cDNA primed using oligo-dT
                     primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and
                     cloned into the EcoRV/NotI sites of pExpress-1. This
                     primary library is non-normalized (normalized primary
                     library is NIH_MGC_259). It was constructed by Express
                     Genomics (Frederick, MD). Sequence ends have been trimmed
```

to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Note: this is a Mammalian Gene Collection library."

ORIGIN

Alignment Scores:

Pred. No.:	1.04e-18	Length:	858
Score:	229.00	Matches:	44
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x CX164545 (1-858)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 117

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 177

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        178 GTTCTTGAATAC 189
```

RESULT 11
CA453561
LOCUS CA453561 916 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT_10579185 NIH_MGC_127 Homo sapiens cDNA clone IMAGE:6717528 5', mRNA sequence.
ACCESSION CA453561
VERSION CA453561.1 GI:24902433
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 916)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM3034 row: f column: 24
High quality sequence stop: 215.

FEATURES Location/Qualifiers
source 1. .916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6717528"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_127"
/note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidnney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 1-2 kb size fraction (other fractions present in NIH_MGC_126 and NIH_MGC_128). Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	1.12e-18	Length:	916
Score:	229.00	Matches:	44
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x CA453561 (1-916)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 111

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 171

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        172 GTTCTTGAATAC 183
```

RESULT 12
BQ230874
LOCUS BQ230874 1006 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7260801 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5785266 5', mRNA sequence.
ACCESSION BQ230874
VERSION BQ230874.1 GI:20412274
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1006)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12872 row: b column: 19
High quality sequence stop: 647.

FEATURES Location/Qualifiers
source 1. .1006
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5785266"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb. "

```
ORIGIN

Alignment Scores:
Pred. No.:                1.24e-18        Length:         1006
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BQ230874 (1-1006)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 231

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 291

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        292 GTTCTTGAATAC 303

RESULT 13
AL044854
LOCUS       AL044854                325 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp434D063_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434D063 5', mRNA sequence.
ACCESSION   AL044854
VERSION     AL044854.1  GI:5433060
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 325)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     EST (Wambutt, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp434D063) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .325
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434D063"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="434 (synonym: htes3)"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN

Alignment Scores:
Pred. No.:                1.99e-18        Length:         325
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       1               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x AL044854 (1-325)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 133

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        134 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 193

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        194 GTTCTTGAATAC 205

RESULT 14
AW865717
LOCUS       AW865717                382 bp    mRNA    linear   EST 22-MAY-2000
DEFINITION  QV3-SN0021-100500-185-c01 SN0021 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW865717
VERSION     AW865717.1  GI:7999767
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=QV3-SN0021-100
            500-185-c01&t3=2000-05-10&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 15
            High quality sequence stop: 381.
FEATURES             Location/Qualifiers
     source          1. .382
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="SN0021"
                     /note="Organ: stomach_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                2.39e-18        Length:         382
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
```

```
Best Local Similarity:  97.7%          Mismatches:    1
Query Match:            97.4%          Indels:        0
DB:                     1              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x AW865717 (1-382)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 117

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        118 TCAACAACTGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 177

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        178 GTTCTTGAATAC 189

RESULT 15
AI017172/c
LOCUS       AI017172                443 bp    mRNA    linear   EST 27-AUG-1998
DEFINITION  ou28b06.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1627571 3', mRNA sequence.
ACCESSION   AI017172
VERSION     AI017172.1  GI:3231508
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 554   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 428.
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1627571"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN

Alignment Scores:
Pred. No.:              2.83e-18       Length:        443
Score:                  223.00         Matches:       43
Percent Similarity:     97.7%          Conservative:  0
Best Local Similarity:  97.7%          Mismatches:    1
Query Match:            97.4%          Indels:        0
DB:                     1              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x AI017172 (1-443)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 221

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        220 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 161

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        160 GTTCTTGAATAC 149

RESULT 16
BM749349
LOCUS       BM749349                514 bp    mRNA    linear   EST 04-MAR-2002
DEFINITION  K-EST0024500 S11SNU1 Homo sapiens cDNA clone S11SNU1-3-D10 5', mRNA
            sequence.
ACCESSION   BM749349
VERSION     BM749349.1  GI:19078967
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 514)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 3  row: D  column: 10
            High quality sequence stop: 514.
FEATURES             Location/Qualifiers
     source          1. .514
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S11SNU1-3-D10"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Lymphoblast-like"
                     /cell_line="SNU-1"
                     /lab_host="Top10F'"
                     /clone_lib="S11SNU1"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
ORIGIN

Alignment Scores:
Pred. No.:              3.35e-18       Length:        514
```

```
Score:                 223.00        Matches:       43
Percent Similarity:    97.7%         Conservative:  0
Best Local Similarity: 97.7%         Mismatches:    1
Query Match:           97.4%         Indels:        0
DB:                    3             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x BM749349 (1-514)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 138

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        139 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 198

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        199 GTTCTTGAATAC 210

RESULT 17
CD627106/c
LOCUS       CD627106                551 bp    mRNA    linear   EST 12-JAN-2004
DEFINITION  56066665J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD627106
VERSION     CD627106.1  GI:40275372
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 551)
  AUTHORS   Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
  TITLE     Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
  JOURNAL   Genomics 84 (1), 205-210 (2004)
   PUBMED   15203218
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES             Location/Qualifiers
     source          1. .551
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="FLP"
                     /note="Vector: pDrive Cloning Vector"
ORIGIN

Alignment Scores:
Pred. No.:             3.62e-18      Length:        551
Score:                 223.00        Matches:       43
Percent Similarity:    97.7%         Conservative:  0
Best Local Similarity: 97.7%         Mismatches:    1
Query Match:           97.4%         Indels:        0
DB:                    6             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x CD627106 (1-551)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 428

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        427 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 368

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        367 GTTCTTGAATAC 356

RESULT 18
CD627110/c
LOCUS       CD627110                564 bp    mRNA    linear   EST 12-JAN-2004
DEFINITION  56066681J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD627110
VERSION     CD627110.1  GI:40275376
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 564)
  AUTHORS   Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
  TITLE     Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
  JOURNAL   Genomics 84 (1), 205-210 (2004)
   PUBMED   15203218
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES             Location/Qualifiers
     source          1. .564
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="FLP"
                     /note="Vector: pDrive Cloning Vector"
ORIGIN

Alignment Scores:
Pred. No.:             3.72e-18      Length:        564
Score:                 223.00        Matches:       43
Percent Similarity:    97.7%         Conservative:  0
Best Local Similarity: 97.7%         Mismatches:    1
Query Match:           97.4%         Indels:        0
DB:                    6             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x CD627110 (1-564)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 441

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        440 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 381

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        380 GTTCTTGAATAC 369

RESULT 19
CD627109
LOCUS       CD627109                568 bp    mRNA    linear   EST 12-JAN-2004
DEFINITION  56066681H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD627109
VERSION     CD627109.1  GI:40275375
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
  TITLE     Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
  JOURNAL   Genomics 84 (1), 205-210 (2004)
   PUBMED   15203218
```

```
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES             Location/Qualifiers
     source          1. .568
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="FLP"
                     /note="Vector: pDrive Cloning Vector"
ORIGIN

Alignment Scores:
Pred. No.:              3.75e-18        Length:         568
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CD627109 (1-568)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 124

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        125 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 184

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        185 GTTCTTGAATAC 196

RESULT 20
BP229655
LOCUS       BP229655                583 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP229655 Sugano cDNA library, fetal brain Homo sapiens cDNA clone
            FBR03988, mRNA sequence.
ACCESSION   BP229655
VERSION     BP229655.1  GI:52102565
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="FBR03988"
                     /tissue_type="brain"
                     /dev_stage="fetal"
                     /clone_lib="Sugano cDNA library, fetal brain"
ORIGIN

Alignment Scores:
Pred. No.:              3.86e-18        Length:         583
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BP229655 (1-583)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 205

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        206 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 265

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        266 GTTCTTGAATAC 277

RESULT 21
BP248461
LOCUS       BP248461                583 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP248461 Sugano cDNA library, embryo kidney Homo sapiens cDNA clone
            HKR11085, mRNA sequence.
ACCESSION   BP248461
VERSION     BP248461.1  GI:52130740
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HKR11085"
                     /tissue_type="kidney"
                     /cell_line="293"
                     /dev_stage="embryo"
                     /clone_lib="Sugano cDNA library, embryo kidney"
ORIGIN

Alignment Scores:
Pred. No.:              3.86e-18        Length:         583
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BP248461 (1-583)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 241

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        242 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 301

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        302 GTTCTTGAGTAC 313

RESULT 22
CN481259
LOCUS       CN481259                 604 bp    mRNA    linear   EST 26-APR-2004
DEFINITION  hw05c03.y1 Human primary human ocular pericytes. Unamplified (hw)
            Homo sapiens cDNA clone hw05c03 5', mRNA sequence.
ACCESSION   CN481259
VERSION     CN481259.1  GI:46562763
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 604)
  AUTHORS   Tsai,J.Y. and Wistow,G.
  TITLE     Expressed sequence tag analysis of cultured primary human ocular
            pericytes
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 05  row: c  column: 03
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..604
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="hw05c03"
                     /cell_type="pericytes"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human primary human ocular pericytes.
                     Unamplified (hw)"
                     /note="Organ: Eye; Vector: pSport1; RNA was extracted from
                     primary human pericytes in culture. A directionally cloned
                     cDNA library in the pSPORT1 vector (Invitrogen) was
                     constructed at Bioserve Biotechnology (Laurel MD)
                     essentially following the protocols of the SuperScript
                     Plasmid System full details of which are contained in the
                     manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. cDNA was
                     cloned in Not I/Sal I sites. EST analysis was performed at
                     the NIH Intramural Sequencing Center (NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              4.02e-18        Length:         604
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CN481259 (1-604)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 118
```

```
Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        119 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 178

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        179 GTTCTTGAATAC 190

RESULT 23
CV024153
LOCUS       CV024153                 630 bp    mRNA    linear   EST 20-AUG-2004
DEFINITION  1345 Full Length cDNA from the Mammalian Gene Collection Homo
            sapiens cDNA 5' similar to BC021712, mRNA sequence.
ACCESSION   CV024153
VERSION     CV024153.1  GI:51481981
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
            Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
            Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
            Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
            Vandenhaute,J., Cusick,M.E., Albala,J.S., Hill,D.E. and Vidal,M.
  TITLE     Human ORFeome Version 1.1: a Platform for Reverse Proteomics
  JOURNAL   Genome Res. (2004) In press
COMMENT     Contact: Vidal M
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 5739
            Email: Marc_Vidal@dfci.harvard.edu
            ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
            results from a PCR reaction using an MGC full-length cDNA as
            template DNA and ORF specific primers
            PCR PRimers
            FORWARD: ATGGGAGACCCGGGGTCG
            BACKWARD: TAAGAGGTGGGAAAAAGTTTTCTTTTG
            Insert Length: 630   Std Error: 25.00
            Plate: 11025  row: 12  column: D
            Seq primer: ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
            High quality sequence start: 99
            High quality sequence stop: 629
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..630
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="mixed"
                     /clone_lib="Full Length cDNA from the Mammalian Gene
                     Collection"
                     /note="Vector: mixed; The ORFs were PCR amplified from the
                     MGC (Mammalian Gene Collection) as of April 2004 and
                     cloned by recombinational Gateway cloning into pDONR223
                     Donor vector. Reference : MGC (Mammalian Gene Collection)
                     Program Team, Generation and Initial Analysis of more than
                     15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
                     2002, 99(26), 16899-16903"
ORIGIN

Alignment Scores:
Pred. No.:              4.22e-18        Length:         630
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     7               Gaps:           0
```

```
US-10-757-745-6_COPY_2_45 (1-44) x CV024153 (1-630)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 24
DN991049
LOCUS       DN991049                643 bp    mRNA    linear   EST 17-MAY-2005
DEFINITION  TC111228 Human adult whole brain, large insert, pCMV expression
            library Homo sapiens cDNA clone TC111228 5' similar to Homo sapiens
            zinc finger protein 451 (ZNF451), mRNA sequence.
ACCESSION   DN991049
VERSION     DN991049.1  GI:66250880
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 643)
  AUTHORS   Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L.,
            Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M.,
            Zhang,X., Jay,G. and He,W.
  TITLE     High-throughput cloning of full-length human cDNAs directly from
            cDNA libraries optimized for large and rare transcripts
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Kovacs, KF
            High Throughput cDNA Cloning
            OriGene Technologies, Inc. ( www.origene.com )
            6 Taft Court, Suite 100, Rockville, MD 20850, USA
            Tel: 301 340 3188
            Fax: 301 340 8606
            Email: cDNA@origene.com
            This EST submission is part of an on-going human full-length
            cloning project at OriGene Technologies, Inc.
            Please contact OriGene for access.
            OriGene Technologies, Inc.
            6 Taft Ct. Suite 100
            Rockville, MD 20850
            Tel: (301) 340-3188
            http://www.origene.com
            Seq primer: pCMV6 5prime forward vector primer, OriGene
            Technologies Inc.
FEATURES             Location/Qualifiers
     source          1. .643
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TC111228"
                     /tissue_type="Whole brain"
                     /clone_lib="Human adult whole brain, large insert, pCMV
                     expression library"
                     /note="Organ: Brain; Vector: pCMV6-XL5; Site_1: EcoR1;
                     Site_2: Xho1/Sal1 compatible end ligatio; Oligo-dT primed
                     reverse transcription optimized for large and GC rich mRNA
                     transcripts, cDNA size selection, optimized ligation for
                     large inserts into mammalian expression vector, random
                     clones selected for end sequence verification of
                     full-length genes"
ORIGIN

Alignment Scores:
Pred. No.:              4.32e-18        Length:         643
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x DN991049 (1-643)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 114

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        115 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 174

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        175 GTTCTTGAATAC 186

RESULT 25
BG722501
LOCUS       BG722501                671 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602693752F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4825788 5',
            mRNA sequence.
ACCESSION   BG722501
VERSION     BG722501.1  GI:14001688
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 671)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10739  row: h  column: 13
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .671
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4825788"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              4.53e-18        Length:         671
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
```

```
Query Match:            97.4%          Indels:        0
DB:                     2              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x BG722501 (1-671)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 115

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 175

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        176 GTTCTTGAATAC 187

RESULT 26
CN314844
LOCUS       CN314844                693 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000424984291 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN314844
VERSION     CN314844.1  GI:47331258
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 693)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 693   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .693
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, cell lines H1, H7, and
                     H9"
                     /clone_lib="GRN_ES"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from undifferentiated hES cell lines H1 (p32), H7 (p29),
                     and H9 (p26) maintained in feeder-free conditions"
ORIGIN

Alignment Scores:
Pred. No.:              4.7e-18        Length:        693
Score:                  223.00         Matches:       43
Percent Similarity:     97.7%          Conservative:  0
Best Local Similarity:  97.7%          Mismatches:    1
Query Match:            97.4%          Indels:        0
DB:                     7              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x CN314844 (1-693)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 263

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 323

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        324 GTTCTTGAATAC 335

RESULT 27
CN314843
LOCUS       CN314843                694 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000424984285 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN314843
VERSION     CN314843.1  GI:47331257
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 694)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 694   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .694
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, cell lines H1, H7, and
                     H9"
                     /clone_lib="GRN_ES"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from undifferentiated hES cell lines H1 (p32), H7 (p29),
                     and H9 (p26) maintained in feeder-free conditions"
ORIGIN

Alignment Scores:
Pred. No.:              4.71e-18       Length:        694
Score:                  223.00         Matches:       43
Percent Similarity:     97.7%          Conservative:  0
Best Local Similarity:  97.7%          Mismatches:    1
Query Match:            97.4%          Indels:        0
DB:                     7              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x CN314843 (1-694)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 261

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 321

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        322 GTTCTTGAATAC 333

RESULT 28
BI559958
LOCUS       BI559958                695 bp    mRNA    linear   EST 05-SEP-2001
```

```
DEFINITION  603253421F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5295845 5',
            mRNA sequence.
ACCESSION   BI559958
VERSION     BI559958.1  GI:15447272
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 695)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11748  row: j  column: 06
            High quality sequence stop: 693.
FEATURES             Location/Qualifiers
     source          1. .695
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5295845"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN
```

```
Alignment Scores:
Pred. No.:              4.72e-18        Length:         695
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BI559958 (1-695)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 134

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        135 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 194

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        195 GTTCTTGAATAC 206
```

```
RESULT 29
BG701977
LOCUS       BG701977                 749 bp    mRNA    linear   EST 07-MAY-2001
DEFINITION  602683351F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4815756 5',
            mRNA sequence.
ACCESSION   BG701977
VERSION     BG701977.1  GI:13972858
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10713  row: f  column: 13
            High quality sequence stop: 709.
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4815756"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_95"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.5 kb and normalized to ROT 5. This is a
                     primary library enriched for  full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health).  Note: this
                     is a NIH_MGC Library."
ORIGIN
```

```
Alignment Scores:
Pred. No.:              5.13e-18        Length:         749
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     2               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BG701977 (1-749)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 208

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        209 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 268

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        269 GTTCTTGAATAC 280
```

```
RESULT 30
BE613457
LOCUS       BE613457                 752 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601504223F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3906220 5',
            mRNA sequence.
ACCESSION   BE613457
```

```
VERSION     BE613457.1  GI:9895054
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9714  row: m  column: 05
            High quality sequence start: 3
            High quality sequence stop: 636.
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3906220"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

Alignment Scores:
Pred. No.:                5.16e-18        Length:          752
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       2               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BE613457 (1-752)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 91

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 151

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        152 GTTCTTGAATAC 163

RESULT 31
CN314848
LOCUS       CN314848                760 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000531863180 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN314848
VERSION     CN314848.1  GI:47331262
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
  PUBMED    15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 760   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .760
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, embryoid bodies
                     derived from H1, H7 and H9 cells"
                     /clone_lib="GRN_EB"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from embryoid body outgrowths derived from hES cell lines
                     H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                     conditions."
ORIGIN

Alignment Scores:
Pred. No.:                5.22e-18        Length:          760
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x CN314848 (1-760)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 80

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 140

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        141 GTTCTTGAATAC 152

RESULT 32
CT002116
LOCUS       CT002116                829 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT002116 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016I2032 5',
            mRNA sequence.
ACCESSION   CT002116
VERSION     CT002116.1  GI:68295999
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 829)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016I2032.
```

```
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016I2032
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .829
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp9016I2032"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9016"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

Alignment Scores:
Pred. No.:                5.76e-18        Length:         829
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CT002116 (1-829)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 114

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 174

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        175 GTTCTTGAATAC 186

RESULT 33
CR767845
LOCUS       CR767845               599 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp469C2210_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469C2210 5', mRNA sequence.
ACCESSION   CR767845
VERSION     CR767845.1  GI:52610372
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,
            Deutschenbaur,S., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix
            (Martinsried/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp469C2210) is available at
            the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
            Berlin, Germany. Please contact RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469C2210
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES              Location/Qualifiers
     source           1. .599
                      /organism="Pongo pygmaeus"
                      /mol_type="mRNA"
                      /db_xref="taxon:9600"
                      /clone="DKFZp469C2210"
                      /tissue_type="kidney"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="469 (synonym: pkid1)"
                      /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:                5.34e-18        Length:         599
Score:                    222.00          Matches:        42
Percent Similarity:       97.7%           Conservative:   1
Best Local Similarity:    95.5%           Mismatches:     1
Query Match:              96.9%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CR767845 (1-599)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        147 GGAGACCCGGGGTCGGAGGTAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 206

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 266

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        267 GTTCTTGAATAC 278

RESULT 34
BG777259
LOCUS       BG777259               691 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602664490F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4809575 5',
            mRNA sequence.
ACCESSION   BG777259
VERSION     BG777259.1  GI:14047576
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM1664 row: d column: 24
High quality sequence stop: 691.

FEATURES Location/Qualifiers
source 1. .691
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4809575"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_59"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	6.28e-18	Length:	691
Score:	222.00	Matches:	43
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	96.9%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BG777259 (1-691)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         38 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 97

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 157

Qy         41 ValLeuGlu 43
              |||||||||
Db        158 GTTCTTGAA 166
```

RESULT 35
CR858845
LOCUS CR858845 5087 bp mRNA linear HTC 12-NOV-2004
DEFINITION Pongo pygmaeus mRNA; cDNA DKFZp469C2210 (from clone DKFZp469C2210).
ACCESSION CR858845
VERSION CR858845.1 GI:55728610
KEYWORDS HTC.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE 1 (bases 1 to 5087)
AUTHORS Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
CONSRTM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp469C2210) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469C2210
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES Location/Qualifiers
source 1. .5087
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469C2210"
/tissue_type="kidney"
/clone_lib="469 (synonym: pkid1). Vector pSport1_Sfi; host DH10B; sites SfilA + SfilB"
/dev_stage="adult"
/note="zinc finger protein 451 (Homo sapiens)"
gene 1. .5087
/gene="DKFZp469C2210"
CDS join(144. .1484,1483. .3330)
/gene="DKFZp469C2210"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH91046.1"
/db_xref="GI:55728611"
/translation="MGDPGSEVIESVPPAGPEASESTTDENEDDIQFVSEGPLRPVLE
YIDLVSSDDEEPSTSYTDENVKRKDHIDHQKDKVALTLARLARHVEVGKQQKEEKNRA
FREKVDFQHAHGLQELEFIRGHSDTEAARLCVDQWLKMPGLKTGTINCGTKSSFRRGG
QTRVSGKPILCPIMHCNKEFDNGHLLLGHLKRFDHSPCDPTITLHGPPFSSFACVVCY
KKFVTQQQYRDHLFDKEATDDGHNNNLLPQIIQCFACPNCFLLFSSKEECSKHMSGKN
HFHQSFRLGDNKGIAHPISFPSFAKKLLISLCKDVPFQVKCVACHQTLRSHMELTAHF
RVRCRNAGPVAVAEKSITQVAEKFILRGYCPDCDQVFVDETSTQNHKQNSGHKVRVIN
SVEESVLLYCHSSEGNKDPSSDLHLLLDQSKFSSLKRTMSMKESSSLECIAIPKKKKM
NLKDKSLEGVACVQKEKSVVKTWFCECNQRFPSEDAVEKHVFSANTMGYKCVVCGKVC
DDSGVIRLHMSRIHGGAHLNNFLFWCRTCKKELTRKDTIMAHVTEFHNGHRYFYEMDE
VEGETLPSSSTTLGNLTANKPSSAITIIDHSPANSSPRGKWQCRICEDMFDSQEYVKQ
HCMSLASHKFHRYSCAHCRKPFHKIETLYRHCQDEHDNEIKIKYFCGLCDLIFNVEEA
FLSHYEEHHSIDYVFVSEKTETSIKTEDDFPVIETSNQLTCGCRESYICKVNRKEDYS
RCLQIMLDKGKLWFRCSLCSATAQNLTDMNTHIHQVHKEKSDEEEQQYVIKCGTCTKA
FHDPESAQQHFHRKHCFLQKLSVAHFGCEKSNLYKFTASASHTERKLKQAINYSKNLD
MEKGVENDLSYQNIEEEIVELPDLDYLRTMTHIVFVDFDNWSNFFGHLPGHLNQGTFI
WGFQGGNTNWKPPLNCKIYNYLNRIGCFFLHPRCSKRKDAADFAICMHAGRLDEQLPK
QIPPTILSGDQGFLELENQFKKTQRPAHILNPHHLEGDMMCALLNSISDTTKECDSDD
NMGAKNTSIGEEFISTEDVELEEAIRRSLEEM"

ORIGIN

Alignment Scores:

Pred. No.:	6.04e-17	Length:	5087
Score:	222.00	Matches:	42
Percent Similarity:	97.7%	Conservative:	1
Best Local Similarity:	95.5%	Mismatches:	1
Query Match:	96.9%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x CR858845 (1-5087)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        147 GGAGACCCGGGGTCGGAGGTAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 206

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        207 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 266

Qy         41 ValLeuGluTyr 44
              |||||||||||
Db        267 GTTCTTGAATAC 278
```

RESULT 36
AL703808
LOCUS AL703808 601 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686K1327_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone

DKFZp686K1327 5', mRNA sequence.
ACCESSION AL703808
VERSION AL703808.1 GI:19687163
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 601)
AUTHORS Ansorge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
TITLE EST (Ansorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
No s1 sequence available.
This clone (DKFZp686K1327) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES Location/Qualifiers
source 1. .601
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686K1327"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"

ORIGIN

Alignment Scores:

Pred. No.:	1.29e-17	Length:	601
Score:	219.00	Matches:	42
Percent Similarity:	97.7%	Conservative:	1
Best Local Similarity:	95.5%	Mismatches:	1
Query Match:	95.6%	Indels:	0
DB:	1	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x AL703808 (1-601)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 100

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||||||||||||||||||||||||||||:::||||||||||||||||||    ||||||
Db        101 TCAACAACGGATGAAAATGAAGACGACATTCGGTTTGTCAGTGAAGGACCATTACGACCT 160

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        161 GTTCTTGAATAC 172
```

RESULT 37
BY300743
LOCUS BY300743 362 bp mRNA linear EST 11-DEC-2002
DEFINITION BY300743 RIKEN full-length enriched, 14.5 days embryo RP+/+ Rathke's pouches Mus musculus cDNA clone K720012A01 5', mRNA sequence.
ACCESSION BY300743
VERSION BY300743.1 GI:26491080
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 362)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Michelle Brinkmeier and Sally Camper (Dept. Human Genetics University of Michigan Medical School 4301 MSRB 3 1500 W. Medical Center Dr. Ann Arbor,MI 48109-0638 USA) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .362
/organism="Mus musculus"

```
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="K720012A01"
                     /tissue_type="Rathke's pouches"
                     /dev_stage="14.5 days embryo RP+/+"
                     /clone_lib="RIKEN full-length enriched, 14.5 days embryo
                     RP+/+ Rathke's pouches"
ORIGIN

Alignment Scores:
Pred. No.:                9.7e-18         Length:         362
Score:                    218.00          Matches:        42
Percent Similarity:       95.5%           Conservative:   0
Best Local Similarity:    95.5%           Mismatches:     2
Query Match:              95.2%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BY300743 (1-362)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 151

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 211

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        212 GTTTTGGAATAC 223

RESULT 38
BY230029
LOCUS       BY230029                366 bp    mRNA    linear   EST 10-DEC-2002
DEFINITION  BY230029 RIKEN full-length enriched, adult inner ear Mus musculus
            cDNA clone F930012B05 5', mRNA sequence.
ACCESSION   BY230029
VERSION     BY230029.1  GI:26411139
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 366)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
            Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
            Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
            Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
            Hayashizaki,Y. Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Tissues were provided by Kirk W. Beisel ( Boys Town National
            Research Hospital 555 North 30th Street Omaha,NE 68131 USA ) whose
            assistance we gratefully acknowledge. Please visit our web site
            (http://genome.gsc.riken.go.jp) for further details.
FEATURES             Location/Qualifiers
     source          1. .366
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="F930012B05"
                     /tissue_type="inner ear"
                     /dev_stage="adult"
                     /clone_lib="RIKEN full-length enriched, adult inner ear"
ORIGIN

Alignment Scores:
Pred. No.:                9.83e-18        Length:         366
Score:                    218.00          Matches:        42
Percent Similarity:       95.5%           Conservative:   0
Best Local Similarity:    95.5%           Mismatches:     2
Query Match:              95.2%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BY230029 (1-366)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 134

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 194

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        195 GTTTTGGAATAC 206

RESULT 39
BY307103
```

LOCUS BY307103 378 bp mRNA linear EST 11-DEC-2002
DEFINITION BY307103 RIKEN full-length enriched, 12.5 days embryo Rathke's pouches Mus musculus cDNA clone K920013E19 5', mRNA sequence.
ACCESSION BY307103
VERSION BY307103.1 GI:26497440
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 378)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Michelle Brinkmeier and Sally Camper (Dept. Human Genetics University of Michigan Medical School 4301 MSRB 3 1500 W. Medical Center Dr. Ann Arbor,MI 48109-0638 USA) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .378
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"
/db_xref="taxon:10090"
/clone="K920013E19"
/tissue_type="Rathke's pouches"
/dev_stage="12.5 days embryo"
/clone_lib="RIKEN full-length enriched, 12.5 days embryo Rathke's pouches"
ORIGIN

Alignment Scores:

Pred. No.:	1.02e-17	Length:	378
Score:	218.00	Matches:	42
Percent Similarity:	95.5%	Conservative:	0
Best Local Similarity:	95.5%	Mismatches:	2
Query Match:	95.2%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BY307103 (1-378)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         80 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 139

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 199

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        200 GTTTTGGAATAC 211
```

RESULT 40
BY094255
LOCUS BY094255 388 bp mRNA linear EST 07-DEC-2002
DEFINITION BY094255 RIKEN full-length enriched, pooled tissues, adult spleen, etc. Mus musculus cDNA clone K630099K07 5', mRNA sequence.
ACCESSION BY094255
VERSION BY094255.1 GI:26204872
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 388)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,

Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .388
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="K630099K07"
/clone_lib="RIKEN full-length enriched, pooled tissues, adult spleen, etc."
/note="(dev_stage=adult,tissue_type=spleen,sex=male), (dev_stage=adult,tissue_type=kidney,sex=male), (dev_stage=adult,tissue_type=testis,sex=male), (dev_stage=adult,tissue_type=thymus,sex=male), (dev_stage=adult,tissue_type=heart,sex=male), (dev_stage=adult,tissue_type=colon,sex=male), (dev_stage=adult,tissue_type=stomach,sex=male), (dev_stage=adult,tissue_type=liver,sex=male), (dev_stage=13 days embryo,tissue_type=whole body,sex=mix),(dev_stage=14 days embryo,tissue_type=whole body,sex=mix),(dev_stage=16 days embryo,tissue_type=whole body,sex=mix),(dev_stage=17 days embryo,tissue_type=whole body,sex=mix),(dev_stage=15 days pregnant, adult,tissue_type=amnion,sex=female),(dev_stage=10 days neonate,tissue_type=brain,sex=mix),(dev_stage=10 days neonate,tissue_type=thymus,sex=mix),(dev_stage=10 days neonate,tissue_type=heart,sex=mix)"

ORIGIN

Alignment Scores:

Pred. No.:	1.05e-17	Length:	388
Score:	218.00	Matches:	42
Percent Similarity:	95.5%	Conservative:	0
Best Local Similarity:	95.5%	Mismatches:	2
Query Match:	95.2%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BY094255 (1-388)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        102 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 161

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 221

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        222 GTTCTGGAATAC 233
```

RESULT 41
BY024214
LOCUS BY024214 392 bp mRNA linear EST 06-DEC-2002
DEFINITION BY024214 RIKEN full-length enriched, mammary gland RCB-0527 Jyg-MC(B) cDNA Mus musculus cDNA clone G930021N16 5', mRNA sequence.
ACCESSION BY024214
VERSION BY024214.1 GI:26129657
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 392)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222

Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

```
FEATURES             Location/Qualifiers
     source          1. .392
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="G930021N16"
                     /tissue_type="mammary gland"
                     /cell_line="RCB-0527 Jyg-MC(B)"
                     /clone_lib="RIKEN full-length enriched, mammary gland
                     RCB-0527 Jyg-MC(B) cDNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.06e-17       Length:          392
Score:                    218.00         Matches:         42
Percent Similarity:       95.5%          Conservative:    0
Best Local Similarity:    95.5%          Mismatches:      2
Query Match:              95.2%          Indels:          0
DB:                       5              Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BY024214 (1-392)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         63 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 122

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        123 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 182

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        183 GTTTTGGAATAC 194

RESULT 42
BY019311
LOCUS       BY019311                 398 bp    mRNA    linear   EST 06-DEC-2002
DEFINITION  BY019311 RIKEN full-length enriched, mammary gland RCB-0526
            Jyg-MC(A) cDNA Mus musculus cDNA clone G830034K22 5', mRNA
            sequence.
ACCESSION   BY019311
VERSION     BY019311.1  GI:26079560
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 398)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
    TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
            Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
            Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
            Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
            Hayashizaki,Y. Direct Submission
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .398
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="G830034K22"
                     /tissue_type="mammary gland"
                     /cell_line="RCB-0526 Jyg-MC(A)"
```

/clone_lib="RIKEN full-length enriched, mammary gland RCB-0526 Jyg-MC(A) cDNA"

ORIGIN

Alignment Scores:

Pred. No.:	1.08e-17	Length:	398
Score:	218.00	Matches:	42
Percent Similarity:	95.5%	Conservative:	0
Best Local Similarity:	95.5%	Mismatches:	2
Query Match:	95.2%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BY019311 (1-398)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 130

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        131 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 190

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        191 GTTTTGGAATAC 202
```

RESULT 43
BI463017
LOCUS BI463017 620 bp mRNA linear EST 21-AUG-2001
DEFINITION 603204540F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5270624 5', mRNA sequence.
ACCESSION BI463017
VERSION BI463017.1 GI:15253673
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 620)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11682 row: o column: 09
High quality sequence stop: 620.

FEATURES Location/Qualifiers
source 1..620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5270624"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	1.79e-17	Length:	620
Score:	218.00	Matches:	42
Percent Similarity:	95.5%	Conservative:	0
Best Local Similarity:	95.5%	Mismatches:	2
Query Match:	95.2%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BI463017 (1-620)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 284

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||  ||||||||||||||||||||||  ||||||
Db        285 TCAACAACGGATGAAAATGAAGACGACACTCAGTTTGTCAGTGAAGGACCATTACGACCT 344

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        345 GTTCTTGAATAC 356
```

RESULT 44
BB657575
LOCUS BB657575 644 bp mRNA linear EST 26-OCT-2001
DEFINITION BB657575 RIKEN full-length enriched, 12 days embryo eyeball Mus musculus cDNA clone D230008I20 5', mRNA sequence.
ACCESSION BB657575
VERSION BB657575.1 GI:16491401
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 644)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,

Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.

FEATURES Location/Qualifiers
source 1. .644
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D230008I20"
/tissue_type="eyeball"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 12 days embryo eyeball"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Alignment Scores:

Pred. No.:	1.87e-17	Length:	644
Score:	218.00	Matches:	42
Percent Similarity:	95.5%	Conservative:	0
Best Local Similarity:	95.5%	Mismatches:	2
Query Match:	95.2%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BB657575 (1-644)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        206 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 265

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        266 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 325

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        326 GTTTTGGAATAC 337
```

RESULT 45
BB626942
LOCUS BB626942 653 bp mRNA linear EST 31-AUG-2001
DEFINITION BB626942 RIKEN full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck Mus musculus cDNA clone 9430043D14 5', mRNA sequence.
ACCESSION BB626942
VERSION BB626942.1 GI:15398976
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 653)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ishii,Y. and Hayashizaki,Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, L72-L86 (2001
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.

FEATURES Location/Qualifiers
source 1. .653
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9430043D14"
/tissue_type="embryonic body between diaphragm region and neck"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI."

ORIGIN

Alignment Scores:

Pred. No.:	1.9e-17	Length:	653
Score:	218.00	Matches:	42

```
Percent Similarity:     95.5%          Conservative:  0
Best Local Similarity:  95.5%          Mismatches:    2
Query Match:            95.2%          Indels:        0
DB:                     2              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x BB626942 (1-653)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        192 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 251

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        252 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 311

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        312 GTTTTGGAATAC 323

RESULT 46
BB616798
LOCUS       BB616798                671 bp    mRNA    linear   EST 31-AUG-2001
DEFINITION  BB616798 RIKEN full-length enriched, adult male testis Mus musculus
            cDNA clone 4932441A15 5', mRNA sequence.
ACCESSION   BB616798
VERSION     BB616798.1  GI:15396205
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 671)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A.,
            Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K.,
            Arakawa,T., Ishii,Y. and Hayashizaki,Y.
             Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
            Func. Genomics 2 pre, L72-L86 (2001
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .671
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="4932441A15"
                     /sex="male"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="SOLR"
                     /clone_lib="RIKEN full-length enriched, adult male testis"
                     /note="Site_1: XhoI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGAGCGGCCGCAATTAATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC
                     3']. cDNA was cloned into the XhoI and BamHI sites. "
ORIGIN

Alignment Scores:
Pred. No.:              1.95e-17       Length:        671
Score:                  218.00         Matches:       42
Percent Similarity:     95.5%          Conservative:  0
Best Local Similarity:  95.5%          Mismatches:    2
Query Match:            95.2%          Indels:        0
DB:                     2              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x BB616798 (1-671)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db         74 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 133

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        134 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 193

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        194 GTTTTGGAATAC 205

RESULT 47
BY715342
LOCUS       BY715342                715 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  BY715342 RIKEN full-length enriched, adult male testis Mus musculus
            cDNA clone 4930515K21 5', mRNA sequence.
ACCESSION   BY715342
VERSION     BY715342.1  GI:27128459
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 715)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
```

Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .715
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4930515K21"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, adult male testis"
/note="Site_1: XhoI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGCGGCCGCAATTAATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. "

ORIGIN

Alignment Scores:

Pred. No.:	2.1e-17	Length:	715
Score:	218.00	Matches:	42
Percent Similarity:	95.5%	Conservative:	0
Best Local Similarity:	95.5%	Mismatches:	2
Query Match:	95.2%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BY715342 (1-715)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        216 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 275

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 335

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        336 GTTTTGGAATAC 347
```

RESULT 48
BQ952426
LOCUS BQ952426 914 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8750083 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6331597 5', mRNA sequence.
ACCESSION BQ952426
VERSION BQ952426.1 GI:22367904
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 914)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13787 row: n column: 14
High quality sequence stop: 731.

FEATURES Location/Qualifiers
source 1. .914
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6331597"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_130"
/note="Organ: otocysts; Vector: pCMV-SPORT6.1; Site_1:

EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	2.78e-17	Length:	914
Score:	218.00	Matches:	42
Percent Similarity:	95.5%	Conservative:	0
Best Local Similarity:	95.5%	Mismatches:	2
Query Match:	95.2%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BQ952426 (1-914)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 89

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db         90 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 149

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        150 GTTTTGGAATAC 161
```

RESULT 49
BQ886508
LOCUS BQ886508 929 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT_8753185 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6394545 5', mRNA sequence.
ACCESSION BQ886508
VERSION BQ886508.1 GI:22278522
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 929)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13887 row: m column: 10
High quality sequence stop: 663.

FEATURES Location/Qualifiers
source 1. .929
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6394545"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_130"
/note="Organ: oocysts; Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	2.83e-17	Length:	929
Score:	218.00	Matches:	42
Percent Similarity:	95.5%	Conservative:	0
Best Local Similarity:	95.5%	Mismatches:	2
Query Match:	95.2%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BQ886508 (1-929)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 216

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        217 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 276

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        277 GTTTTGGAATAC 288
```

RESULT 50
AK015797
LOCUS AK015797 1750 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4930515K21 product:hypothetical Zinc finger, C2H2 type containing protein, full insert sequence.
ACCESSION AK015797
VERSION AK015797.1 GI:12854273
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1750)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,

Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised longer than 7 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES Location/Qualifiers

source 1. .1750
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4930515K21"
/db_xref="taxon:10090"
/clone="4930515K21"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

CDS 213. .>1748
/note="unnamed protein product; hypothetical Zinc finger, C2H2 type containing protein (InterPro|IPR000822, evidence: InterPro)
putative"
/codon_start=1
/protein_id="BAB29981.2"
/db_xref="GI:26381581"
/translation="MGDPGPEIIESVPPAGPEASESTTDENEDDIQFVSEGPLRPVLE
YIDLVSSDDEEPSTSHSDRMPESKVPSSENHRPEMCSSCSVPLPIGDSSSSSGSCANS
PQRIVSQPSSVENPLENQKNDHNNSDVSVSEKETLKPQSNQTLPSASLLGSQESLAFS
EVKEDLPGTMSSASQSGQDAILYLQTQVAEMSRVIRDLQSRSCFRFHHSRPSESSSVP
WDISSSKEDSLSTVEEESDCKSPSADDKGQPPDPMQSSFTGLLKRMEQRGVIKRVTFQ
SEAETCEGKPDWMTSKKRLVPPLHPLLRIATTEVFKDPADCHPSSFMGHRVYPVAKDT
SPFQPNPPAEGPIIDVLEHSKRGNTASPLDSTSKEMEVMGCRFYHAASIAARAASYMA
YMTQYQRKLWEDMEDLVHDPEFDRGKARCIISDGMDAGLWQLCTTRDIMDSVVRVMAM
AVDYRRQAWLRLTSLTKKTQEKILHLPFDGTSLFGQDVNAIVAEDNSIKENEYKDHNK
YYTQ"

ORIGIN

Alignment Scores:

Pred. No.:	5.8e-17	Length:	1750
Score:	218.00	Matches:	42
Percent Similarity:	95.5%	Conservative:	0
Best Local Similarity:	95.5%	Mismatches:	2
Query Match:	95.2%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x AK015797 (1-1750)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        216 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 275

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 335

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        336 GTTTTGGAATAC 347
```

RESULT 51
AK030088
LOCUS AK030088 3651 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4932441A15 product:hypothetical Zinc finger, C2H2 type containing protein, full insert sequence.
ACCESSION AK030088
VERSION AK030088.1 GI:26326068
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636

REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159

REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861

REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3651)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,

```
          Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
          Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
          Physical and Chemical Research (RIKEN), Laboratory for Genome
          Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
          RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
          Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
          URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
          Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
          Encyclopedia Project of Genome Exploration Research Group in Riken
          Genomic Sciences Center and Genome Science Laboratory in RIKEN.
          Division of Experimental Animal Research in Riken contributed to
          prepare mouse tissues.
          Please visit our web site for further details.
          URL:http://genome.gsc.riken.jp/
          URL:http://fantom.gsc.riken.jp/.
FEATURES             Location/Qualifiers
     source          1. .3651
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4932441A15"
                     /db_xref="taxon:10090"
                     /clone="4932441A15"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             71. .3241
                     /note="unnamed protein product; hypothetical Zinc finger,
                     C2H2 type containing protein (InterPro|IPR000822,
                     evidence: InterPro)
                     putative"
                     /codon_start=1
                     /protein_id="BAC26778.1"
                     /db_xref="GI:26326069"
                     /translation="MGDPGPEIIESVPPAGPEASESTTDENEDDIQFVSEGPLRPVLE
                     YIDLVSSDDEEPSTSHSDENFKCKDYIDHQKDKVALTLARLARHVEVEKQQKEEKNRA
                     FREKIDFQHAHGLQELEFIQGHSETEAARQCVDQWLKMPGLRTNAANSGTKRSFQRGG
                     RMWRSEKPILCPIMHCNKEFDNGHLLLGHLKRFDHSPCDPTITLHGPLANSFACAVCY
                     EHFVTQQQYKDHLLSRTAAADGHSNSLLPQIIQCYACPQCFLLFSTKDECLKHMSTKN
                     HFHQSFKLSDNKGTARPISFPSFAKKRLVSLCKDVPFQVKCVACHQTLRSHMELTAHF
                     RVRCQNAGPVAIAEKSITQVAKEFIVRGYCSDCNQVFMDVASTQSHKNSGHKITLANS
                     VEESVLLYCHISEGSRPPCDLHLFSQPKISSLKRILSVKESSAEDCIVPTKKVNLGVE
                     SLGGATRVQRQSPAVTAWFCECRRQFPSEEAVEKHVFSANTMCYKCVVCGKVCEDSGV
                     MRLHMSRFHGGAHLNNFLFWCRTCKKELVKKDAIMAHITEFHSGHRYFYEMDEVEEEE
                     EEAMPSSSVESHLNTDKPPSPIAVVDHCPANSPPRGRWQCRICEDMFESQECVKQHCM
                     SLTSHRFHRYSCAHCRKTFHKVETLYRHCQDEHDSEIMMKYFCGLCDLIFNKEEEFLS
                     HYKEHHSIDYVFVSEKTKTSIKTEGDFKIVETSSLLSCGCHESYMCKINRKEDYDRCL
                     PVLLEKGRLWFRCSSCSATAQNVTDINTHVCQVHRKEKSEEEQQYVIKCGICTKAFQN
                     TESAQQHFHRKHAALQKPTATPGGANRSSTCQLAASASHAEKNLKQPSSQKHSDVEKG
                     AEHDVRCQNIEEEVELPDVDYLRTMTHIVFVDFDNWSNFFGHLPGHLNQGTFIWGFQG
                     GNTNWKPPLSCKVYNYLSRIGCFFLHPRCSKRKDAADFAICMHAGRLDEQLPKQIPFT
                     ILSGDQGFLELENQFKKTQRPAHILNPHHLEGDMMCALLNSISDTTKECDSDDSSGMK
                     GSPAEELRATEDVELEEAIRRSLEEM"
     polyA_signal    3634. .3639
                     /note="putative"
     polyA_site      3651
                     /note="putative"
ORIGIN

Alignment Scores:
Pred. No.:                1.34e-16        Length:          3651
Score:                    218.00          Matches:         42
Percent Similarity:       95.5%           Conservative:    0
Best Local Similarity:    95.5%           Mismatches:      2
Query Match:              95.2%           Indels:          0
DB:                       4               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AK030088 (1-3651)
```

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 133

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 193

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        194 GTTTTGGAATAC 205

RESULT 52
AW118103/c
LOCUS       AW118103                 460 bp    mRNA    linear   EST 20-OCT-1999
DEFINITION  xe03d02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2606019 3', mRNA sequence.
ACCESSION   AW118103
VERSION     AW118103.1  GI:6086687
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2606019"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN

Alignment Scores:
Pred. No.:                1.71e-17        Length:          460
Score:                    217.00          Matches:         42
Percent Similarity:       95.5%           Conservative:    0
Best Local Similarity:    95.5%           Mismatches:      2
Query Match:              94.8%           Indels:          0
DB:                       1               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x AW118103 (1-460)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATTTGAG 221

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        220 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 161

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        160 GTTCTTGAATAC 149

RESULT 53
CD627108/c
LOCUS       CD627108                538 bp    mRNA    linear   EST 12-JAN-2004
DEFINITION  56066673J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD627108
VERSION     CD627108.1  GI:40275374
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
  TITLE     Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
  JOURNAL   Genomics 84 (1), 205-210 (2004)
   PUBMED   15203218
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="FLP"
                     /note="Vector: pDrive Cloning Vector"
ORIGIN

Alignment Scores:
Pred. No.:                2.04e-17       Length:         538
Score:                    217.00         Matches:        42
Percent Similarity:       95.5%          Conservative:   0
Best Local Similarity:    95.5%          Mismatches:     2
Query Match:              94.8%          Indels:         0
DB:                       6              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CD627108 (1-538)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        474 GGAGACCCGGGGTCGGAGATAATAGAATTTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 415

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        414 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 355

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        354 GTTCTTGAATAC 343

RESULT 54
CD627107
LOCUS       CD627107                567 bp    mRNA    linear   EST 12-JAN-2004
DEFINITION  56066673H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD627107
VERSION     CD627107.1  GI:40275373
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 567)
  AUTHORS   Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
  TITLE     Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
  JOURNAL   Genomics 84 (1), 205-210 (2004)
   PUBMED   15203218
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES             Location/Qualifiers
     source          1. .567
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="FLP"
                     /note="Vector: pDrive Cloning Vector"
ORIGIN

Alignment Scores:
Pred. No.:                2.16e-17       Length:         567
Score:                    217.00         Matches:        42
Percent Similarity:       95.5%          Conservative:   0
Best Local Similarity:    95.5%          Mismatches:     2
Query Match:              94.8%          Indels:         0
DB:                       6              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CD627107 (1-567)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db         65 GGAGACCCGGGGTCGGAGATAATAGAATTTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 124

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        125 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 184

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        185 GTTCTTGAATAC 196

RESULT 55
BM352924
LOCUS       BM352924                627 bp    mRNA    linear   EST 07-JAN-2002
DEFINITION  ig69a09.y1 HR85 islet Homo sapiens cDNA 5' similar to TR:Q9Y4E5
            Q9Y4E5 KIAA0576 PROTEIN ;, mRNA sequence.
ACCESSION   BM352924
VERSION     BM352924.1  GI:18085282
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 627)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ig69a09.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
```

Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 481.

```
FEATURES             Location/Qualifiers
     source          1. .627
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
ORIGIN

Alignment Scores:
Pred. No.:                 2.42e-17      Length:          627
Score:                     217.00        Matches:         42
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               94.8%         Indels:          0
DB:                        3             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BM352924 (1-627)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 CCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACA 71

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 ACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCTGTTCTT 131

Qy         43 GluTyr 44
              ||||||
Db        132 GAATAC 137

RESULT 56
BF667883
LOCUS       BF667883                768 bp    mRNA    linear   EST 21-DEC-2000
DEFINITION  602122117F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4279180 5',
            mRNA sequence.
ACCESSION   BF667883
VERSION     BF667883.1  GI:11941778
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 768)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1104  row: a  column: 05
            High quality sequence stop: 678.
FEATURES             Location/Qualifiers
     source          1. .768
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4279180"
                     /tissue_type="primitive neuroectoderm"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_56"
                     /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.65 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
ORIGIN

Alignment Scores:
Pred. No.:                 4.09e-17      Length:          768
Score:                     216.00        Matches:         44
Percent Similarity:        97.8%         Conservative:    0
Best Local Similarity:     97.8%         Mismatches:      0
Query Match:               94.3%         Indels:          1
DB:                        2             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BF667883 (1-768)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProPro-AlaGlyProGluAlaSerGl 20
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         64 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGGCTGGCCCTGAGGCATCTGA 123

Qy         20 uSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 GTCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACC 183

Qy         40 oValLeuGluTyr 44
              |||||||||||||
Db        184 TGTTCTTGAATAC 196

RESULT 57
BF030066
LOCUS       BF030066                814 bp    mRNA    linear   EST 10-OCT-2000
DEFINITION  601557142F1 NIH_MGC_58 Homo sapiens cDNA clone IMAGE:3826883 5',
            mRNA sequence.
ACCESSION   BF030066
VERSION     BF030066.1  GI:10737778
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 814)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM495  row: c  column: 12
            High quality sequence stop: 473.
```

```
FEATURES             Location/Qualifiers
     source          1. .814
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3826883"
                     /tissue_type="hypernephroma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_58"
                     /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.35 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
ORIGIN

Alignment Scores:
Pred. No.:                4.37e-17        Length:          814
Score:                    216.00          Matches:         42
Percent Similarity:       95.5%           Conservative:    0
Best Local Similarity:    95.5%           Mismatches:      2
Query Match:              94.3%           Indels:          0
DB:                       2               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x BF030066 (1-814)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 118

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||   ||||||   |||||||||
Db        119 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCCGTGAAGGAACATCGAGACCT 178

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        179 GTTCTTGAATAC 190

RESULT 58
BY064671
LOCUS       BY064671                362 bp    mRNA    linear   EST 06-DEC-2002
DEFINITION  BY064671 RIKEN full-length enriched, pooled tissues, 16 days
            embryo, etc. Mus musculus cDNA clone I920036I20 5', mRNA sequence.
ACCESSION   BY064671
VERSION     BY064671.1  GI:26182352
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 362)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
    TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
            Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
            Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
            Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
            Hayashizaki,Y. Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .362
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="I920036I20"
                     /clone_lib="RIKEN full-length enriched, pooled tissues, 16
                     days embryo, etc."
                     /note="pooled tissues: (dev_stage=16 days
                     embryo,tissue_type=heart,sex=mix),(dev_stage=16 days
                     embryo,tissue_type=kidney,sex=mix),(dev_stage=17 days
                     embryo,tissue_type=heart,sex=mix),(dev_stage=17 days
                     embryo,tissue_type=stomach,sex=mix),(dev_stage=17 days
                     embryo,tissue_type=kidney,sex=mix),(dev_stage=17 days
                     pregnant, adult,tissue_type=amnion,sex=female),
                     (dev_stage=13 days embryo,tissue_type=liver,sex=mix)"
ORIGIN

Alignment Scores:
Pred. No.:                2.33e-17        Length:          362
Score:                    215.00          Matches:         41
Percent Similarity:       95.5%           Conservative:    1
Best Local Similarity:    93.2%           Mismatches:      2
Query Match:              93.9%           Indels:          0
DB:                       5               Gaps:            0
```

```
US-10-757-745-6_COPY_2_45 (1-44) x BY064671 (1-362)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 210

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||:::||||||||||||||||||||||||||||||||||||    ||||||
Db        211 TCTACCACGGATGACAACGAAGATGACATTCAGTTTGTCAGTGAAGGACCGTTAAGACCC 270

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        271 GTTTTGGAATAC 282

RESULT 59
CO738423
LOCUS       CO738423                 702 bp    mRNA    linear   EST 29-JUL-2004
DEFINITION  S1LE04c20g21f1 squirrel embryo library 1 Spermophilus lateralis
            cDNA clone 20g21 5', mRNA sequence.
ACCESSION   CO738423
VERSION     CO738423.1  GI:50825693
KEYWORDS    EST.
SOURCE      Spermophilus lateralis (golden-mantled ground squirrel)
  ORGANISM  Spermophilus lateralis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Sciuridae; Xerinae; Marmotini; Spermophilus.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
            Rogers,J. and Cossins,A.R.
  TITLE     Microarray analysis of transcriptional changes during hibernation
            in the golden mantled ground squirrel, Spermophilus lateralis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Andrew R. Cossins
            Laboratory for Environmental Gene Regulation
            University of Liverpool
            School of Biological Sciences, The Biosciences Building, Crown
            Street, Liverpool, United Kingdom, L69 7ZB
            Tel: +44(0)151-795-4510
            Fax: +44(0)151-795-4431
            Email: cossins@liv.ac.uk
            Vector has been trimmed from this EST.
            Plate: 20  row: g  column: 21
            Seq primer: pflc T7 (5'-AATACGACTCACTATAGGG-3')
            High quality sequence stop: 702.
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="Spermophilus lateralis"
                     /mol_type="mRNA"
                     /db_xref="taxon:76772"
                     /clone="20g21"
                     /sex="male and female"
                     /tissue_type="embryo"
                     /dev_stage="embryonic"
                     /lab_host="E.coli Electromax DH10B"
                     /clone_lib="squirrel embryo library 1"
                     /note="Vector: pFLC; Site_1: SalI GTCGAG; Site_2: BamHI
                     GGATCC; Normalized and subtracted cDNA library prepared
                     from embryos"
ORIGIN

Alignment Scores:
Pred. No.:              4.95e-17        Length:         702
Score:                  215.00          Matches:        42
Percent Similarity:     95.5%           Conservative:   0
Best Local Similarity:  95.5%           Mismatches:     2
Query Match:            93.9%           Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CO738423 (1-702)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        206 GGAGACCCAGGGTCGGAGATAATAGAATCTGCCCCTCCAGCTGGGCCTGAGGCATCTGAA 265

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||    |||||||||
Db        266 TCAACAACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACTGTCGAGACCT 325

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        326 GTTCTTGAATAC 337

RESULT 60
BQ769546
LOCUS       BQ769546                 796 bp    mRNA    linear   EST 26-AUG-2002
DEFINITION  UI-M-FI0-byr-h-11-0-UI.r1 NIH_BMAP_FI0 Mus musculus cDNA clone
            IMAGE:5698234 5', mRNA sequence.
ACCESSION   BQ769546
VERSION     BQ769546.1  GI:21978020
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            The following repetitive elements were found in this cDNA
            sequence: 1-42, >(CGG)n#Simple_repeat
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .796
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5698234"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 12.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_FI0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CAGCCACGAC. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

Alignment Scores:
Pred. No.:              5.7e-17         Length:         796
Score:                  215.00          Matches:        41
```

```
Percent Similarity:     95.5%          Conservative:  1
Best Local Similarity:  93.2%          Mismatches:    2
Query Match:            93.9%          Indels:        0
DB:                     5              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x BQ769546 (1-796)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   |||||||||:::||||||||||||||||||||||||||||||||
Db         70 GGAGACCCAGGGCCGGAGATAATTCAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 129

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        130 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 189

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        190 GTTTTGGAATAC 201

RESULT 61
BY009270
LOCUS       BY009270                368 bp    mRNA    linear   EST 06-DEC-2002
DEFINITION  BY009270 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus
            musculus cDNA clone G730011G18 5', mRNA sequence.
ACCESSION   BY009270
VERSION     BY009270.1  GI:26069519
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 368)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
            Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
            Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
            Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
            Hayashizaki,Y. Direct Submission
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES              Location/Qualifiers
     source           1. .368
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /db_xref="taxon:10090"
                      /clone="G730011G18"
                      /tissue_type="lung"
                      /cell_line="RCB-0558 LLC"
                      /clone_lib="RIKEN full-length enriched, lung RCB-0558 LLC
                      cDNA"
ORIGIN

Alignment Scores:
Pred. No.:              4.26e-17       Length:        368
Score:                  213.00         Matches:       41
Percent Similarity:     93.2%          Conservative:  0
Best Local Similarity:  93.2%          Mismatches:    3
Query Match:            93.0%          Indels:        0
DB:                     5              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x BY009270 (1-368)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   |||||||||||||||||||||||||||||||   |||||||||
Db         86 GGAGACCCAGGGNCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGCGGCATCAGAG 145

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        146 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 205

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        206 GTTTTGGAATAC 217

RESULT 62
CD627104/c
LOCUS       CD627104                547 bp    mRNA    linear   EST 12-JAN-2004
DEFINITION  56066565J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD627104
VERSION     CD627104.1  GI:40275370
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
  TITLE     Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
```

```
  JOURNAL   Genomics 84 (1), 205-210 (2004)
   PUBMED   15203218
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="FLP"
                     /note="Vector: pDrive Cloning Vector"
ORIGIN

Alignment Scores:
Pred. No.:              6.69e-17        Length:         547
Score:                  213.00          Matches:        41
Percent Similarity:     93.2%           Conservative:   0
Best Local Similarity:  93.2%           Mismatches:     3
Query Match:            93.0%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CD627104 (1-547)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 437

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||   |||   |||||||||   ||||||
Db        436 TCAACAACGGATGAAAATGAAGACGACATTCAGCTTGTCGGTGAAGGACCATTACGACCT 377

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        376 GTTCTTGAATAC 365

RESULT 63
CD627103
LOCUS       CD627103                553 bp    mRNA    linear   EST 12-JAN-2004
DEFINITION  56066565H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION   CD627103
VERSION     CD627103.1  GI:40275369
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 553)
  AUTHORS   Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
  TITLE     Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
  JOURNAL   Genomics 84 (1), 205-210 (2004)
   PUBMED   15203218
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES             Location/Qualifiers
     source          1. .553
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="FLP"
                     /note="Vector: pDrive Cloning Vector"
ORIGIN

Alignment Scores:
Pred. No.:              6.77e-17        Length:         553
Score:                  213.00          Matches:        41
Percent Similarity:     93.2%           Conservative:   0
Best Local Similarity:  93.2%           Mismatches:     3
Query Match:            93.0%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CD627103 (1-553)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 111

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||   |||   |||||||||   ||||||
Db        112 TCAACAACGGATGAAAATGAAGACGACATTCAGCTTGTCGGTGAAGGACCATTACGACCT 171

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        172 GTTCTTGAATAC 183

RESULT 64
BY210989
LOCUS       BY210989                353 bp    mRNA    linear   EST 10-DEC-2002
DEFINITION  BY210989 RIKEN full-length enriched, B6-derived CD11 +ve dendritic
            cells Mus musculus cDNA clone F730320O06 5', mRNA sequence.
ACCESSION   BY210989
VERSION     BY210989.1  GI:26391573
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 353)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
```

Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .353
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F730320O06"
/cell_type="B6-derived CD11 +ve dendritic cells"
/clone_lib="RIKEN full-length enriched, B6-derived CD11 +ve dendritic cells"

ORIGIN

Alignment Scores:

Pred. No.:	5.45e-17	Length:	353
Score:	212.00	Matches:	41
Percent Similarity:	93.2%	Conservative:	0
Best Local Similarity:	93.2%	Mismatches:	3
Query Match:	92.6%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BY210989 (1-353)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         44 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 103

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        104 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 163

Qy         41 ValLeuGluTyr 44
              |||   ||||||
Db        164 GTTTGGGAATAC 175
```

RESULT 65
BY307104
LOCUS BY307104 378 bp mRNA linear EST 11-DEC-2002
DEFINITION BY307104 RIKEN full-length enriched, 12.5 days embryo Rathke's pouches Mus musculus cDNA clone K920013E20 5', mRNA sequence.
ACCESSION BY307104
VERSION BY307104.1 GI:26497441
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 378)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Michelle Brinkmeier and Sally Camper (Dept. Human Genetics University of Michigan Medical School 4301 MSRB 3 1500 W. Medical Center Dr. Ann Arbor,MI 48109-0638 USA) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .378
/organism="Mus musculus"
/mol_type="mRNA"

/strain="CD-1"
/db_xref="taxon:10090"
/clone="K920013E20"
/tissue_type="Rathke's pouches"
/dev_stage="12.5 days embryo"
/clone_lib="RIKEN full-length enriched, 12.5 days embryo Rathke's pouches"

ORIGIN

Alignment Scores:

Pred. No.:	5.89e-17	Length:	378
Score:	212.00	Matches:	41
Percent Similarity:	93.2%	Conservative:	0
Best Local Similarity:	93.2%	Mismatches:	3
Query Match:	92.6%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BY307104 (1-378)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         80 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 139

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        140 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 199

Qy         41 ValLeuGluTyr 44
              |||   ||||||
Db        200 GTTTGGGAATAC 211
```

RESULT 66
BI560244
LOCUS BI560244 1165 bp mRNA linear EST 05-SEP-2001
DEFINITION 603253376F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5295787 5', mRNA sequence.
ACCESSION BI560244
VERSION BI560244.1 GI:15447558
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1165)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11748 row: g column: 20
High quality sequence stop: 671.
FEATURES Location/Qualifiers
source 1. .1165
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5295787"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	2.83e-16	Length:	1165
Score:	211.00	Matches:	43
Percent Similarity:	97.8%	Conservative:	1
Best Local Similarity:	95.6%	Mismatches:	0
Query Match:	92.1%	Indels:	1
DB:	3	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BI560244 (1-1165)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 129

Qy         21 SerThrThrAspGlu-AsnGluAspAspIleGlnPheValSerGluGlyProSerArgPr 40
              |||||||||||||||| :::|||||||||||||||||||||||||||||||||||||||
Db        130 TCAACAACGGATGAAAGATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACC 189

Qy         40 oValLeuGluTyr 44
              ||||||||||||
Db        190 TGTTCTTGAATAC 202
```

RESULT 67
DN371969
LOCUS DN371969 307 bp mRNA linear EST 07-MAR-2005
DEFINITION LIB3733-037-A1-K1-C2 LIB3733 Canis familiaris cDNA clone CLN12920479, mRNA sequence.
ACCESSION DN371969
VERSION DN371969.1 GI:60553189
KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 307)
AUTHORS Staten,N.R.
TITLE Direct Submission (Staten,N.R.)
JOURNAL Unpublished (2005)
COMMENT Contact: Nick Staten
Tel: 636 247 6855
Email: nicholas.r.staten@pfizer.com.
FEATURES Location/Qualifiers
source 1. .307
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="beagle"
/db_xref="taxon:9615"
/clone="CLN12920479"
/tissue_type="brain"
/lab_host="BHK"
/clone_lib="LIB3733"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; brain - dog"

ORIGIN

Alignment Scores:

Pred. No.:	8.35e-17	Length:	307
Score:	210.00	Matches:	40
Percent Similarity:	97.7%	Conservative:	3
Best Local Similarity:	90.9%	Mismatches:	1
Query Match:	91.7%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x DN371969 (1-307)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db         93 GGAGACCCGGGGTCGGAGATGATAGAATCTGTCCCTCCAGCTGGGCCCGAGGCGTCTGAG 152

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        153 GCAACGACGGATGAAAACGAAGATGACATTCAGTTTGTTAGCGAAGGACCATCAAAATGT 212

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        213 GTTCTTGAATAC 224

RESULT 68
DN371450
LOCUS       DN371450                764 bp    mRNA    linear   EST 07-MAR-2005
DEFINITION  LIB3733-028-A1-K1-A12 LIB3733 Canis familiaris cDNA clone
            CLN12919601, mRNA sequence.
ACCESSION   DN371450
VERSION     DN371450.1  GI:60552398
KEYWORDS    EST.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 764)
  AUTHORS   Staten,N.R.
  TITLE     Direct Submission (Staten,N.R.)
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Nick Staten
            Tel: 636 247 6855
            Email: nicholas.r.staten@pfizer.com.
FEATURES             Location/Qualifiers
     source          1. .764
                     /organism="Canis familiaris"
                     /mol_type="mRNA"
                     /strain="beagle"
                     /db_xref="taxon:9615"
                     /clone="CLN12919601"
                     /tissue_type="brain"
                     /lab_host="BHK"
                     /clone_lib="LIB3733"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; brain
                     - dog"
ORIGIN

Alignment Scores:
Pred. No.:                2.35e-16       Length:          764
Score:                    210.00         Matches:         40
Percent Similarity:       97.7%          Conservative:    3
Best Local Similarity:    90.9%          Mismatches:      1
Query Match:              91.7%          Indels:          0
DB:                       8              Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x DN371450 (1-764)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        110 GGAGACCCGGGGTCGGAGATGATAGAATCTGTCCCTCCAGCTGGGCCCGAGGCGTCTGAG 169

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        170 GCAACGACGGATGAAAACGAAGATGACATTCAGTTTGTTAGCGAAGGACCATCAAAATGT 229

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        230 GTTCTTGAATAC 241

RESULT 69
BY028275
LOCUS       BY028275                377 bp    mRNA    linear   EST 06-DEC-2002
DEFINITION  BY028275 RIKEN full-length enriched, mammary gland RCB-0527
            Jyg-MC(B) cDNA Mus musculus cDNA clone G930047P03 5', mRNA
            sequence.
ACCESSION   BY028275
VERSION     BY028275.1  GI:26133718
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 377)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
            Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
            Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
            Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
            Hayashizaki,Y. Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
```

```
FEATURES             Location/Qualifiers
     source          1. .377
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="G930047P03"
                     /tissue_type="mammary gland"
                     /cell_line="RCB-0527 Jyg-MC(B)"
                     /clone_lib="RIKEN full-length enriched, mammary gland
                     RCB-0527 Jyg-MC(B) cDNA"
ORIGIN

Alignment Scores:
Pred. No.:              1.89e-16        Length:         377
Score:                  208.00          Matches:        40
Percent Similarity:     93.2%           Conservative:   1
Best Local Similarity:  90.9%           Mismatches:     3
Query Match:            90.8%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x BY028275 (1-377)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   :::||||||||||||||||||||||||||||||||||||||||||
Db         63 GGAGACCCAGGGCCGCAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 122

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||   ||||||
Db        123 TCTACCACGGATGAAAACGAACATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 182

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        183 GTTTTGGAATAC 194

RESULT 70
CB808042
LOCUS       CB808042                406 bp    mRNA    linear   EST 16-MAY-2003
DEFINITION  AMGNNUC:NRPI4-00011-G3-A W Rat pituitary (10472) Rattus norvegicus
            cDNA clone nrpi4-00011-g3 5', mRNA sequence.
ACCESSION   CB808042
VERSION     CB808042.1  GI:29927623
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Amgen EST Program.
  TITLE     Amgen Rat EST Program
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00011  row: g  column: 3.
FEATURES             Location/Qualifiers
     source          1. .406
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="nrpi4-00011-g3"
                     /clone_lib="W Rat pituitary (10472)"
                     /note="Vector: pSPORT1; Site_1: Sal1; Site_2: Not1; W Rat
                     pituitary adult female Wistar rat avg insert size 2.1 kb"
ORIGIN

Alignment Scores:
Pred. No.:              2.06e-16        Length:         406
Score:                  208.00          Matches:        40
Percent Similarity:     90.9%           Conservative:   0
Best Local Similarity:  90.9%           Mismatches:     4
Query Match:            90.8%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CB808042 (1-406)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||   ||||||
Db        122 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGCCATCAGAG 181

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||   ||||||||||||||||||||||||||||||||||||||||   ||||||
Db        182 TCAACGGCGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 241

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        242 GTTCTTGAATAT 253

RESULT 71
CB789225
LOCUS       CB789225                443 bp    mRNA    linear   EST 16-MAY-2003
DEFINITION  AMGNNUC:MRPE3-00073-H3-A placenta embryo D17 (10379) Rattus
            norvegicus cDNA clone mrpe3-00073-h3 5', mRNA sequence.
ACCESSION   CB789225
VERSION     CB789225.1  GI:29877617
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   Amgen EST Program.
  TITLE     Amgen Rat EST Program
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00073  row: h  column: 3.
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="mrpe3-00073-h3"
                     /tissue_type="placenta embryo"
                     /clone_lib="placenta embryo D17 (10379)"
                     /note="Vector: pSPORT1; placenta embryo D17"
ORIGIN

Alignment Scores:
Pred. No.:              2.27e-16        Length:         443
Score:                  208.00          Matches:        40
Percent Similarity:     90.9%           Conservative:   0
Best Local Similarity:  90.9%           Mismatches:     4
Query Match:            90.8%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CB789225 (1-443)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||   ||||||
Db         83 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGCCATCAGAG 142

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||   ||||||||||||||||||||||||||||||||||||||||   ||||||
Db        143 TCAACGGCGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 202

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        203 GTTCTTGAATAT 214

RESULT 72
CK357744
```

```
LOCUS       CK357744                578 bp    mRNA    linear   EST 23-DEC-2003
DEFINITION  AGENCOURT_17196436 NIH_MGC_233 Rattus norvegicus cDNA clone
            IMAGE:7104890 5', mRNA sequence.
ACCESSION   CK357744
VERSION     CK357744.1  GI:40323676
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Howard Jacobs
             cDNA Library Preparation: Express Genomics
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14963  row: b  column: 24
            High quality sequence stop: 520.
FEATURES             Location/Qualifiers
     source          1. .578
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="IMAGE:7104890"
                     /tissue_type="heart, pooled"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_233"
                     /note="Organ: heart; Vector: pExpress-1; Site_1: EcoRV;
                     Site_2: NotI; RNA obtained from pooled heart tissue from a
                     mix of male and female animals at 8 wk old. Tissues were
                     snap-frozen and kept at -80C for two days before RNA
                     extraction and purification (Tri-reagent method). cDNA was
                     primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb
                     resulted in an average insert size of 2 kb. This primary
                     library is not normalized (normalized primary library is
                     NIH_MGC_234) and was constructed by Express Genomics
                     (Frederick, MD). Note: this is a NIH_MGC library."
ORIGIN

Alignment Scores:
Pred. No.:                3.07e-16        Length:         578
Score:                    208.00          Matches:        40
Percent Similarity:       90.9%           Conservative:   0
Best Local Similarity:    90.9%           Mismatches:     4
Query Match:              90.8%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x CK357744 (1-578)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||   ||||||
Db        212 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGCCATCAGAG 271

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||   ||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        272 TCAACGGCGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 331

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        332 GTTCTTGAATAT 343
```

```
RESULT 73
DN933911
LOCUS       DN933911                878 bp    mRNA    linear   EST 27-APR-2005
DEFINITION  AGENCOURT_49984462 NCI_CGAP_Pr50 Rattus norvegicus cDNA clone
            IMAGE:7938755 5', mRNA sequence.
ACCESSION   DN933911
VERSION     DN933911.1  GI:62917338
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 878)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeff Green/Paturu Kondaiah, NCI.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM16372  row: c  column: 09
            High quality sequence start: 9
            High quality sequence stop: 691.
FEATURES             Location/Qualifiers
     source          1. .878
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="IMAGE:7938755"
                     /sex="male"
                     /tissue_type="dorsolateral prostate, pool of 3-, 5-, and
                     7-days post-castration"
                     /dev_stage="adult, 11 week"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Pr50"
                     /note="Organ: prostate; Vector: pCMV-SPORT6.1; Site_1:
                     NotI; Site_2: EcoRV; Cloned unidirectionally. Primer:
                     Oligo dT. Pool of 3 primary libraries: NCI_CGAP_Pr39
                     (dorsolateral prostate from 11 wk male, 3 days
                     post-castration, average insert size 2.7 kb),
                     NCI_CGAP_Pr29 (dorsolateral prostate from 11 wk male, 5
                     days post-castration, average insert size 2.2 kb) and
                     NCI_CGAP_Pr42 (dorsolateral prostate from 11 wk male, 7
                     days post-castration, average insert size 2.2 kb).
                     Constructed by Life Technologies/Invitrogen. Note: this is
                     a NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:                4.93e-16        Length:         878
Score:                    208.00          Matches:        40
Percent Similarity:       90.9%           Conservative:   0
Best Local Similarity:    90.9%           Mismatches:     4
Query Match:              90.8%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x DN933911 (1-878)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||   ||||||
Db        200 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGCCATCAGAG 259

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
```

```
                ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db        260 TCAACGGCGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGACCC 319

Qy         41 ValLeuGluTyr 44
                ||||||||||||
Db        320 GTTCTTGAATAT 331
```

RESULT 74
BB854831
LOCUS BB854831 343 bp mRNA linear EST 26-NOV-2001
DEFINITION BB854831 RIKEN full-length enriched, B16 F10Y cells Mus musculus cDNA clone G370021E15 5', mRNA sequence.
ACCESSION BB854831
VERSION BB854831.1 GI:17096285
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 343)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.

FEATURES Location/Qualifiers
source 1. .343
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G370021E15"
/cell_type="B16 F10Y cells"
/clone_lib="RIKEN full-length enriched, B16 F10Y cells"

ORIGIN

Alignment Scores:

Pred. No.:	7.33e-16	Length:	343
Score:	203.00	Matches:	40
Percent Similarity:	93.0%	Conservative:	0
Best Local Similarity:	93.0%	Mismatches:	3
Query Match:	88.6%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BB854831 (1-343)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 GGAGACCCAGGGCCGGAGATAATAGAATCTGTCCCACCAGCTGGGCCCGAGGCATCAGAG 135

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        136 TCTACCACGGATGAAAACGAAGATGACATTCAGTTTGTTAGTGAAGGACCGTTAAGAACC 195

Qy         41 ValLeuGlu 43
              |||||||||
Db        196 GTTTTGGAA 204
```

RESULT 75
BG286553
LOCUS BG286553 741 bp mRNA linear EST 21-FEB-2001
DEFINITION 602382555F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4500277 5', mRNA sequence.
ACCESSION BG286553
VERSION BG286553.1 GI:13039536
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 741)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM10365 row: m column: 14
High quality sequence stop: 737.

FEATURES Location/Qualifiers
source 1. .741
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4500277"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_93"
/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	4.22e-15	Length:	741
Score:	200.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	87.3%	Indels:	1
DB:	2	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x BG286553 (1-741)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         45 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 104

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        105 TCAACAACGGATGAAAATGAAGACGACATTCAGTT-GTCAGTGAAGGACCATCGAGACCT 163

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        164 GTTCTTGAATAC 175
```

Search completed: March 11, 2006, 03:41:13
Job time : 1188.43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 03:44:14 ; Search time 127.337 Seconds
(without alignments)
797.263 Million cell updates/sec

Title: US-10-757-745-6_COPY_2_45
Perfect score: 229
Sequence: 1 GDPGSEIIESVPPAGPEASE.....ENEDDIQFVSEGPSRPVLEY 44

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7673375 seqs, 1153648444 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170339_4330/app_query.fasta_1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1676_@runat_09032006_170339_4330 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	223	97.4	4961	8	US-10-775-169-139	Sequence 139, App
2	66	28.8	1203	7	US-10-932-182A-1033	Sequence 1033, Ap
3	66	28.8	1203	7	US-10-932-182A-1033	Sequence 1033, Ap
4	65	28.4	1388	8	US-10-750-185-60725	Sequence 60725, A
5	65	28.4	1388	8	US-10-750-623-60725	Sequence 60725, A
c 6	62	27.1	4984	7	US-10-116-788A-32	Sequence 32, Appl
c 7	61.5	26.9	1968	8	US-10-750-185-57152	Sequence 57152, A
c 8	61.5	26.9	1968	8	US-10-750-623-57152	Sequence 57152, A
c 9	61	26.6	162173	12	US-11-121-086-72	Sequence 72, Appl
c 10	60.5	26.4	756	6	US-09-925-065A-78676	Sequence 78676, A
c 11	60.5	26.4	756	6	US-09-925-065A-78677	Sequence 78677, A
c 12	60.5	26.4	35962	8	US-10-775-169-234	Sequence 234, App
c 13	60	26.2	571	6	US-09-925-065A-503760	Sequence 503760,
14	60	26.2	6277	12	US-11-136-527-707	Sequence 707, App
15	59.5	26.0	1062	12	US-11-137-465-11	Sequence 11, Appl
16	59.5	26.0	1347	12	US-11-137-465-12	Sequence 12, Appl
c 17	59.5	26.0	88873	8	US-10-995-561-13383	Sequence 13383, A
18	59	25.8	1114	12	US-11-226-657-39	Sequence 39, Appl
c 19	59	25.8	1235	8	US-10-750-185-54635	Sequence 54635, A
c 20	59	25.8	1235	8	US-10-750-623-54635	Sequence 54635, A
21	59	25.8	3835	12	US-11-091-883-313	Sequence 313, App
22	59	25.8	3848	8	US-10-821-234-258	Sequence 258, App
23	59	25.8	4990	12	US-11-128-061-932	Sequence 932, App
24	59	25.8	4990	12	US-11-128-049-932	Sequence 932, App
c 25	58.5	25.5	894	12	US-11-039-398-3	Sequence 3, Appli
c 26	58.5	25.5	1356	12	US-11-039-398-1	Sequence 1, Appli
c 27	58.5	25.5	1461	12	US-11-039-398-5	Sequence 5, Appli
c 28	58.5	25.5	3642	12	US-11-039-398-13	Sequence 13, Appl
c 29	58.5	25.5	3651	12	US-11-039-398-11	Sequence 11, Appl
c 30	58.5	25.5	3660	12	US-11-039-398-9	Sequence 9, Appli
c 31	58.5	25.5	3669	12	US-11-039-398-7	Sequence 7, Appli
c 32	58.5	25.5	3699	12	US-11-039-398-17	Sequence 17, Appl
c 33	58.5	25.5	3708	12	US-11-039-398-15	Sequence 15, Appl
c 34	58.5	25.5	3750	12	US-11-039-398-21	Sequence 21, Appl
c 35	58.5	25.5	3759	12	US-11-039-398-19	Sequence 19, Appl
c 36	58.5	25.5	4213	8	US-10-624-932-25	Sequence 25, Appl
c 37	58.5	25.5	4277	12	US-11-039-398-23	Sequence 23, Appl
c 38	57.5	25.1	530	6	US-09-925-065A-16151	Sequence 16151, A
c 39	57	24.9	859	6	US-09-925-065A-80196	Sequence 80196, A
40	57	24.9	1185	7	US-10-932-182A-76811	Sequence 76811, A
41	57	24.9	1185	7	US-10-932-182A-76811	Sequence 76811, A
c 42	57	24.9	1225	8	US-10-750-185-53105	Sequence 53105, A
c 43	57	24.9	1225	8	US-10-750-623-53105	Sequence 53105, A
44	57	24.9	1725	7	US-10-932-182A-4234	Sequence 4234, Ap
45	57	24.9	1725	7	US-10-932-182A-4234	Sequence 4234, Ap
c 46	57	24.9	28693	8	US-10-995-561-13341	Sequence 13341, A
c 47	57	24.9	116856	12	US-11-143-980-1	Sequence 1, Appli
c 48	57	24.9	153376	12	US-11-121-086-5	Sequence 5, Appli
c 49	56.5	24.7	530	6	US-09-925-065A-16150	Sequence 16150, A
50	56.5	24.7	636	6	US-09-925-065A-748950	Sequence 748950,
51	56.5	24.7	641	6	US-09-925-065A-749153	Sequence 749153,
52	56.5	24.7	641	6	US-09-925-065A-821963	Sequence 821963,
53	56.5	24.7	1507	9	US-11-096-568A-6495	Sequence 6495, Ap
54	56.5	24.7	2176	7	US-10-330-773-340	Sequence 340, App
55	56.5	24.7	215126	7	US-10-330-773-339	Sequence 339, App
c 56	56.5	24.7	348101	7	US-10-330-773-122	Sequence 122, App
c 57	56	24.5	541	6	US-09-925-065A-306831	Sequence 306831,
c 58	56	24.5	625	6	US-09-925-065A-7171	Sequence 7171, Ap
c 59	56	24.5	625	6	US-09-925-065A-7172	Sequence 7172, Ap
c 60	56	24.5	625	6	US-09-925-065A-7173	Sequence 7173, Ap
c 61	56	24.5	635	6	US-09-925-065A-82701	Sequence 82701, A
c 62	56	24.5	635	6	US-09-925-065A-82702	Sequence 82702, A
63	56	24.5	6429	9	US-11-179-624-2	Sequence 2, Appli
64	56	24.5	62444	7	US-10-330-773-551	Sequence 551, App
65	55.5	24.2	321	7	US-10-932-182A-78518	Sequence 78518, A
66	55.5	24.2	321	7	US-10-932-182A-78518	Sequence 78518, A
67	55.5	24.2	632	6	US-09-925-065A-123746	Sequence 123746,
68	55.5	24.2	1062	7	US-10-932-182A-3520	Sequence 3520, Ap
69	55.5	24.2	1062	7	US-10-932-182A-3520	Sequence 3520, Ap
70	55.5	24.2	2431	12	US-11-000-688-366	Sequence 366, App
c 71	55.5	24.2	2776	8	US-10-750-185-24916	Sequence 24916, A
c 72	55.5	24.2	2776	8	US-10-750-623-24916	Sequence 24916, A
73	55.5	24.2	2787	12	US-11-182-016-4	Sequence 4, Appli
74	55.5	24.2	34214	12	US-11-127-832-27	Sequence 27, Appl
75	55	24.0	201	12	US-11-124-368A-13139	Sequence 13139, A

ALIGNMENTS

RESULT 1
US-10-775-169-139
; Sequence 139, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dorner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 139
; LENGTH: 4961
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-139

Alignment Scores:

Pred. No.:	6.3e-20	Length:	4961
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-775-169-139 (1-4961)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 106

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 166

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        167 GTTCTTGAATAC 178
```

RESULT 2
US-10-932-182A-1033
; Sequence 1033, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1033
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-1033

Alignment Scores:

Pred. No.:	29.6	Length:	1203
Score:	66.00	Matches:	12
Percent Similarity:	58.1%	Conservative:	6
Best Local Similarity:	38.7%	Mismatches:	13
Query Match:	28.8%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-932-182A-1033 (1-1203)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||:::|||   :::   ::::::|||   :::|||          |||      |||
Db       1063 GGTGAGCCTGATGCCGGTGTCGTAGAAGAAATGCCATCAAAAAATGCTGAAGATGTTGAA 1122

Qy         21 SerThrThrAspGluAsnGluAspAspIleGln 31
                  ||||||||||||||:::   |||
Db       1123 GACGAAAACGATGAAAATGAAGATGAAGATCAG 1155
```

RESULT 3
US-10-932-182A-1033
; Sequence 1033, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1033
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-1033

Alignment Scores:

Pred. No.:	29.6	Length:	1203
Score:	66.00	Matches:	12
Percent Similarity:	58.1%	Conservative:	6
Best Local Similarity:	38.7%	Mismatches:	13
Query Match:	28.8%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-932-182A-1033 (1-1203)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||:::|||   :::   :::::::|||   :::|||          |||      |||
Db       1063 GGTGAGCCTGATGCCGGTGTCGTAGAAGAAATGCCATCAAAAAATGCTGAAGATGTTGAA 1122

Qy         21 SerThrThrAspGluAsnGluAspAspIleGln 31
                  ||||||||||||||:::   |||
Db       1123 GACGAAAACGATGAAAATGAAGATGAAGATCAG 1155
```

RESULT 4
US-10-750-185-60725
; Sequence 60725, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1

```
; SEQ ID NO 60725
;   LENGTH: 1388
;   TYPE: DNA
;   ORGANISM: Bovine  19866881392376
US-10-750-185-60725

Alignment Scores:
Pred. No.:              47          Length:         1388
Score:                  65.00       Matches:        15
Percent Similarity:     47.6%       Conservative:   5
Best Local Similarity:  35.7%       Mismatches:     14
Query Match:            28.4%       Indels:         8
DB:                     8           Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-750-185-60725 (1-1388)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||     |||                        |||   :::|||
Db        963 GGGCATCCCGCAGGAGAG----------------------GGGCGAGACGCTGGCACA 998

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||:::   |||      :::||||||:::   |||||||||       |||       |||
Db        999 AGCAGTGGTGACAGTGAGGACGATGACGTTGTTTTTGTTTCGTCTAAGCCAAGGAGCCCC 1058

Qy         41 ValLeu 42
              :::|||
Db       1059 TTACTC 1064

RESULT 5
US-10-750-623-60725
; Sequence 60725, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 60725
;   LENGTH: 1388
;   TYPE: DNA
;   ORGANISM: Bovine  19866881392376
US-10-750-623-60725

Alignment Scores:
Pred. No.:              47          Length:         1388
Score:                  65.00       Matches:        15
Percent Similarity:     47.6%       Conservative:   5
Best Local Similarity:  35.7%       Mismatches:     14
Query Match:            28.4%       Indels:         8
DB:                     8           Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-750-623-60725 (1-1388)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||     |||                        |||   :::|||
Db        963 GGGCATCCCGCAGGAGAG----------------------GGGCGAGACGCTGGCACA 998

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||:::   |||      :::||||||:::   |||||||||       |||       |||
Db        999 AGCAGTGGTGACAGTGAGGACGATGACGTTGTTTTTGTTTCGTCTAAGCCAAGGAGCCCC 1058

Qy         41 ValLeu 42
              :::|||
Db       1059 TTACTC 1064

RESULT 6
US-10-116-788A-32/c
; Sequence 32, Application US/10116788A
; Publication No. US20060024804A9
; GENERAL INFORMATION:
;  APPLICANT: McDonald, Thomas L.
;  APPLICANT:  Larson, Marilyn A.
;  APPLICANT:  Weber, Annika
;  TITLE OF INVENTION: Genomic Mammary Amyloid A Sequence
;  FILE REFERENCE: P04557US1
;  CURRENT APPLICATION NUMBER: US/10/116,788A
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: 09/425,679
;  PRIOR FILING DATE: 1999-10-22
;  PRIOR APPLICATION NUMBER: 60/218,482
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/218,611
;  PRIOR FILING DATE: 2000-07-17
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
;   LENGTH: 4984
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3191, 3222, 3229, 3323, 3402
;   OTHER INFORMATION: n can be a, t, c, or g
US-10-116-788A-32

Alignment Scores:
Pred. No.:              460         Length:         4984
Score:                  62.00       Matches:        11
Percent Similarity:     50.0%       Conservative:   3
Best Local Similarity:  39.3%       Mismatches:     14
Query Match:            27.1%       Indels:         0
DB:                     7           Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-116-788A-32 (1-4984)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||                    :::||||||:::   |||       |||
Db       4169 GGAGATCCAGGTGGCAGTTTCTGGAGCAGGATTCCACCATCCAGTCCTCCTGTGTCCATG 4110

Qy         21 SerThrThrAspGluAsnGluAsp 28
                 |||:::      ||||||
Db       4109 CCCCACACAGAGAGCACAGAGGAC 4086

RESULT 7
US-10-750-185-57152/c
; Sequence 57152, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 57152
;   LENGTH: 1968
```

```
;   TYPE: DNA
;   ORGANISM: Bovine  19866881011548
US-10-750-185-57152

Alignment Scores:
Pred. No.:              203           Length:        1968
Score:                  61.50         Matches:       14
Percent Similarity:     53.7%         Conservative:  8
Best Local Similarity:  34.1%         Mismatches:    10
Query Match:            26.9%         Indels:        9
DB:                     8             Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-750-185-57152 (1-1968)

Qy          1 GlyAspProGlySerGluIleIleGluSerValPro-----------ProAlaGly--- 15
              ||||||||||||      :::|||:::|||   |||           |||:::|||
Db        828 GGAGACCCAGGCGGCTCAGTGATCAAGAGCCCGCCTGCGCGTGCAGGACCCTCAGGAGAG 769

Qy         16 -----------ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
                       ||||||:::      |||      :::   |||   :::   :::|||
Db        768 GTGGGTTCGATCCCTGAGTCAGGAAGATCCCCTGGAGAAAGAAATGGCAACCCCCTCCAG 709

Qy         32 Phe 32
              :::
Db        708 TAT 706

RESULT 8
US-10-750-623-57152/c
; Sequence 57152, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 57152
;   LENGTH: 1968
;   TYPE: DNA
;   ORGANISM: Bovine  19866881011548
US-10-750-623-57152

Alignment Scores:
Pred. No.:              203           Length:        1968
Score:                  61.50         Matches:       14
Percent Similarity:     53.7%         Conservative:  8
Best Local Similarity:  34.1%         Mismatches:    10
Query Match:            26.9%         Indels:        9
DB:                     8             Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-750-623-57152 (1-1968)

Qy          1 GlyAspProGlySerGluIleIleGluSerValPro-----------ProAlaGly--- 15
              ||||||||||||      :::|||:::|||   |||           |||:::|||
Db        828 GGAGACCCAGGCGGCTCAGTGATCAAGAGCCCGCCTGCGCGTGCAGGACCCTCAGGAGAG 769

Qy         16 -----------ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
                       ||||||:::      |||      :::   |||   :::   :::|||
Db        768 GTGGGTTCGATCCCTGAGTCAGGAAGATCCCCTGGAGAAAGAAATGGCAACCCCCTCCAG 709

Qy         32 Phe 32
              :::
Db        708 TAT 706

RESULT 9
US-11-121-086-72/c
; Sequence 72, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 72
;   LENGTH: 162173
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-72

Alignment Scores:
Pred. No.:              2.43e+04      Length:        162173
Score:                  61.00         Matches:       12
Percent Similarity:     55.6%         Conservative:  3
Best Local Similarity:  44.4%         Mismatches:    12
Query Match:            26.6%         Indels:        0
DB:                     12            Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-121-086-72 (1-162173)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||   |||       :::           ||||||       ||||||   ||||||   :::
Db     120438 CCCCAGAGCCATGGGCTCGCTCACGCCCCTCCGCACTCCCCGGAGGGCTCCGAGGACAGC 120379

Qy         23 ThrAspGluAsnGluAspAsp 29
                 :::|||   |||||||||
Db     120378 GGAGAGGAGGAGGAAGACGAC 120358

RESULT 10
US-09-925-065A-78676/c
; Sequence 78676, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78676
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-78676

Alignment Scores:
Pred. No.:              101           Length:        756
Score:                  60.50         Matches:       17
Percent Similarity:     51.2%         Conservative:  5
```

```
Best Local Similarity:  39.5%          Mismatches:    17
Query Match:            26.4%          Indels:        4
DB:                     6              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-78676 (1-756)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||   |||   |||            ||||||  |||||||||   |||
Db        593 GGCAAGCCTCACAGTACCATT---------GAGCCCCCAGGTGGTCCGGAGCAAAGCAGT 543

Qy         21 SerThrThrAspGlu-AsnGluAspAspIleGlnPheValSerGluGlyProSerArgPr 40
                  |||    :::||||||   |||:::   :::|||      ||||||
Db        542 CAGCCACAGGACCATCAGCGAGGACAGTATTGAAGTCCTCAGTACCTGCCCCTCTGAGGC 483

Qy         40 oValLeu 42
                ::::::
Db        482 CCTCATC 476

RESULT 11
US-09-925-065A-78677/c
; Sequence 78677, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78677
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-78677

Alignment Scores:
Pred. No.:              101            Length:        756
Score:                  60.50          Matches:       17
Percent Similarity:     51.2%          Conservative:  5
Best Local Similarity:  39.5%          Mismatches:    17
Query Match:            26.4%          Indels:        4
DB:                     6              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-78677 (1-756)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||   |||   |||            ||||||  |||||||||   |||
Db        593 GGCAAGCCTCACAGTACCATT---------GAGCCCCCAGGTGGTCCGGAGCAAAGCAGT 543

Qy         21 SerThrThrAspGlu-AsnGluAspAspIleGlnPheValSerGluGlyProSerArgPr 40
                  |||    :::||||||   |||:::   :::|||      ||||||
Db        542 CAGCCACAGGACCATCAGCGAGGACAGTATTGAAGTCCTCAGTACCTGCCCCTCTGAGGC 483

Qy         40 oValLeu 42
                ::::::
Db        482 CCTCATC 476

RESULT 12
US-10-775-169-234/c
; Sequence 234, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Burczynski, Michael
;  APPLICANT:  Twine, Natalie
;  APPLICANT:  Dorner, Andrew
;  APPLICANT:  Trepicchio, William
;  TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
;  FILE REFERENCE: AM101080 (031896-013000)
;  CURRENT APPLICATION NUMBER: US/10/775,169
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 5278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 234
;   LENGTH: 35962
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-775-169-234

Alignment Scores:
Pred. No.:              5.85e+03       Length:        35962
Score:                  60.50          Matches:       13
Percent Similarity:     62.5%          Conservative:  7
Best Local Similarity:  40.6%          Mismatches:    11
Query Match:            26.4%          Indels:        1
DB:                     8              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-775-169-234 (1-35962)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp 28
              |||:::   ||||||      |||:::|||   ::::::||||||         :::
Db      28065 GAGACACCCCCTCCCCCGGCCCCGGACGCT---GACACCACCACCGCCCTCAGGCAGTCA 28009

Qy         29 AspIleGlnPheValSerGluGlyProSerArgPro 40
                  :::         :::||||||||||||||||
Db      28008 CATCTCCCGAGCTCCGCCGAGGGACCCAGCCGCCCC 27973

RESULT 13
US-09-925-065A-503760/c
; Sequence 503760, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 503760
;   LENGTH: 571
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-503760

Alignment Scores:
Pred. No.:              88.3           Length:        571
Score:                  60.00          Matches:       14
Percent Similarity:     50.0%          Conservative:  7
Best Local Similarity:  33.3%          Mismatches:    21
Query Match:            26.2%          Indels:        0
DB:                     6              Gaps:          0
```

```
US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-503760 (1-571)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||  |||  |||       |||      ||||||||||||   |||   ::::::
Db        336 GGACAGCCGCCTTCCAGCCACAGGGAGCCCACACCACCCGCGGGAAAGGAGACAGCCAAA 277

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ::::::      :::   :::           :::||||||   |||      |||
Db        276 GCCTCTAACAGACAGAAGGATTTATTCTGGGCGCTTATGTCTGAGAACCCCTTTCCCCCA 217

Qy         41 ValLeu 42
                |||
Db        216 ACCCTT 211

RESULT 14
US-11-136-527-707
; Sequence 707, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 707
;   LENGTH: 6277
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-707

Alignment Scores:
Pred. No.:              1.09e+03        Length:         6277
Score:                  60.00           Matches:        16
Percent Similarity:     51.3%           Conservative:   4
Best Local Similarity:  41.0%           Mismatches:     17
Query Match:            26.2%           Indels:         2
DB:                     12              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-136-527-707 (1-6277)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              |||||||||   |||       |||      |||         ||||||:::
Db        852 GACCCGGGAAAGGAGCCCGCCCCCTCTCAGAAGCCAGTG------GAAGCCGCCAGAGAA 905

Qy         22 ThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||:::     |||:::      |||      ||||||   |||:::   |||
Db        906 ACCACGGAGGTTCCTGAGGAGCAGACCCAGCCCCCATCCGAAGCCCCTACAGTGCCT 962

RESULT 15
US-11-137-465-11
; Sequence 11, Application US/11137465
; Publication No. US20050255558A1
; GENERAL INFORMATION:
;  APPLICANT: Agarwal, Pankaj
;  APPLICANT:  Murdoch, Paul R.
;  APPLICANT:  Rizvi, Safia, K.
;  APPLICANT:  Smith, Randall, F.
;  APPLICANT:  Xiang, Zhaoying
;  APPLICANT:  Kabnick, Karen
;  TITLE OF INVENTION: NOVEL COMPOUNDS
;  FILE REFERENCE: GP50018
;  CURRENT APPLICATION NUMBER: US/11/137,465
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US/10/239,663
;  PRIOR FILING DATE: 2002-09-24
;  PRIOR APPLICATION NUMBER: PCT/US01/09226
;  PRIOR FILING DATE: 2001-03-22
;  PRIOR APPLICATION NUMBER: 60/192,158
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,668
;  PRIOR FILING DATE: 2000-03-27
;  PRIOR APPLICATION NUMBER: 60/200,166
;  PRIOR FILING DATE: 2000-04-27
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11
;   LENGTH: 1062
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-137-465-11

Alignment Scores:
Pred. No.:              198             Length:         1062
Score:                  59.50           Matches:        15
Percent Similarity:     56.2%           Conservative:   3
Best Local Similarity:  46.9%           Mismatches:     11
Query Match:            26.0%           Indels:         3
DB:                     12              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-137-465-11 (1-1062)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||       |||   |||:::   ||||||   |||       |||   ||||||
Db        688 GGGATGTCAGGGGAAGAACTCAGTGAATCTACTCCAGAGCCTCAAAAAGAAATTTCTGAG 747

Qy         21 Ser---------ThrThrAspGluAsnGluAspAsp 29
              |||         |||   |||:::::|||||||||
Db        748 TCATTGAGTGTCACCAGAGACCAGGATGAAGATGAT 783

RESULT 16
US-11-137-465-12
; Sequence 12, Application US/11137465
; Publication No. US20050255558A1
; GENERAL INFORMATION:
;  APPLICANT: Agarwal, Pankaj
;  APPLICANT:  Murdoch, Paul R.
;  APPLICANT:  Rizvi, Safia, K.
;  APPLICANT:  Smith, Randall, F.
;  APPLICANT:  Xiang, Zhaoying
;  APPLICANT:  Kabnick, Karen
;  TITLE OF INVENTION: NOVEL COMPOUNDS
;  FILE REFERENCE: GP50018
;  CURRENT APPLICATION NUMBER: US/11/137,465
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US/10/239,663
;  PRIOR FILING DATE: 2002-09-24
;  PRIOR APPLICATION NUMBER: PCT/US01/09226
;  PRIOR FILING DATE: 2001-03-22
;  PRIOR APPLICATION NUMBER: 60/192,158
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,668
;  PRIOR FILING DATE: 2000-03-27
;  PRIOR APPLICATION NUMBER: 60/200,166
;  PRIOR FILING DATE: 2000-04-27
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
;   LENGTH: 1347
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-137-465-12

Alignment Scores:
Pred. No.:              254             Length:         1347
Score:                  59.50           Matches:        15
Percent Similarity:     56.2%           Conservative:   3
Best Local Similarity:  46.9%           Mismatches:     11
Query Match:            26.0%           Indels:         3
DB:                     12              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-137-465-12 (1-1347)
```

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||        |||   |||:::   ||||||   |||           |||   ||||||
Db        973 GGGATGTCAGGGGAAGAACTCAGTGAATCTACTCCAGAGCCTCAAAAAGAAATTTCTGAG 1032

Qy         21 Ser---------ThrThrAspGluAsnGluAspAsp 29
              |||         |||   |||::::::|||||||||
Db       1033 TCATTGAGTGTCACCAGAGACCAGGATGAAGATGAT 1068
```

RESULT 17
US-10-995-561-13383/c
; Sequence 13383, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13383
; LENGTH: 88873
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(88873)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13383

Alignment Scores:

Pred. No.:	2.07e+04	Length:	88873
Score:	59.50	Matches:	14
Percent Similarity:	50.0%	Conservative:	6
Best Local Similarity:	35.0%	Mismatches:	19
Query Match:	26.0%	Indels:	1
DB:	8	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-995-561-13383 (1-88873)

```
Qy          2 AspPro---GlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||   ||||||   :::|||   :::   ||||||      |||   |||   :::
Db      13718 GACCCCACAGGAAGCAGACTCATCCACACGAGCCCCCCACCCCAGCCCCCCGCAGGGGAT 13659

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              :::   ||||||   |||                  ::::::      |||      |||
Db      13658 GCGCACACGGACATGAACCCCCAGCGTACCTCCACCCTCACAGGGAGCCCACCCTACCCC 13599
```

RESULT 18
US-11-226-657-39
; Sequence 39, Application US/11226657
; Publication No. US20060025574A1
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: PZ006G13AP1C1D2
; CURRENT APPLICATION NUMBER: US/11/226,657
; CURRENT FILING DATE: 2005-09-15
; PRIOR APPLICATION NUMBER: 10/062,831
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/690,454
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US98/10868
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: 1997-05-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-226-657-39

Alignment Scores:

Pred. No.:	244	Length:	1114
Score:	59.00	Matches:	11
Percent Similarity:	70.6%	Conservative:	1
Best Local Similarity:	64.7%	Mismatches:	5
Query Match:	25.8%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-226-657-39 (1-1114)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              |||||||||:::|||      |||   ||||||   |||||||||   |||
Db        309 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGC 359
```

RESULT 19
US-10-750-185-54635/c
; Sequence 54635, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 54635
; LENGTH: 1235
; TYPE: DNA
; ORGANISM: Bovine 19866880948445
US-10-750-185-54635

Alignment Scores:

Pred. No.:	271	Length:	1235
Score:	59.00	Matches:	9
Percent Similarity:	76.5%	Conservative:	4
Best Local Similarity:	52.9%	Mismatches:	4
Query Match:	25.8%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-750-185-54635 (1-1235)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              ||||||::::::   :::      :::||||||   ||||||||||||||||
Db        197 CCCGGAGCCCAGACTGTCCGGGGCCTGCCGCCCACAGGGCCGGAGGCCTCT 147
```

RESULT 20
US-10-750-623-54635/c

```
; Sequence 54635, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 54635
;   LENGTH: 1235
;   TYPE: DNA
;   ORGANISM: Bovine  19866880948445
US-10-750-623-54635

Alignment Scores:
Pred. No.:                271             Length:         1235
Score:                    59.00           Matches:        9
Percent Similarity:       76.5%           Conservative:   4
Best Local Similarity:    52.9%           Mismatches:     4
Query Match:              25.8%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-750-623-54635 (1-1235)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSer 19
              ||||||::::::   :::      :::|||||||   ||||||||||||||||
Db        197 CCCGGAGCCCAGACTGTCCGGGGCCTGCCGCCCACAGGGCCGGAGGCCTCT 147

RESULT 21
US-11-091-883-313
; Sequence 313, Application US/11091883
; Publication No. US20060024693A1
; GENERAL INFORMATION:
;  APPLICANT: CIBELLI, JOSE
;  APPLICANT:  FERNANDEZ, EMILIO O.
;  APPLICANT:  JORDAO DE MEGALHAES, GUILHERME
;  APPLICANT:  KOCABAS, ARIF
;  APPLICANT:  CROSBY, JAVIER A.
;  TITLE OF INVENTION: IDENTIFICATION OF GENES OR POLYPEPTIDES THE EXPRESSION OF WHICH
;  TITLE OF INVENTION:  CORRELATES TO FERTILITY, OVARIAN FUNCTION AND/OR FETAL/NEWBORN
;  TITLE OF INVENTION:  VIABILITY
;  FILE REFERENCE: 53942US
;  CURRENT APPLICATION NUMBER: US/11/091,883
;  CURRENT FILING DATE:  2005-03-29
;  PRIOR APPLICATION NUMBER: 60/556,875
;  PRIOR FILING DATE: 2004-03-29
;  NUMBER OF SEQ ID NOS: 513
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 313
;   LENGTH: 3835
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-091-883-313

Alignment Scores:
Pred. No.:                892             Length:         3835
Score:                    59.00           Matches:        15
Percent Similarity:       51.1%           Conservative:   8
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              25.8%           Indels:         6
DB:                       12              Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-11-091-883-313 (1-3835)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              ::::||||||:::|||      |||:::                  :::   |||||||:::
Db       2481 GAACCAGGTGCTGAATCTTCTGAGAACGCTGATGATCCCAACAAAGATACAAGTGAAAAC 2540

Qy         22 Thr---------ThrAspGluAsnGluAspAspIleGlnPheValSerGlu--------- 35
                          :::||||||||||:::||||||                 |||
Db       2541 GCAGATGGTCAAAGTGATGAGAACAAGGACGACTATACAATCCCAGATGAGTATAGAATT 2600

Qy         36 GlyProSerArgPro 40
              ||||||   :::|||
Db       2601 GGACCATATCAGCCC 2615

RESULT 22
US-10-821-234-258
; Sequence 258, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
;  APPLICANT: Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Andarmani, Susan
;  APPLICANT:  Tang, Y. Tom
;  TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;  FILE REFERENCE: 821A
;  CURRENT APPLICATION NUMBER: US/10/821,234
;  CURRENT FILING DATE:  2004-04-07
;  PRIOR APPLICATION NUMBER: US 60/462,047
;  PRIOR FILING DATE: 2003-04-07
;  NUMBER OF SEQ ID NOS: 1704
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 258
;   LENGTH: 3848
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-821-234-258

Alignment Scores:
Pred. No.:                895             Length:         3848
Score:                    59.00           Matches:        15
Percent Similarity:       51.1%           Conservative:   8
Best Local Similarity:    33.3%           Mismatches:     16
Query Match:              25.8%           Indels:         6
DB:                       8               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-821-234-258 (1-3848)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              ::::||||||:::|||      |||:::                  :::   |||||||:::
Db       2478 GAACCAGGTGCTGAATCTTCTGAGAACGCTGATGATCCCAACAAAGATACAAGTGAAAAC 2537

Qy         22 Thr---------ThrAspGluAsnGluAspAspIleGlnPheValSerGlu--------- 35
                          :::||||||||||:::||||||                 |||
Db       2538 GCAGATGGTCAAAGTGATGAGAACAAGGACGACTATACAATCCCAGATGAGTATAGAATT 2597

Qy         36 GlyProSerArgPro 40
              ||||||   :::|||
Db       2598 GGACCATATCAGCCC 2612

RESULT 23
US-11-128-061-932
; Sequence 932, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
;  APPLICANT: Melville, Mark W.
;  APPLICANT:  Charlebois, Timothy S.
;  APPLICANT:  Mounts, William M.
;  APPLICANT:  Hann, Louane E.
;  APPLICANT:  Sinacore, Martin S.
;  APPLICANT:  Leonard, Mark W.
;  APPLICANT:  Brown, Eugene L.
;  APPLICANT:  Miller, Christopher P.
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
;  TITLE OF INVENTION:  TO MONITOR GENE EXPRESSION
```

```
;  FILE REFERENCE: 01997.027701
;  CURRENT APPLICATION NUMBER: US/11/128,061
;  CURRENT FILING DATE:  2005-05-11
;  PRIOR APPLICATION NUMBER: US 60/570,425
;  PRIOR FILING DATE: 2004-05-11
;  NUMBER OF SEQ ID NOS: 7285
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 932
;   LENGTH: 4990
;   TYPE: DNA
;   ORGANISM: Cricetulus griseus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4888)..(4948)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-932

Alignment Scores:
Pred. No.:              1.18e+03        Length:         4990
Score:                  59.00           Matches:        19
Percent Similarity:     39.3%           Conservative:   5
Best Local Similarity:  31.1%           Mismatches:     19
Query Match:            25.8%           Indels:         18
DB:                     12              Gaps:           3

US-10-757-745-6_COPY_2_45 (1-44) x US-11-128-061-932 (1-4990)

Qy          1 GlyAspProGlySerGluIleIleGluSerValPro---ProAlaGlyProGluAlaSer 19
              |||   ||||||            :::      |||   ||||||||||||
Db       4067 GGTGCCCCCGGTGAGAGAGGGGAGCAGGGGCGACCAGGACCTGCTGGCCCTCGTGGGGAG 4126

Qy         20 GluSerThrThrAspGluAsnGluAspAspIle-------------------------- 30
              :::                 ||||||||||||
Db       4127 AAAGGAGAAGCTGCACTGACTGAGGATGACATTCGGGGCTTTGTACGTCAGGAGATGAGT 4186

Qy         31 --------------------GlnPheValSerGluGlyProSerArgProValLeuGlu 43
                                   ||||||::::::   |||   |||||||||:::
Db       4187 CAGCATTGTGCCTGCCAGGGCCAGTTTATTGCGTCTGGA---TCACGACCCCTCCCCAGC 4243

Qy         44 Tyr 44
              |||
Db       4244 TAC 4246

RESULT 24
US-11-128-049-932
; Sequence 932, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
;  APPLICANT: Melville, Mark W.
;  APPLICANT:  Charlebois, Timothy S.
;  APPLICANT:  Mounts, William M.
;  APPLICANT:  Hann, Louane E.
;  APPLICANT:  Sinacore, Martin S.
;  APPLICANT:  Leonard, Mark W.
;  APPLICANT:  Brown, Eugene L.
;  APPLICANT:  Miller, Christopher P.
;  TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
;  TITLE OF INVENTION:  MAKING AND USING SAME
;  FILE REFERENCE: 01997.027700
;  CURRENT APPLICATION NUMBER: US/11/128,049
;  CURRENT FILING DATE:  2005-05-11
;  PRIOR APPLICATION NUMBER: US 60/570,425
;  PRIOR FILING DATE: 2004-05-11
;  NUMBER OF SEQ ID NOS: 7285
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 932
;   LENGTH: 4990
;   TYPE: DNA
;   ORGANISM: Cricetulus griseus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4888)..(4948)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-128-049-932

Alignment Scores:
Pred. No.:              1.18e+03        Length:         4990
Score:                  59.00           Matches:        19
Percent Similarity:     39.3%           Conservative:   5
Best Local Similarity:  31.1%           Mismatches:     19
Query Match:            25.8%           Indels:         18
DB:                     12              Gaps:           3

US-10-757-745-6_COPY_2_45 (1-44) x US-11-128-049-932 (1-4990)

Qy          1 GlyAspProGlySerGluIleIleGluSerValPro---ProAlaGlyProGluAlaSer 19
              |||   ||||||            :::      |||   ||||||||||||
Db       4067 GGTGCCCCCGGTGAGAGAGGGGAGCAGGGGCGACCAGGACCTGCTGGCCCTCGTGGGGAG 4126

Qy         20 GluSerThrThrAspGluAsnGluAspAspIle-------------------------- 30
              :::                 ||||||||||||
Db       4127 AAAGGAGAAGCTGCACTGACTGAGGATGACATTCGGGGCTTTGTACGTCAGGAGATGAGT 4186

Qy         31 --------------------GlnPheValSerGluGlyProSerArgProValLeuGlu 43
                                   ||||||::::::   |||   |||||||||:::
Db       4187 CAGCATTGTGCCTGCCAGGGCCAGTTTATTGCGTCTGGA---TCACGACCCCTCCCCAGC 4243

Qy         44 Tyr 44
              |||
Db       4244 TAC 4246

RESULT 25
US-11-039-398-3/c
; Sequence 3, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Walke, D. Wade
;  APPLICANT:  Hilbun, Erin
;  APPLICANT:  Nepomnichy, Boris
;  APPLICANT:  Hu, Yi
;  TITLE OF INVENTION: Novel Human Proteases and
;  TITLE OF INVENTION:  Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0221-USA
;  CURRENT APPLICATION NUMBER: US/11/039,398
;  CURRENT FILING DATE:  2005-01-20
;  PRIOR APPLICATION NUMBER: US/09/938,330
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: US 60/227,104
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: US 60/233,796
;  PRIOR FILING DATE: 2000-09-19
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(894)
US-11-039-398-3

Alignment Scores:
Pred. No.:              226             Length:         894
Score:                  58.50           Matches:        12
Percent Similarity:     60.0%           Conservative:   6
Best Local Similarity:  40.0%           Mismatches:     11
Query Match:            25.5%           Indels:         1
DB:                     12              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-3 (1-894)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
              ||||||   ||||||:::|||                    ||||||   |||   ::::::
```

```
Db        378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy         31 GlnPheValSerGluGlyProSerArgPro 40
              :::   |||:::      |||:::||||||
Db        318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289

RESULT 26
US-11-039-398-1/c
; Sequence 1, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Walke, D. Wade
;  APPLICANT:  Hilbun, Erin
;  APPLICANT:  Nepomnichy, Boris
;  APPLICANT:  Hu, Yi
;  TITLE OF INVENTION: Novel Human Proteases and
;  TITLE OF INVENTION:  Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0221-USA
;  CURRENT APPLICATION NUMBER: US/11/039,398
;  CURRENT FILING DATE:  2005-01-20
;  PRIOR APPLICATION NUMBER: US/09/938,330
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: US 60/227,104
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: US 60/233,796
;  PRIOR FILING DATE: 2000-09-19
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1356
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1356)
US-11-039-398-1

Alignment Scores:
Pred. No.:                350             Length:         1356
Score:                    58.50           Matches:        12
Percent Similarity:       60.0%           Conservative:   6
Best Local Similarity:    40.0%           Mismatches:     11
Query Match:              25.5%           Indels:         1
DB:                       12              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-1 (1-1356)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
              ||||||  ||||||:::|||                ||||||   |||   ::::::
Db        378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy         31 GlnPheValSerGluGlyProSerArgPro 40
              :::   |||:::      |||:::||||||
Db        318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289

RESULT 27
US-11-039-398-5/c
; Sequence 5, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Walke, D. Wade
;  APPLICANT:  Hilbun, Erin
;  APPLICANT:  Nepomnichy, Boris
;  APPLICANT:  Hu, Yi
;  TITLE OF INVENTION: Novel Human Proteases and
;  TITLE OF INVENTION:  Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0221-USA
;  CURRENT APPLICATION NUMBER: US/11/039,398
;  CURRENT FILING DATE:  2005-01-20
;  PRIOR APPLICATION NUMBER: US/09/938,330
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: US 60/227,104
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: US 60/233,796
;  PRIOR FILING DATE: 2000-09-19
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 1461
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1461)
US-11-039-398-5

Alignment Scores:
Pred. No.:                379             Length:         1461
Score:                    58.50           Matches:        12
Percent Similarity:       60.0%           Conservative:   6
Best Local Similarity:    40.0%           Mismatches:     11
Query Match:              25.5%           Indels:         1
DB:                       12              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-5 (1-1461)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
              ||||||  ||||||:::|||                ||||||   |||   ::::::
Db        378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy         31 GlnPheValSerGluGlyProSerArgPro 40
              :::   |||:::      |||:::||||||
Db        318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289

RESULT 28
US-11-039-398-13/c
; Sequence 13, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Walke, D. Wade
;  APPLICANT:  Hilbun, Erin
;  APPLICANT:  Nepomnichy, Boris
;  APPLICANT:  Hu, Yi
;  TITLE OF INVENTION: Novel Human Proteases and
;  TITLE OF INVENTION:  Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0221-USA
;  CURRENT APPLICATION NUMBER: US/11/039,398
;  CURRENT FILING DATE:  2005-01-20
;  PRIOR APPLICATION NUMBER: US/09/938,330
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: US 60/227,104
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: US 60/233,796
;  PRIOR FILING DATE: 2000-09-19
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 3642
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3642)
US-11-039-398-13

Alignment Scores:
Pred. No.:                988             Length:         3642
Score:                    58.50           Matches:        12
Percent Similarity:       60.0%           Conservative:   6
Best Local Similarity:    40.0%           Mismatches:     11
```

```
Query Match:            25.5%          Indels:        1
DB:                     12             Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-13 (1-3642)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
              ||||||   |||||||:::|||                  ||||||   |||   ::::::
Db        378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy         31 GlnPheValSerGluGlyProSerArgPro 40
              :::   |||:::      |||:::||||||
Db        318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289

RESULT 29
US-11-039-398-11/c
; Sequence 11, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Walke, D. Wade
;  APPLICANT:  Hilbun, Erin
;  APPLICANT:  Nepomnichy, Boris
;  APPLICANT:  Hu, Yi
;  TITLE OF INVENTION: Novel Human Proteases and
;  TITLE OF INVENTION:  Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0221-USA
;  CURRENT APPLICATION NUMBER: US/11/039,398
;  CURRENT FILING DATE:  2005-01-20
;  PRIOR APPLICATION NUMBER: US/09/938,330
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: US 60/227,104
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: US 60/233,796
;  PRIOR FILING DATE: 2000-09-19
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 3651
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3651)
US-11-039-398-11

Alignment Scores:
Pred. No.:              990            Length:        3651
Score:                  58.50          Matches:       12
Percent Similarity:     60.0%          Conservative:  6
Best Local Similarity:  40.0%          Mismatches:    11
Query Match:            25.5%          Indels:        1
DB:                     12             Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-11 (1-3651)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
              ||||||   |||||||:::|||                  ||||||   |||   ::::::
Db        378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy         31 GlnPheValSerGluGlyProSerArgPro 40
              :::   |||:::      |||:::||||||
Db        318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289

RESULT 30
US-11-039-398-9/c
; Sequence 9, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Walke, D. Wade
;  APPLICANT:  Hilbun, Erin
;  APPLICANT:  Nepomnichy, Boris
;  APPLICANT:  Hu, Yi
;  TITLE OF INVENTION: Novel Human Proteases and
;  TITLE OF INVENTION:  Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0221-USA
;  CURRENT APPLICATION NUMBER: US/11/039,398
;  CURRENT FILING DATE:  2005-01-20
;  PRIOR APPLICATION NUMBER: US/09/938,330
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: US 60/227,104
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: US 60/233,796
;  PRIOR FILING DATE: 2000-09-19
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 3660
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3660)
US-11-039-398-9

Alignment Scores:
Pred. No.:              993            Length:        3660
Score:                  58.50          Matches:       12
Percent Similarity:     60.0%          Conservative:  6
Best Local Similarity:  40.0%          Mismatches:    11
Query Match:            25.5%          Indels:        1
DB:                     12             Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-9 (1-3660)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
              ||||||   |||||||:::|||                  ||||||   |||   ::::::
Db        378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy         31 GlnPheValSerGluGlyProSerArgPro 40
              :::   |||:::      |||:::||||||
Db        318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289

RESULT 31
US-11-039-398-7/c
; Sequence 7, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Walke, D. Wade
;  APPLICANT:  Hilbun, Erin
;  APPLICANT:  Nepomnichy, Boris
;  APPLICANT:  Hu, Yi
;  TITLE OF INVENTION: Novel Human Proteases and
;  TITLE OF INVENTION:  Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0221-USA
;  CURRENT APPLICATION NUMBER: US/11/039,398
;  CURRENT FILING DATE:  2005-01-20
;  PRIOR APPLICATION NUMBER: US/09/938,330
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: US 60/227,104
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: US 60/233,796
;  PRIOR FILING DATE: 2000-09-19
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 3669
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3669)
```

US-11-039-398-7

Alignment Scores:

Pred. No.:	995	Length:	3669
Score:	58.50	Matches:	12
Percent Similarity:	60.0%	Conservative:	6
Best Local Similarity:	40.0%	Mismatches:	11
Query Match:	25.5%	Indels:	1
DB:	12	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-7 (1-3669)

```
Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
               ||||||   ||||||:::|||                    ||||||   |||   ::::::
Db         378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy          31 GlnPheValSerGluGlyProSerArgPro 40
               :::   |||:::      |||:::||||||
Db         318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289
```

RESULT 32
US-11-039-398-17/c
; Sequence 17, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: Novel Human Proteases and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0221-USA
; CURRENT APPLICATION NUMBER: US/11/039,398
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3699
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3699)
US-11-039-398-17

Alignment Scores:

Pred. No.:	1e+03	Length:	3699
Score:	58.50	Matches:	12
Percent Similarity:	60.0%	Conservative:	6
Best Local Similarity:	40.0%	Mismatches:	11
Query Match:	25.5%	Indels:	1
DB:	12	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-17 (1-3699)

```
Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
               ||||||   ||||||:::|||                    ||||||   |||   ::::::
Db         378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy          31 GlnPheValSerGluGlyProSerArgPro 40
               :::   |||:::      |||:::||||||
Db         318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289
```

RESULT 33
US-11-039-398-15/c
; Sequence 15, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: Novel Human Proteases and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0221-USA
; CURRENT APPLICATION NUMBER: US/11/039,398
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3708
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3708)
US-11-039-398-15

Alignment Scores:

Pred. No.:	1.01e+03	Length:	3708
Score:	58.50	Matches:	12
Percent Similarity:	60.0%	Conservative:	6
Best Local Similarity:	40.0%	Mismatches:	11
Query Match:	25.5%	Indels:	1
DB:	12	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-15 (1-3708)

```
Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
               ||||||   ||||||:::|||                    ||||||   |||   ::::::
Db         378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy          31 GlnPheValSerGluGlyProSerArgPro 40
               :::   |||:::      |||:::||||||
Db         318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289
```

RESULT 34
US-11-039-398-21/c
; Sequence 21, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: Novel Human Proteases and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0221-USA
; CURRENT APPLICATION NUMBER: US/11/039,398
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 4.0

```
; SEQ ID NO 21
;   LENGTH: 3750
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3750)
US-11-039-398-21

Alignment Scores:
Pred. No.:              1.02e+03       Length:         3750
Score:                  58.50          Matches:        12
Percent Similarity:     60.0%          Conservative:   6
Best Local Similarity:  40.0%          Mismatches:     11
Query Match:            25.5%          Indels:         1
DB:                     12             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-21 (1-3750)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
               ||||||   ||||||:::|||                ||||||   |||   ::::::
Db         378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy          31 GlnPheValSerGluGlyProSerArgPro 40
               :::   |||:::      |||:::||||||
Db         318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289

RESULT 35
US-11-039-398-19/c
; Sequence 19, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
;  APPLICANT: Friddle, Carl Johan
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Walke, D. Wade
;  APPLICANT:  Hilbun, Erin
;  APPLICANT:  Nepomnichy, Boris
;  APPLICANT:  Hu, Yi
;  TITLE OF INVENTION: Novel Human Proteases and
;  TITLE OF INVENTION:  Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0221-USA
;  CURRENT APPLICATION NUMBER: US/11/039,398
;  CURRENT FILING DATE:  2005-01-20
;  PRIOR APPLICATION NUMBER: US/09/938,330
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: US 60/227,104
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: US 60/233,796
;  PRIOR FILING DATE: 2000-09-19
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
;   LENGTH: 3759
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(3759)
US-11-039-398-19

Alignment Scores:
Pred. No.:              1.02e+03       Length:         3759
Score:                  58.50          Matches:        12
Percent Similarity:     60.0%          Conservative:   6
Best Local Similarity:  40.0%          Mismatches:     11
Query Match:            25.5%          Indels:         1
DB:                     12             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-19 (1-3759)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
               ||||||   ||||||:::|||                ||||||   |||   ::::::
Db         378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy          31 GlnPheValSerGluGlyProSerArgPro 40
               :::   |||:::      |||:::||||||
Db         318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289

RESULT 36
US-10-624-932-25/c
; Sequence 25, Application US/10624932
; Publication No. US20040096877A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier, Raymond
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Spaderna, Steven
;  APPLICANT:  Shimkets, Richard
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Li, Li
;  APPLICANT:  Gusev, Vladimir
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Stone, David
;  APPLICANT:  Smithson, Glennda
;  TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-074 US
;  CURRENT APPLICATION NUMBER: US/10/624,932
;  CURRENT FILING DATE:  2003-07-21
;  PRIOR APPLICATION NUMBER: US/09/918,779
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 60/221,409
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: 60/222,840
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,752
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,762
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,770
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/223,769
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: 60/225,146
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/225,392
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/225,470
;  PRIOR FILING DATE: 2000-08-15
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
;   LENGTH: 4213
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-624-932-25

Alignment Scores:
Pred. No.:              1.15e+03       Length:         4213
Score:                  58.50          Matches:        12
Percent Similarity:     60.0%          Conservative:   6
Best Local Similarity:  40.0%          Mismatches:     11
Query Match:            25.5%          Indels:         1
DB:                     8              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-624-932-25 (1-4213)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
```

```
              |||||||   |||||||:::|||                        ||||||   |||   ::::::
Db        378 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 319

Qy         31 GlnPheValSerGluGlyProSerArgPro 40
              :::   |||:::       |||:::||||||
Db        318 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 289
```

RESULT 37
US-11-039-398-23/c
; Sequence 23, Application US/11039398
; Publication No. US20050267297A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: Novel Human Proteases and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0221-USA
; CURRENT APPLICATION NUMBER: US/11/039,398
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 4277
; TYPE: DNA
; ORGANISM: homo sapiens
US-11-039-398-23

Alignment Scores:

Pred. No.:	1.17e+03	Length:	4277
Score:	58.50	Matches:	12
Percent Similarity:	60.0%	Conservative:	6
Best Local Similarity:	40.0%	Mismatches:	11
Query Match:	25.5%	Indels:	1
DB:	12	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-039-398-23 (1-4277)

```
Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp---AspIle 30
              ||||||   |||||||:::|||                        ||||||   |||   ::::::
Db        575 CCTCCGATTGGGCCGCAGGCGCAAGTGCAGTTCCTTCCCGAAAACAGTGACATTGAAGTA 516

Qy         31 GlnPheValSerGluGlyProSerArgPro 40
              :::   |||:::       |||:::||||||
Db        515 GAGGGAGTGGCGCCCCACCCTGCCAGGCCA 486
```

RESULT 38
US-09-925-065A-16151/c
; Sequence 16151, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16151
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-16151

Alignment Scores:

Pred. No.:	178	Length:	530
Score:	57.50	Matches:	14
Percent Similarity:	48.8%	Conservative:	7
Best Local Similarity:	32.6%	Mismatches:	16
Query Match:	25.1%	Indels:	6
DB:	6	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-16151 (1-530)

```
Qy          1 GlyAspProGlySerGluIleIle-GluSerValProProAlaGlyProGluAlaSerGl 20
              ||||||||||||:::             |||:::|||||||||||||||||      :::
Db        326 GGGGATCCTGGTGCCTCCACCACGGCACTCCCTCCCTCCTGCCGGCCCTCGCTGTGCCCG 267

Qy         20 uSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPr 40
              :::          |||                 :::   :::|||      ||
Db        266 CGCTCGCCCAGACCCAGGC--------------CTTCTGGGTCAAGGAGGGGAGSCTCC 222

Qy         40 oValLeu 42
              |:::|||
Db        221 CCTTCTG 215
```

RESULT 39
US-09-925-065A-80196/c
; Sequence 80196, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80196
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-80196

Alignment Scores:

Pred. No.:	346	Length:	859
Score:	57.00	Matches:	16
Percent Similarity:	52.3%	Conservative:	7
Best Local Similarity:	36.4%	Mismatches:	15
Query Match:	24.9%	Indels:	6
DB:	6	Gaps:	2

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-80196 (1-859)

```
Qy          4 GlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThr 23
```

```
                ||||      :::|||   ||||||   |||||||||   :::   |||
Db        390 GGTCGGTCGGTGATACGGTCTGTGAGCCCAGCTGGTGCAGACTGCTCTGGGCAAGCTATG 331

Qy         24 AspGluAsnGluAsp----------------AspIleGlnPheValSerGluGlyProSer 38
               :::|||       |||               ::::::   |||:::   ||||||   |||
Db        330 GAAGAAGTAGGGGATGGGACCCACCTTCAGGAGCTGCTGTTCATCGTAGAGGGA---AGC 274

Qy         39 ArgProValLeu 42
                 |||:::|||
Db        273 CCTCCATTGCTT 262
```

RESULT 40
US-10-932-182A-76811
; Sequence 76811, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76811
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-76811

Alignment Scores:

Pred. No.:	486	Length:	1185
Score:	57.00	Matches:	10
Percent Similarity:	51.7%	Conservative:	5
Best Local Similarity:	34.5%	Mismatches:	14
Query Match:	24.9%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-932-182A-76811 (1-1185)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||:::||||||:::      :::|||   :::|||            |||      |||
Db       1063 GGCGAGCCAGGTGCGGGAGCAGTGGAAGAGATGCCATCAAAGAATGCGGAAGAGGAAGAA 1122

Qy         21 SerThrThrAspGluAsnGluAspAsp 29
                        |||   |||:::|||
Db       1123 GAGGAGGAGAGTGAAGAAGAAGAAGAT 1149
```

RESULT 41
US-10-932-182A-76811
; Sequence 76811, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76811
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-76811

Alignment Scores:

Pred. No.:	486	Length:	1185
Score:	57.00	Matches:	10
Percent Similarity:	51.7%	Conservative:	5
Best Local Similarity:	34.5%	Mismatches:	14
Query Match:	24.9%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-932-182A-76811 (1-1185)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||:::||||||:::      :::|||   :::|||            |||      |||
Db       1063 GGCGAGCCAGGTGCGGGAGCAGTGGAAGAGATGCCATCAAAGAATGCGGAAGAGGAAGAA 1122

Qy         21 SerThrThrAspGluAsnGluAspAsp 29
                        |||   |||:::|||
Db       1123 GAGGAGGAGAGTGAAGAAGAAGAAGAT 1149
```

RESULT 42
US-10-750-185-53105/c
; Sequence 53105, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53105
; LENGTH: 1225
; TYPE: DNA
; ORGANISM: Bovine 19866881074560
US-10-750-185-53105

Alignment Scores:

Pred. No.:	503	Length:	1225
Score:	57.00	Matches:	14
Percent Similarity:	44.2%	Conservative:	9
Best Local Similarity:	26.9%	Mismatches:	17
Query Match:	24.9%	Indels:	12
DB:	8	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-750-185-53105 (1-1225)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||       |||      :::   ||||||   |||   ::::::::::::   :::
Db        383 CCTGGACGTTGGATTACACGGGCCACGCCCCCTGAGGGCTCCGACAGCACAGACCCCAGC 324

Qy         23 ThrAspGluAsnGlu----------------------------------AspAspIle 30
              |||   |||   |||                                    |||:::
Db        323 ACCCAGGAGCCCGAGGTACCTCCAGAGCAAGATCTGGTAACCAGCACTGTGTCAGATGTG 264

Qy         31 GlnPheValSerGluGlyProSerArgProValLeu 42
                       |||   |||:::||||||:::
Db        263 GTGACCACGGTGATGGGCAGCTCCCAGCCTGTGGTG 228
```

RESULT 43
US-10-750-623-53105/c
; Sequence 53105, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.

```
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53105
;   LENGTH: 1225
;   TYPE: DNA
;   ORGANISM: Bovine  19866881074560
US-10-750-623-53105

Alignment Scores:
Pred. No.:                503             Length:         1225
Score:                    57.00           Matches:        14
Percent Similarity:       44.2%           Conservative:   9
Best Local Similarity:    26.9%           Mismatches:     17
Query Match:              24.9%           Indels:         12
DB:                       8               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-750-623-53105 (1-1225)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||       |||      :::   ||||||   |||   ::::::::::::   :::
Db        383 CCTGGACGTTGGATTACACGGGCCACGCCCCCTGAGGGCTCCGACAGCACAGACCCCAGC 324

Qy         23 ThrAspGluAsnGlu----------------------------------AspAspIle 30
              |||   |||   |||                                  |||:::
Db        323 ACCCAGGAGCCCGAGGTACCTCCAGAGCAAGATCTGGTAACCAGCACTGTGTCAGATGTG 264

Qy         31 GlnPheValSerGluGlyProSerArgProValLeu 42
                    |||   |||:::||||||:::
Db        263 GTGACCACGGTGATGGGCAGCTCCCAGCCTGTGGTG 228

RESULT 44
US-10-932-182A-4234
; Sequence 4234, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4234
;   LENGTH: 1725
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-4234

Alignment Scores:
Pred. No.:                720             Length:         1725
Score:                    57.00           Matches:        10
Percent Similarity:       60.7%           Conservative:   7
Best Local Similarity:    35.7%           Mismatches:     11
Query Match:              24.9%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-932-182A-4234 (1-1725)

Qy          7 IleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn 26
              :::   |||   :::||||||:::   ||||||      |||   :::      |||
Db        190 GTTAGTGAGTTCATACCGCCGTCCGTTCCTGAGGACCAAGAGATATCGTTAACAGAAGGA 249

Qy         27 GluAspAspIleGlnPheValSer 34
              |||   ::::::   |||:::|||
Db        250 GAGCAAAATTTGAATTTCATTTCC 273

RESULT 45
US-10-932-182A-4234
; Sequence 4234, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4234
;   LENGTH: 1725
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-4234

Alignment Scores:
Pred. No.:                720             Length:         1725
Score:                    57.00           Matches:        10
Percent Similarity:       60.7%           Conservative:   7
Best Local Similarity:    35.7%           Mismatches:     11
Query Match:              24.9%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-932-182A-4234 (1-1725)

Qy          7 IleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn 26
              :::   |||   :::||||||:::   ||||||      |||   :::      |||
Db        190 GTTAGTGAGTTCATACCGCCGTCCGTTCCTGAGGACCAAGAGATATCGTTAACAGAAGGA 249

Qy         27 GluAspAspIleGlnPheValSer 34
              |||   ::::::   |||:::|||
Db        250 GAGCAAAATTTGAATTTCATTTCC 273

RESULT 46
US-10-995-561-13341/c
; Sequence 13341, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13341
;   LENGTH: 28693
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-13341

Alignment Scores:
Pred. No.:                1.38e+04        Length:         28693
Score:                    57.00           Matches:        13
Percent Similarity:       68.0%           Conservative:   4
Best Local Similarity:    52.0%           Mismatches:     6
```

```
Query Match:            24.9%          Indels:        2
DB:                     8              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-995-561-13341 (1-28693)

Qy         10 SerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAsp 29
              |||:::|||          ||||||          :::|||:::|||||||||||||||
Db      15027 TCTCTCCCTGCCACAGCCCCTGAAGTA------ACCACTTCAGATGAAAATGAAGATTTG 14974

Qy         30 IleGlnPheValSer 34
              |||   |||:::|||
Db      14973 ATATTTTTCATTTCC 14959

RESULT 47
US-11-143-980-1/c
; Sequence 1, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
;  APPLICANT: He, Min
;  APPLICANT:  Hucul, John
;  APPLICANT:  Haltli, Bradley A.
;  APPLICANT:  Wagenaar, Melissa M.
;  APPLICANT:  Graziani, Edmund
;  APPLICANT:  Summers, Mia
;  APPLICANT:  Kulowski, Kerry
;  APPLICANT:  Pong, Kevin
;  TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
;  TITLE OF INVENTION:  Polyketide
;  FILE REFERENCE: AM-101426US
;  CURRENT APPLICATION NUMBER: US/11/143,980
;  CURRENT FILING DATE:  2005-06-03
;  PRIOR APPLICATION NUMBER: US 60/664,483
;  PRIOR FILING DATE: 2005-03-23
;  PRIOR APPLICATION NUMBER: US 60/576,895
;  PRIOR FILING DATE: 2004-06-03
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
;   LENGTH: 116856
;   TYPE: DNA
;   ORGANISM: Streptomyces sp.
US-11-143-980-1

Alignment Scores:
Pred. No.:              6.02e+04       Length:        116856
Score:                  57.00          Matches:       15
Percent Similarity:     47.7%          Conservative:  6
Best Local Similarity:  34.1%          Mismatches:    13
Query Match:            24.9%          Indels:        10
DB:                     12             Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x US-11-143-980-1 (1-116856)

Qy          7 IleIleGluSerValProProAla--------------GlyProGluAlaSerGluSer 21
              :::|||   :::   |||||||||              |||         :::|||
Db      60041 GTCATCACCGCCACACCGCCCGCCAGCGCCAGTGTGCTTTCGCCCCGGCGCAGGGACTCA 59982

Qy         22 --------------ThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGly 36
                           |||   |||:::|||     :::     |||:::   |||
Db      59981 CACGCCCAGTGCACGCTGACCAGCGAGGACGAGCACGCGGTGTCCACGGTGACGGCGGGG 59922

Qy         37 ProSerArgPro 40
              ||||||||||||
Db      59921 CCTTCGAGGCCG 59910

RESULT 48
US-11-121-086-5/c
; Sequence 5, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
;   LENGTH: 153376
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-5

Alignment Scores:
Pred. No.:              8.02e+04       Length:        153376
Score:                  57.00          Matches:       14
Percent Similarity:     42.2%          Conservative:  5
Best Local Similarity:  31.1%          Mismatches:    20
Query Match:            24.9%          Indels:        6
DB:                     12             Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-121-086-5 (1-153376)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyPro---------------- 16
              |||||||||      :::      |||||||||||||||
Db      22761 CCCGGCTCCAGCGGAGTCACGGATGCCCCTCCGGCAGGGCCCATGCGGAAACGCCAACAC 22702

Qy         17 GluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGly 36
              ||||||         ::::::      |||   :::         |||
Db      22701 CGCCAGAGTGAGCTCTGCCCCGAACAAAGAGGCGACTGTGTCGACAGAACCAGTCCGGCT 22642

Qy         37 ProSerArgProVal 41
              |||       |||:::
Db      22641 CCCAGAGGCCCCCTC 22627

RESULT 49
US-09-925-065A-16150/c
; Sequence 16150, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16150
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-16150

Alignment Scores:
Pred. No.:              244            Length:        530
Score:                  56.50          Matches:       14
Percent Similarity:     48.8%          Conservative:  7
Best Local Similarity:  32.6%          Mismatches:    16
Query Match:            24.7%          Indels:        6
DB:                     6              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-16150 (1-530)
```

```
Qy          1 GlyAspProGlySerGluIleIle-GluSerValProProAlaGlyProGluAlaSerGl 20
              ||||||||||||:::          |||:::|||||||||||||||      :::
Db        326 GGGGATCCTGGTGCCTCCACCACGGCACTCCCTCCCTCCTGCCGGCCCTCGCTGTGCCCG 267

Qy         20 uSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPr 40
                :::     |||                          :::   :::|||         ||
Db        266 CGCTCGCCCAGACCCAGGC--------------CTTCTGGGTCAAGGAGGGGAGCCTCC 222

Qy         40 oValLeu 42
                |:::|||
Db        221 CCTTCTG 215

RESULT 50
US-09-925-065A-748950
; Sequence 748950, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 748950
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-748950

Alignment Scores:
Pred. No.:                295           Length:          636
Score:                    56.50         Matches:         12
Percent Similarity:       50.0%         Conservative:    3
Best Local Similarity:    40.0%         Mismatches:      4
Query Match:              24.7%         Indels:          11
DB:                       6             Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-748950 (1-636)

Qy          11 ValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIle 30
               :::|||||||:::   |||:::      |||   |||
Db        403 ATCCCTCCTTCTTCACCTCAGGTTTGCGAGTTAACC----------------------- 438

Qy          31 GlnPheValSerGluGlyProSerArgPro 40
                  ||||||||||||||||||||
Db        439 ---------AGTGAGGGGCCTTCCAGGCCT 459

RESULT 51
US-09-925-065A-749153
; Sequence 749153, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 749153
;   LENGTH: 641
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-749153

Alignment Scores:
Pred. No.:                298           Length:          641
Score:                    56.50         Matches:         12
Percent Similarity:       50.0%         Conservative:    3
Best Local Similarity:    40.0%         Mismatches:      4
Query Match:              24.7%         Indels:          11
DB:                       6             Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-749153 (1-641)

Qy          11 ValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIle 30
               :::|||||||:::   |||:::      |||   |||
Db        404 ATCCCTCCTTCTTCACCTCAGGTTTGCGAGTTAACC----------------------- 439

Qy          31 GlnPheValSerGluGlyProSerArgPro 40
                  ||||||||||||||||||||
Db        440 ---------AGTGAGGGGCCTTCCAGGCCT 460

RESULT 52
US-09-925-065A-821963
; Sequence 821963, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 821963
;   LENGTH: 641
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-821963

Alignment Scores:
Pred. No.:                298           Length:          641
Score:                    56.50         Matches:         12
Percent Similarity:       50.0%         Conservative:    3
Best Local Similarity:    40.0%         Mismatches:      4
Query Match:              24.7%         Indels:          11
DB:                       6             Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-821963 (1-641)
```

```
Qy         11 ValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIle 30
              :::||||||:::   |||:::      |||   |||
Db        404 ATCCCTCCTTCTTCACCTCAGGTTTGCGAGTTAACC---------------------- 439

Qy         31 GlnPheValSerGluGlyProSerArgPro 40
                        ||||||||||||||||||||
Db        440 ---------AGTGAGGGGCCTTCCAGGCCT 460

RESULT 53
US-11-096-568A-6495
; Sequence 6495, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6495
;   LENGTH: 1507
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1507)
;   OTHER INFORMATION: Ceres Seq. ID no. 14315403
US-11-096-568A-6495

Alignment Scores:
Pred. No.:              731          Length:          1507
Score:                  56.50        Matches:         12
Percent Similarity:     61.1%        Conservative:    10
Best Local Similarity:  33.3%        Mismatches:      11
Query Match:            24.7%        Indels:          3
DB:                     9            Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-096-568A-6495 (1-1507)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp 28
              :::|||:::||||||:::   |||::::::      :::   |||   |||
Db        295 CAATCCCTCCCCCCTTCCTCGCCGGACTCGCAACGCGCTTGGACGCGCGAG--------- 345

Qy         29 AspIleGlnPheValSerGluGlyProSerArgProValLeuGluTyr 44
              ||||||::::::|||   :::   ||||||          :::   |||
Db        346 GACATCCGCTACGTGAAGGACGTGCCGTCGATCGCGCCGGTGGCGTAC 393

RESULT 54
US-10-330-773-340
; Sequence 340, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 340
;   LENGTH: 2176
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-330-773-340

Alignment Scores:
Pred. No.:              1.07e+03     Length:          2176
Score:                  56.50        Matches:         13
Percent Similarity:     51.4%        Conservative:    6
Best Local Similarity:  35.1%        Mismatches:      17
Query Match:            24.7%        Indels:          1
DB:                     7            Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-330-773-340 (1-2176)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||   :::      |||      |||:::|||
Db          2 CCGGGTTCCCCAGTGCCAGCCAGCGCGGCCCCCTCGGGGCTCCGGCAGCAGCGCCGGCAT 61

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArg 39
                           |||::::::|||::::::||||||||||||   |||
Db         62 GTC---GTCCGGCACGATGAAGTTCAATGGCTATCTGAGGGTCCGCATCGG 109

RESULT 55
US-10-330-773-339
; Sequence 339, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 339
;   LENGTH: 215126
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(215126)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-339

Alignment Scores:
Pred. No.:              1.34e+05     Length:          215126
Score:                  56.50        Matches:         13
Percent Similarity:     51.4%        Conservative:    6
Best Local Similarity:  35.1%        Mismatches:      17
Query Match:            24.7%        Indels:          1
DB:                     7            Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-330-773-339 (1-215126)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||   :::      |||      |||:::|||
Db      10002 CCGGGTTCCCCAGTGCCAGCCAGCGCGGCCCCCTCGGGGCTCCGGCAGCAGCGCCGGCAT 10061

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArg 39
                           |||::::::|||::::::||||||||||||   |||
Db      10062 GTC---GTCCGGCACGATGAAGTTCAATGGCTATCTGAGGGTCCGCATCGG 10109

RESULT 56
US-10-330-773-122/c
; Sequence 122, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 122
;   LENGTH: 348101
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(348101)
```

```
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-122

Alignment Scores:
Pred. No.:                2.2e+05        Length:         348101
Score:                    56.50          Matches:        16
Percent Similarity:       51.3%          Conservative:   4
Best Local Similarity:    41.0%          Mismatches:     14
Query Match:              24.7%          Indels:         5
DB:                       7              Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-10-330-773-122 (1-348101)

Qy          4 GlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThr 23
              |||::::::  ::   ||||||||||||||   |||           |||:::
Db     158338 GGAAATGATGGCCTGTCCCTTGTGCCCCCTGCTGGAAGTGAG----------ACCAGT 158291

Qy         24 AspGluAsnGluAspAspIleGlnPheValSerGluGlyPro---SerArgProVal 41
              |||         |||               |||  ||||||     |||||||||
Db     158290 GATGCTCGGAGGGACAGTGGTCTCACTGCAAGCAGGGGTCCAAGGGACAGGCCCGTC 158234

RESULT 57
US-09-925-065A-306831/c
; Sequence 306831, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 306831
;   LENGTH: 541
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-306831

Alignment Scores:
Pred. No.:                291            Length:         541
Score:                    56.00          Matches:        14
Percent Similarity:       51.5%          Conservative:   3
Best Local Similarity:    42.4%          Mismatches:     16
Query Match:              24.5%          Indels:         0
DB:                       6              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-306831 (1-541)

Qy          8 IleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGlu 27
              |||   |||   |||              |||:::|||           :::
Db        333 ATTATTTCATTTCCTGCTTTTAATTTCCTAGTTTCTGATTCTGMAGAGCTCCACTTTAAA 274

Qy         28 AspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||  ||||||:::  ||||||||||||  |||
Db        273 GATGACAGTCAATTCATAAAGGAAGGACCTAGTAGTCCT 235

RESULT 58
US-09-925-065A-7171/c
; Sequence 7171, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7171
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-7171

Alignment Scores:
Pred. No.:                339            Length:         625
Score:                    56.00          Matches:        13
Percent Similarity:       50.0%          Conservative:   7
Best Local Similarity:    32.5%          Mismatches:     8
Query Match:              24.5%          Indels:         12
DB:                       6              Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-7171 (1-625)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||||||||   :::      |||:::|||||||:::   |||::::::|||:::
Db        612 GGCAGACCTGGGTCAACACTCAATCACTCCCTCCCTCCCTCCCTCCCTCAATCATCAGAC 553

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                                                  :::|||      ||||||   |||
Db        552 ----------------------------------ATTTCCTGGGCACCTTCTATGCCC 529

RESULT 59
US-09-925-065A-7172/c
; Sequence 7172, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7172
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-7172

Alignment Scores:
Pred. No.:                339            Length:         625
```

```
Score:                 56.00          Matches:       13
Percent Similarity:    50.0%          Conservative:  7
Best Local Similarity: 32.5%          Mismatches:    8
Query Match:           24.5%          Indels:        12
DB:                    6              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-7172 (1-625)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||||||||   :::      |||:::||||||:::   |||::::::|||:::
Db        612 GGCAGACCTGGGTCAACACTCAATCACTCCCTCCCTCCCTCCCTCCCTCAATCATCAGAC 553

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                                                 :::|||      ||||||   |||
Db        552 ----------------------------------ATTTCCTGGGCACCTTCTATGCCC 529

RESULT 60
US-09-925-065A-7173/c
; Sequence 7173, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7173
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-7173

Alignment Scores:
Pred. No.:             339            Length:        625
Score:                 56.00          Matches:       13
Percent Similarity:    50.0%          Conservative:  7
Best Local Similarity: 32.5%          Mismatches:    8
Query Match:           24.5%          Indels:        12
DB:                    6              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-7173 (1-625)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||||||||   :::      |||:::||||||:::   |||::::::|||:::
Db        612 GGCAGACCTGGGTCAACACTCAATCACTCCCTCCCTCCCTCCCTCCCTCAATCATCAGAC 553

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                                                 :::|||      ||||||   |||
Db        552 ----------------------------------ATTTCCTGGGCACCTTCTATGCCC 529

RESULT 61
US-09-925-065A-82701/c
; Sequence 82701, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 82701
;   LENGTH: 635
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-82701

Alignment Scores:
Pred. No.:             345            Length:        635
Score:                 56.00          Matches:       13
Percent Similarity:    50.0%          Conservative:  7
Best Local Similarity: 32.5%          Mismatches:    8
Query Match:           24.5%          Indels:        12
DB:                    6              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-82701 (1-635)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||||||||   :::      |||:::||||||:::   |||::::::|||:::
Db        622 GGCAGACCTGGGTCAACACTCAATCACTCCCTCCCTCCCTCCCTCCCTCAATCATCAGAC 563

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                                                 :::|||      ||||||   |||
Db        562 ----------------------------------ATTTCCTGGGCACCTTCTATGCCC 539

RESULT 62
US-09-925-065A-82702/c
; Sequence 82702, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 82702
;   LENGTH: 635
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-82702

Alignment Scores:
Pred. No.:             345            Length:        635
Score:                 56.00          Matches:       13
Percent Similarity:    50.0%          Conservative:  7
Best Local Similarity: 32.5%          Mismatches:    8
Query Match:           24.5%          Indels:        12
DB:                    6              Gaps:          1
```

```
US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-82702 (1-635)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||||||||   :::     |||:::||||||:::   |||::::::|||:::
Db        622 GGCAGACCTGGGTCAACACTCAATCACTCCCTCCCTCCCTCCCTCAATCATCAGAC 563

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                                          :::|||     ||||||   |||
Db        562 -----------------------------------ATTTCCTGGGCACCTTCTATGCCC 539

RESULT 63
US-11-179-624-2
; Sequence 2, Application US/11179624
; Publication No. US20060040372A1
; GENERAL INFORMATION:
;  APPLICANT: Sakaguchi, Nobuo
;  APPLICANT:  Kuwahara, Kazuhiko
;  TITLE OF INVENTION: GANP Protein
;  FILE REFERENCE: 050208-0018
;  CURRENT APPLICATION NUMBER: US/11/179,624
;  CURRENT FILING DATE:  2005-07-13
;  PRIOR APPLICATION NUMBER: PCT/JP99/04634
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: 09/914,272
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: 10/747,133
;  PRIOR FILING DATE: 2003-12-30
;  PRIOR APPLICATION NUMBER: 47035/1999
;  PRIOR FILING DATE: 1999-02-24
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
;   LENGTH: 6429
;   TYPE: DNA
;   ORGANISM: Mouse
US-11-179-624-2

Alignment Scores:
Pred. No.:              3.92e+03        Length:         6429
Score:                  56.00           Matches:        13
Percent Similarity:     48.6%           Conservative:   4
Best Local Similarity:  37.1%           Mismatches:     18
Query Match:            24.5%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-179-624-2 (1-6429)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              |||   |||       |||       |||||||||:::|||||||||
Db       2153 GACTGTGGCAGACACATCTGAGGAGAAGTACCGCCTTCTGGACCAGAGAGACCGCATCAT 2212

Qy         22 ThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGly 36
                        :::|||:::|||       |||::|||
Db       2213 GCGGCAAGCTCGAGTGAAGAGGACGGACCTGGACAAAGCCAGGGC 2257

RESULT 64
US-10-330-773-551
; Sequence 551, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 551
;   LENGTH: 62444
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(62444)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-551

Alignment Scores:
Pred. No.:              4.27e+04        Length:         62444
Score:                  56.00           Matches:        11
Percent Similarity:     60.7%           Conservative:   6
Best Local Similarity:  39.3%           Mismatches:     11
Query Match:            24.5%           Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-330-773-551 (1-62444)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
              |||||||||||||||:::||||||      :::     :::       :::|||
Db       9992 CCGCCTGCAGGTCCCGACGCGTCTCCGCCGTCGCCGTCGCCGTCGCCGTCCGCCGTACAG 10051

Qy         32 PheValSerGluGlyProSerArg 39
                  |||   ||||||::::::
Db      10052 CCGAATAGCCGAGGACCCACGCAG 10075

RESULT 65
US-10-932-182A-78518
; Sequence 78518, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78518
;   LENGTH: 321
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-78518

Alignment Scores:
Pred. No.:              197             Length:         321
Score:                  55.50           Matches:        13
Percent Similarity:     43.5%           Conservative:   7
Best Local Similarity:  28.3%           Mismatches:     9
Query Match:            24.2%           Indels:         17
DB:                     7               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-932-182A-78518 (1-321)

Qy          4 GlySerGluIleIleGluSerValProProAlaGlyPro---------------------- 16
              |||:::|||   :::   :::|||||||   |||||||||
Db        169 GGTAACGAAAAGTTGGCTGCTGTTCCAGCTGCTGGTCCAGCTTCTGCTGGCGGTGCTGCT 228

Qy         17 ----------------------------GluAlaSerGluSerThrThrAspGluAsn 26
                                          |||     |||         :::|||:::
Db        229 GCTGCCTCTGGTGATGCTGCTGCTGAAGAAGAAAAGGAAGAAGAAGCCGCTGAAGAATCT 288

Qy         27 GluAspAspIleGlnPhe 32
              :::||||||:::   |||
Db        289 GACGACGACATGGGTTTC 306

RESULT 66
US-10-932-182A-78518
; Sequence 78518, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78518
;   LENGTH: 321
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-78518

Alignment Scores:
Pred. No.:              197          Length:         321
Score:                  55.50        Matches:        13
Percent Similarity:     43.5%        Conservative:   7
Best Local Similarity:  28.3%        Mismatches:     9
Query Match:            24.2%        Indels:         17
DB:                     7            Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-932-182A-78518 (1-321)

Qy          4 GlySerGluIleIleGluSerValProProAlaGlyPro---------------------- 16
              |||:::|||   :::   :::||||||   |||||||||
Db        169 GGTAACGAAAAGTTGGCTGCTGTTCCAGCTGCTGGTCCAGCTTCTGCTGGCGGTGCTGCT 228

Qy         17 ------------------------------GluAlaSerGluSerThrThrAspGluAsn 26
                                            |||      |||         :::|||:::
Db        229 GCTGCCTCTGGTGATGCTGCTGCTGAAGAAGAAAAGGAAGAAGAAGCCGCTGAAGAATCT 288

Qy         27 GluAspAspIleGlnPhe 32
              :::||||||:::   |||
Db        289 GACGACGACATGGGTTTC 306

RESULT 67
US-09-925-065A-123746
; Sequence 123746, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 123746
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-123746

Alignment Scores:
Pred. No.:              401          Length:         632
Score:                  55.50        Matches:        12
Percent Similarity:     46.2%        Conservative:   6
Best Local Similarity:  30.8%        Mismatches:     12
Query Match:            24.2%        Indels:         9
DB:                     6            Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-09-925-065A-123746 (1-632)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn-------------- 26
              ||||||:::     :::||||||   |||      :::|||
Db        290 CCTCCTAGCCCAAGACAGGCGTCCAGCAGTCTGGGTGAGGAGAGGATTAGGGGTAAATGA 349

Qy         27 -----------GluAspAspIleGlnPheValSerGluGlyProSerArgProVal 41
                         :::||||||      |||      :::|||      |||||||:::
Db        350 ATCTTTTGAGCACAAGATGACTGCTCCTTTGCTGGGGATGGGGCTGGGAGGCCAATC 406

RESULT 68
US-10-932-182A-3520
; Sequence 3520, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3520
;   LENGTH: 1062
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-3520

Alignment Scores:
Pred. No.:              692          Length:         1062
Score:                  55.50        Matches:        18
Percent Similarity:     35.8%        Conservative:   6
Best Local Similarity:  26.9%        Mismatches:     12
Query Match:            24.2%        Indels:         31
DB:                     7            Gaps:           3

US-10-757-745-6_COPY_2_45 (1-44) x US-10-932-182A-3520 (1-1062)

Qy          3 ProGlySerGluIleIleGluSerValProPro-------------------------- 13
              |||      :::::::::      ||||||
Db        433 CCATCCAAGGATTTAATGTTCAAGTTTCCACCAGGGCAGAATGAGCAGACTGACAAAGAA 492

Qy         14 ---------------------------------------AlaGlyProGluAlaSerGlu 20
                                                     :::   |||      ||||||
Db        493 CTAATGCACGACTGGAAAATTGCGCTTGAGAAATTTTTATCGAGTCCAAGTGGAAGTGAA 552

Qy         21 SerThrThr--------------------AspGluAsnGluAspAspIleGlnPheVal 33
              |||   |||                  ||||||:::|||||||||:::      |||
Db        553 AGCGTTACTACAGGCAGTGATTATGACGAAGATGAAGACGAGGATGATTTGATTGTAGTC 612

Qy         34 SerGlu------GlyProSer 38
              |||      |||||||||
Db        613 GATGAAAAGGCGGGTCCTTCT 633

RESULT 69
US-10-932-182A-3520
; Sequence 3520, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
```

```
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3520
;   LENGTH: 1062
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-3520

Alignment Scores:
Pred. No.:              692          Length:          1062
Score:                  55.50        Matches:         18
Percent Similarity:     35.8%        Conservative:    6
Best Local Similarity:  26.9%        Mismatches:      12
Query Match:            24.2%        Indels:          31
DB:                     7            Gaps:            3

US-10-757-745-6_COPY_2_45 (1-44) x US-10-932-182A-3520 (1-1062)

Qy          3 ProGlySerGluIleIleGluSerValProPro---------------------------- 13
              |||     :::::::::             ||||||
Db        433 CCATCCAAGGATTTAATGTTCAAGTTTCCACCAGGGCAGAATGAGCAGACTGACAAAGAA 492

Qy         14 ------------------------------------AlaGlyProGluAlaSerGlu 20
                                                  :::   |||      ||||||
Db        493 CTAATGCACGACTGGAAAATTGCGCTTGAGAAATTTTTATCGAGTCCAAGTGGAAGTGAA 552

Qy         21 SerThrThr--------------------AspGluAsnGluAspAspIleGlnPheVal 33
              |||   |||                  ||||||:::|||||||||:::      |||
Db        553 AGCGTTACTACAGGCAGTGATTATGACGAAGATGAAGACGAGGATGATTTGATTGTAGTC 612

Qy         34 SerGlu------GlyProSer 38
                |||      |||||||||
Db        613 GATGAAAAGGCGGGTCCTTCT 633

RESULT 70
US-11-000-688-366
; Sequence 366, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
;  APPLICANT: BERTUCCI, Francois
;  APPLICANT:  HOULGATTE, Remi
;  APPLICANT:  BIRNBAUM, Daniel
;  TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
;  FILE REFERENCE: 1423-R-03
;  CURRENT APPLICATION NUMBER: US/11/000,688
;  CURRENT FILING DATE:  2004-12-01
;  PRIOR APPLICATION NUMBER: US 60/525,987
;  PRIOR FILING DATE: 2003-12-01
;  NUMBER OF SEQ ID NOS: 1596
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 366
;   LENGTH: 2431
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial sequences:primer
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(2431)
;   OTHER INFORMATION: sp110 nuclear body protein(SP110) gene.
US-11-000-688-366

Alignment Scores:
Pred. No.:              1.65e+03     Length:          2431
Score:                  55.50        Matches:         14
Percent Similarity:     50.0%        Conservative:    3
Best Local Similarity:  41.2%        Mismatches:      16
Query Match:            24.2%        Indels:          1
DB:                     12           Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-000-688-366 (1-2431)

Qy         10 SerValProProAlaGlyProGluAlaSerGlu---SerThrThrAspGluAsnGluAsp 28
              |||       |||       |||       ||||||      |||:::   :::      |||
Db        543 AGCTGTTCACCCTGTGCGCCAAGAGTCAGTGAGCCTGGAACATCCTCCCAGCAAAGCGAT 602

Qy         29 AspIleGlnPheValSerGluGlyProSerArgProValLeu 42
              :::|||       |||      ||||||   |||||||||
Db        603 GAGATCCTGAGTGAGTCGCCCAGCCCATCTGACCCTGTCCTG 644

RESULT 71
US-10-750-185-24916/c
; Sequence 24916, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 24916
;   LENGTH: 2776
;   TYPE: DNA
;   ORGANISM: Bovine  19866881045177
US-10-750-185-24916

Alignment Scores:
Pred. No.:              1.9e+03      Length:          2776
Score:                  55.50        Matches:         13
Percent Similarity:     51.4%        Conservative:    6
Best Local Similarity:  35.1%        Mismatches:      17
Query Match:            24.2%        Indels:          1
DB:                     8            Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-750-185-24916 (1-2776)

Qy          5 SerGluIleIleGluSerValProProAla---GlyProGluAlaSerGluSerThrThr 23
              |||   :::    :::||||||||||   |||         |||:::
Db       2717 TCAGGCCTCTTTCCTGGGATCCCTCCTGCCCCTGGCCGCAGAGTCAGTGACCTCGTCTGT 2658

Qy         24 AspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||:::      ||||||:::      :::|||||      ||||||
Db       2657 GACCAGCGTTACTCAGACATCGAGCCTGCCACTGAGGGAGAGGTCAGGCCC 2607

RESULT 72
US-10-750-623-24916/c
; Sequence 24916, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
```

```
; SEQ ID NO 24916
;   LENGTH: 2776
;   TYPE: DNA
;   ORGANISM: Bovine  19866881045177
US-10-750-623-24916

Alignment Scores:
Pred. No.:              1.9e+03       Length:         2776
Score:                  55.50         Matches:        13
Percent Similarity:     51.4%         Conservative:   6
Best Local Similarity:  35.1%         Mismatches:     17
Query Match:            24.2%         Indels:         1
DB:                     8             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-750-623-24916 (1-2776)

Qy          5 SerGluIleIleGluSerValProProAla---GlyProGluAlaSerGluSerThrThr 23
              |||   :::         :::|||||||||   |||       |||:::
Db       2717 TCAGGCCTCTTTCCTGGGATCCCTCCTGCCCCTGGCCGCAGAGTCAGTGACCTCGTCTGT 2658

Qy         24 AspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              |||:::         ||||||:::      :::||||||      ||||||
Db       2657 GACCAGCGTTACTCAGACATCGAGCCTGCCACTGAGGGAGAGGTCAGGCCC 2607

RESULT 73
US-11-182-016-4
; Sequence 4, Application US/11182016
; Publication No. US20060019294A1
; GENERAL INFORMATION:
;  APPLICANT: SUGEN, INC.
;  TITLE OF INVENTION: TYROSINE KINASE SUBSTRATE (TKS) PROTEINS
;  FILE REFERENCE: 038602/0102
;  CURRENT APPLICATION NUMBER: US/11/182,016
;  CURRENT FILING DATE:  2005-07-15
;  PRIOR APPLICATION NUMBER: US/09/958,359
;  PRIOR FILING DATE: 2002-02-05
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 2787
;   TYPE: DNA
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: Tks 202
US-11-182-016-4

Alignment Scores:
Pred. No.:              1.91e+03      Length:         2787
Score:                  55.50         Matches:        15
Percent Similarity:     47.6%         Conservative:   5
Best Local Similarity:  35.7%         Mismatches:     15
Query Match:            24.2%         Indels:         7
DB:                     12            Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-11-182-016-4 (1-2787)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              ||||||         :::   |||           ||||||||||||      :::|||
Db       535 GATCCCGCGCCCGCCGTGTCTGAG-----------GCGGGGCCGGAGACACACGACAGC 582

Qy         22 ThrThrAspGluAsnGluAspAspIleGln---------PheValSerGluGlyProSer 38
                ||||||   ::::::|||                     |||   ||||||:::
Db       583 GAGGAGGACGAAGGCGACGAGGACGCGGCGGCGGCGCTGGCCACCTCGTGCGGCCCGGCG 642

Qy         39 ArgPro 40
              |||
Db       643 AGTCCC 648

RESULT 74
US-11-127-832-27
; Sequence 27, Application US/11127832
; Publication No. US20060008884A1
; GENERAL INFORMATION:
;  APPLICANT: Hearing, Patrick
;  APPLICANT:  Bahou, Wadie
;  APPLICANT:  Sandalon, Ziv
;  APPLICANT:  Gnatenko, Dmitri
;  TITLE OF INVENTION: Adenoviral Vectors
;  FILE REFERENCE: STONYB-04970
;  CURRENT APPLICATION NUMBER: US/11/127,832
;  CURRENT FILING DATE:  2005-05-12
;  PRIOR APPLICATION NUMBER: US/09/782,378
;  PRIOR FILING DATE: 2001-02-12
;  PRIOR APPLICATION NUMBER: 60/237,747
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
;   LENGTH: 34214
;   TYPE: DNA
;   ORGANISM: Human adenovirus type 40
US-11-127-832-27

Alignment Scores:
Pred. No.:              2.65e+04      Length:         34214
Score:                  55.50         Matches:        12
Percent Similarity:     48.8%         Conservative:   8
Best Local Similarity:  29.3%         Mismatches:     12
Query Match:            24.2%         Indels:         9
DB:                     12            Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-127-832-27 (1-34214)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGlu--------- 25
              :::   :::|||||:::   ||||||      |||:::|||       |||
Db        786 GACGGTTTGCCTCCTAGCGATCCTGAAACTGATGAGGCTACAGAGGCGGAAGAAGAGGCG 845

Qy         26 -----------------AsnGluAspAspIleGlnPheValSerGluGlyProSerArg 39
                              ||||||::::::   :::   |||   :::   |||
Db        846 GCTATGCCGACTTATGTGAATGAAAATGAAAATGAGCTGGTGCTGGACTGTCCAGAGAAC 905

Qy         40 Pro 40
              |||
Db        906 CCT 908

RESULT 75
US-11-124-368A-13139
; Sequence 13139, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  James J. Devlin
;  APPLICANT:  May Luke
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Vascular Diseases, Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001524
;  CURRENT APPLICATION NUMBER: US/11/124,368A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,845
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/625,936
;  PRIOR FILING DATE: 2004-11-09
;  NUMBER OF SEQ ID NOS: 21112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13139
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-368A-13139

Alignment Scores:
Pred. No.:              141           Length:         201
Score:                  55.00         Matches:        12
Percent Similarity:     62.5%         Conservative:   3
Best Local Similarity:  50.0%         Mismatches:     9
Query Match:            24.0%         Indels:         0
```

DB: 12 Gaps: 0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-124-368A-13139 (1-201)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||||||:::   :::   ||||||      |||   |||      |||   ||||||
Db         69 CCTGGATCCCAGACCGTCGGTTCTGTATCTCTYGCTGTTCCCAACCAAAGCCACAGCACC 128

Qy         23 ThrAspGluAsn 26
              :::   ||||||
Db        129 AGCCAGGAAAAT 140
```

Search completed: March 11, 2006, 15:37:52
Job time : 214.337 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 00:16:23 ; Search time 1884.27 Seconds
(without alignments)
1588.688 Million cell updates/sec

Title: US-10-757-745-2
Perfect score: 1902
Sequence: 1 MELGSCLEGGREAAEEEGEP.....DCGRFPSDHWGLLCNLDIIL 362

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170337_4279/app_query.fasta_1
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss08
-USER=US10757745_@CGN_1_1_2715_@runat_09032006_170337_4279 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1902	100.0	1920	9	US-10-757-745-1	Sequence 1, Appli
2	1902	100.0	1940	9	US-10-783-271-28	Sequence 28, Appl
3	1902	100.0	1948	5	US-10-037-270-889	Sequence 889, App
4	1902	100.0	1948	6	US-10-117-722-889	Sequence 889, App
5	1902	100.0	1948	9	US-10-122-851-889	Sequence 889, App
6	1902	100.0	3152	3	US-09-981-353-88	Sequence 88, Appl
7	1894	99.6	1296	3	US-09-925-299-170	Sequence 170, App
8	1894	99.6	1296	3	US-09-925-299-170	Sequence 170, App
9	1258	66.1	1312	9	US-10-757-745-3	Sequence 3, Appli
10	1199	63.0	1088	9	US-10-450-763-22058	Sequence 22058, A
c 11	1177	61.9	1079	3	US-09-745-288-63	Sequence 63, Appl
c 12	1177	61.9	1079	7	US-10-453-919-63	Sequence 63, Appl
13	1119	58.8	674	5	US-10-066-543-1067	Sequence 1067, Ap
14	1069	56.2	644	5	US-10-066-543-1070	Sequence 1070, Ap
15	1068	56.2	1227	9	US-10-450-763-22059	Sequence 22059, A
16	1054	55.4	633	5	US-10-066-543-1084	Sequence 1084, Ap
17	1031	54.2	625	5	US-10-066-543-548	Sequence 548, App
18	1020	53.6	625	5	US-10-066-543-1031	Sequence 1031, Ap
19	902	47.4	553	5	US-10-066-543-983	Sequence 983, App
20	651	34.2	401	5	US-10-066-543-1363	Sequence 1363, Ap
21	590	31.0	400	7	US-10-242-535A-8082	Sequence 8082, Ap
22	590	31.0	400	7	US-10-085-783A-8082	Sequence 8082, Ap
c 23	510	26.8	391	5	US-10-066-543-826	Sequence 826, App
24	299.5	15.7	634	4	US-09-925-065A-783536	Sequence 783536,
25	298	15.7	176	3	US-09-878-722-121	Sequence 121, App
c 26	298	15.7	176	3	US-09-878-178-50	Sequence 50, Appl
27	298	15.7	176	3	US-09-904-456-121	Sequence 121, App
c 28	298	15.7	176	5	US-10-046-935-50	Sequence 50, Appl
c 29	298	15.7	176	5	US-10-146-502-50	Sequence 50, Appl
c 30	298	15.7	176	5	US-10-060-036-1367	Sequence 1367, Ap
31	296	15.6	606	4	US-09-925-065A-783537	Sequence 783537,
32	267	14.0	725	8	US-10-363-345A-21053	Sequence 21053, A
c 33	267	14.0	725	8	US-10-363-345A-21054	Sequence 21054, A
34	267	14.0	725	9	US-10-363-483A-21053	Sequence 21053, A
c 35	267	14.0	725	9	US-10-363-483A-21054	Sequence 21054, A
36	246	12.9	1686	7	US-10-424-599-132360	Sequence 132360,
c 37	221	11.6	268	3	US-09-876-889-312	Sequence 312, App
38	195.5	10.3	1497	7	US-10-437-963-42449	Sequence 42449, A
c 39	184.5	9.7	725	8	US-10-363-345A-21055	Sequence 21055, A
40	184.5	9.7	725	8	US-10-363-345A-21056	Sequence 21056, A
c 41	184.5	9.7	725	9	US-10-363-483A-21055	Sequence 21055, A
42	184.5	9.7	725	9	US-10-363-483A-21056	Sequence 21056, A
43	182.5	9.6	444	3	US-09-864-761-16762	Sequence 16762, A
44	182.5	9.6	691	3	US-09-864-761-33229	Sequence 33229, A
45	169	8.9	267	9	US-10-450-763-5709	Sequence 5709, Ap
46	158.5	8.3	525	7	US-10-021-323-14543	Sequence 14543, A
47	153	8.0	1236	7	US-10-437-963-47256	Sequence 47256, A
48	147	7.7	9737	9	US-10-764-420-1725	Sequence 1725, Ap
c 49	132	6.9	932	8	US-10-425-115-66072	Sequence 66072, A
50	129	6.8	898	8	US-10-425-115-105489	Sequence 105489,
51	128	6.7	1767	9	US-10-998-197-33	Sequence 33, Appl
52	128	6.7	1971	9	US-10-887-553A-839	Sequence 839, App
53	127.5	6.7	1350	10	US-11-097-143-9713	Sequence 9713, Ap
c 54	126	6.6	90043	5	US-10-087-192-1141	Sequence 1141, Ap
55	120	6.3	1644	6	US-10-369-493-24191	Sequence 24191, A
56	117.5	6.2	5299	7	US-10-473-574-33	Sequence 33, Appl
57	117.5	6.2	168050	9	US-10-939-321-1	Sequence 1, Appli
c 58	117	6.2	96960	8	US-10-484-577-662	Sequence 662, App
59	112	5.9	5322	7	US-10-473-574-31	Sequence 31, Appl
60	111	5.8	1608	3	US-09-738-626-2561	Sequence 2561, Ap
61	111	5.8	3309400	3	US-09-738-626-1	Sequence 1, Appli
62	110.5	5.8	2277	9	US-10-450-763-20786	Sequence 20786, A
63	110.5	5.8	2606	9	US-10-450-763-11636	Sequence 11636, A
64	110.5	5.8	5130	9	US-10-450-763-2332	Sequence 2332, Ap
65	110.5	5.8	5130	9	US-10-450-763-3043	Sequence 3043, Ap
66	110.5	5.8	5130	9	US-10-450-763-3970	Sequence 3970, Ap
67	110.5	5.8	5130	9	US-10-450-763-6046	Sequence 6046, Ap
68	110.5	5.8	5130	9	US-10-450-763-6593	Sequence 6593, Ap
69	110.5	5.8	5130	9	US-10-450-763-8626	Sequence 8626, Ap
70	110.5	5.8	5130	9	US-10-450-763-9629	Sequence 9629, Ap
71	110.5	5.8	5130	9	US-10-450-763-10788	Sequence 10788, A
72	110.5	5.8	5130	9	US-10-450-763-19864	Sequence 19864, A
73	110.5	5.8	5130	9	US-10-450-763-20770	Sequence 20770, A
74	110.5	5.8	5130	9	US-10-450-763-20902	Sequence 20902, A
75	110.5	5.8	5130	9	US-10-450-763-25259	Sequence 25259, A

ALIGNMENTS

RESULT 1

```
US-10-757-745-1
; Sequence 1, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1627)..(1627)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1108)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1849)..(1849)
;   OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

Alignment Scores:
Pred. No.:                2.19e-233       Length:          1920
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2 (1-362) x US-10-757-745-1 (1-1920)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 79

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 139

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 199

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 259

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 319

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 379

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1099

Qy        361 IleLeu 362
              ||||||
Db       1100 ATATTG 1105

RESULT 2
US-10-783-271-28
; Sequence 28, Application US/10783271
; Publication No. US20050186577A1
; GENERAL INFORMATION:
;  APPLICANT: Veridex, LLC
;  APPLICANT:  Wang, Yixin
;  TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
;  FILE REFERENCE: VDX-5003 USNP
;  CURRENT APPLICATION NUMBER: US/10/783,271
;  CURRENT FILING DATE:  2004-02-20
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: human
US-10-783-271-28

Alignment Scores:
Pred. No.:                2.22e-233       Length:          1940
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
```

```
DB:                 9              Gaps:          0

US-10-757-745-2 (1-362) x US-10-783-271-28 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
              ||||||
Db       1106 ATATTG 1111

RESULT 3
US-10-037-270-889
; Sequence 889, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/10/037,270
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-037-270-889

Alignment Scores:
Pred. No.:                2.24e-233       Length:          1948
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2 (1-362) x US-10-037-270-889 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124
```

```
RESULT 4
US-10-117-722-889
; Sequence 889, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BCIP
;  CURRENT APPLICATION NUMBER: US/10/117,722
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-117-722-889

Alignment Scores:
Pred. No.:                2.24e-233       Length:          1948
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2 (1-362) x US-10-117-722-889 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124

RESULT 5
US-10-122-851-889
; Sequence 889, Application US/10122851
; Publication No. US20050239060A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BDV3
;  CURRENT APPLICATION NUMBER: US/10/122,851
;  CURRENT FILING DATE:  2002-04-12
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-122-851-889

Alignment Scores:
Pred. No.:              2.24e-233       Length:         1948
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2 (1-362) x US-10-122-851-889 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
                ||||||
Db       1119 ATATTG 1124

RESULT 6
US-09-981-353-88
; Sequence 88, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE: PA-0038 US
;  CURRENT APPLICATION NUMBER: US/09/981,353
;  CURRENT FILING DATE:  2001-10-11
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020160382A1 232992.1
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

Alignment Scores:
Pred. No.:                4.89e-233       Length:          3152
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2 (1-362) x US-09-981-353-88 (1-3152)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 1311

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 1371

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 1431

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 1491

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1492 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 1551

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 1611

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 1671

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1672 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 1731

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1732 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1791

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1792 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1851

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1852 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1911

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1912 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1971

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1972 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 2031

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2032 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 2091

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2092 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 2151

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2152 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 2211

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2212 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 2271

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2272 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 2331

Qy        361 IleLeu 362
                ||||||
Db       2332 ATATTG 2337

RESULT 7
US-09-925-299-170
; Sequence 170, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170
```

```
Alignment Scores:
Pred. No.:                1.24e-232      Length:          1296
Score:                    1894.00        Matches:         361
Percent Similarity:       99.7%          Conservative:    0
Best Local Similarity:    99.7%          Mismatches:      1
Query Match:              99.6%          Indels:          0
DB:                       3              Gaps:            0

US-10-757-745-2 (1-362) x US-09-925-299-170 (1-1296)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 83

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 143

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 203

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 263

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 323

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 383

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 443

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 503

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACA 563

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 623

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTT 683

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 743

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 803

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 863

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 GTGGATGTCTGGGAGTTTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAA 923

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 983

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        984 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1043

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1103

Qy        361 IleLeu 362
              ||||||
Db       1104 ATATTG 1109

RESULT 8
US-09-925-299-170
; Sequence 170, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:                1.24e-232      Length:          1296
Score:                    1894.00        Matches:         361
Percent Similarity:       99.7%          Conservative:    0
Best Local Similarity:    99.7%          Mismatches:      1
Query Match:              99.6%          Indels:          0
DB:                       3              Gaps:            0

US-10-757-745-2 (1-362) x US-09-925-299-170 (1-1296)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 83

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 143

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 203

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 263

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 323

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
```

```
Db        324 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 383

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        384 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 443

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        444 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 503

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db        504 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACA 563

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db        564 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 623

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220

Db        624 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTT 683

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240

Db        684 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 743

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260

Db        744 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 803

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280

Db        804 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 863

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        864 GTGGATGTCTGGGAGTTTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAA 923

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        924 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 983

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340

Db        984 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1043

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360

Db       1044 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1103

Qy        361 IleLeu 362

Db       1104 ATATTG 1109
```

RESULT 9
US-10-757-745-3
; Sequence 3, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/10/757,745
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/697,863A
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (122)..(1234)
; OTHER INFORMATION:
US-10-757-745-3

Alignment Scores:

Pred. No.:	1.03e-150	Length:	1312
Score:	1258.00	Matches:	240
Percent Similarity:	80.6%	Conservative:	43
Best Local Similarity:	68.4%	Mismatches:	64
Query Match:	66.1%	Indels:	4
DB:	9	Gaps:	3

US-10-757-745-2 (1-362) x US-10-757-745-3 (1-1312)

```
Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31

Db        191 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 247

Qy         32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51

Db        248 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 307

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71

Db        308 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 367

Qy         72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91

Db        368 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 427

Qy         92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111

Db        428 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 484

Qy        112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131

Db        485 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 544

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151

Db        545 AGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 604

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171

Db        605 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 664

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191

Db        665 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAAATTT 724

Qy        192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211

Db        725 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 784

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231

Db        785 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 844

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251

Db        845 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 904

Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271

Db        905 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 964

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
```

```
Db        965 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 1024

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311

Db       1025 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1084

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331

Db       1085 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1138

Qy        332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHis 351

Db       1139 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1198

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362

Db       1199 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1231
```

RESULT 10
US-10-450-763-22058
; Sequence 22058, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 22058
; LENGTH: 1088
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1)..(687)
; OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein
; OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
; OTHER INFORMATION: Waterman Score=1199.
US-10-450-763-22058

Alignment Scores:

Pred. No.:	3.02e-143	Length:	1088
Score:	1199.00	Matches:	226
Percent Similarity:	99.1%	Conservative:	1
Best Local Similarity:	98.7%	Mismatches:	2
Query Match:	63.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2 (1-362) x US-10-450-763-22058 (1-1088)

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152

Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172

Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192

Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
```

```
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232

Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252

Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272

Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312

Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332

Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy        333 SerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHisTrp 352

Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 660

Qy        353 GlyLeuLeuCysAsnLeuAspIleIle 361

Db        661 GGTCTTCTGTGCAACTTAGATATAATA 687
```

RESULT 11
US-09-745-288-63/c
; Sequence 63, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446D1
; CURRENT APPLICATION NUMBER: US/09/745,288
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-745-288-63

Alignment Scores:

Pred. No.:	2.03e-140	Length:	1079
Score:	1177.00	Matches:	226
Percent Similarity:	99.6%	Conservative:	0
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	61.9%	Indels:	1
DB:	3	Gaps:	0

US-10-757-745-2 (1-362) x US-09-745-288-63 (1-1079)

```
Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155

Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175

Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958
```

```
Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398
```

RESULT 12
US-10-453-919-63/c
; Sequence 63, Application US/10453919
; Publication No. US20040033230A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C7
; CURRENT APPLICATION NUMBER: US/10/453,919
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-453-919-63

Alignment Scores:

Pred. No.:	2.03e-140	Length:	1079
Score:	1177.00	Matches:	226
Percent Similarity:	99.6%	Conservative:	0
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	61.9%	Indels:	1
DB:	7	Gaps:	0

US-10-757-745-2 (1-362) x US-10-453-919-63 (1-1079)

```
Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398
```

RESULT 13
US-10-066-543-1067
; Sequence 1067, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1067
; LENGTH: 674

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11, 377, 640, 654
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1067

Alignment Scores:
Pred. No.:                2.79e-133       Length:          674
Score:                    1119.00         Matches:         216
Percent Similarity:       98.6%           Conservative:    0
Best Local Similarity:    98.6%           Mismatches:      2
Query Match:              58.8%           Indels:          1
DB:                       5               Gaps:            0

US-10-757-745-2 (1-362) x US-10-066-543-1067 (1-674)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 77

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        253 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        272 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy        292 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

Qy        312 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProAr 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy        332 gSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAsp 350
              ||||||||||||||||||||||   ||||||||   ||||||||||||||||||
Db        618 AAGTTTGGACCTTCTTGGATTANAAAAACTGGACTGNGGTAGATTTCCTAGTGAT 672

RESULT 14
US-10-066-543-1070
; Sequence 1070, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1070
;   LENGTH: 644
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-066-543-1070

Alignment Scores:
Pred. No.:                7.12e-127       Length:          644
Score:                    1069.00         Matches:         203
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    99.5%           Mismatches:      0
Query Match:              56.2%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2 (1-362) x US-10-066-543-1070 (1-644)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||
Db        332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy        333 SerLeuAspLeu 336
              ||||||||||||
Db        632 AGTTTGGACCTT 643

RESULT 15
```

```
US-10-450-763-22059
; Sequence 22059, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22059
;   LENGTH: 1227
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (84)..(203)
;   OTHER INFORMATION: 85% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=164.
US-10-450-763-22059

Alignment Scores:
Pred. No.:                2.73e-126       Length:          1227
Score:                    1068.00         Matches:         248
Percent Similarity:       69.4%           Conservative:    11
Best Local Similarity:    66.5%           Mismatches:      25
Query Match:              56.2%           Indels:          89
DB:                       9               Gaps:            5

US-10-757-745-2 (1-362) x US-10-450-763-22059 (1-1227)

Qy          1 MetGluLeuGlySerCysLeu-GluGlyGlyArgGluAlaAlaGluGluGluGlyGluPr 20
              |||||||||||||||||||||     |||   ||||||||||||||||||||||||||||
Db         74 ATGGAGTTGGGGAGTTGCCTGTTATGGCTTTAGGGAGGCGGCGGAGGAAGAGGGCGAGCC 133

Qy         20 oGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAl 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGC 193

Qy         40 aAlaVal-AlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnS 60
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 CGCAGTGAGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACT 253

Qy         60 erTyrPheGluProProValGluGluSer-AlaLeuGluArgArgProGluThrIleSer 79
              ||||||||||||||||    ||||||||| ||||||||||||||||||||||||||||||
Db        254 CCTACTTCGAGCCTCCGGCGGAGGAGAGCAGCCTTGGAACGCCGACCTGAAACCATCTCT 313

Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAA 373

Qy        100 IleSerProSerGluAspThrGlnGlnGluAsn-GlySerMetPheSerLeu-IleThrT 119
              |||||||||||||||||||||||||||||||| |||||||||||||||||| |||||||
Db        374 ATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGGCAGCATGTTCTCTCTTCATTACCT 433

Qy        119 rp-AsnIle-AspGlyLeuAsp---LeuAsnAsnLeu-SerGlu-ArgAlaArgGlyVal 136
              || |||||| ||||||:::      |||||||||||| |||||| ||||||||||||||||
Db        434 GGGAATATTTGATGGGATTAGGATCTTAAACAATCTGTTCAGAGGAGGGCTCGAGGGGTG 493

Qy        137 CysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProPro 156
              ||||||||||||||||          :::::|||      |||
Db        494 TGTTCCTACTTAGCTTTGTAAGTATTATCACTTCTATTTAATAGTAATGTG--------- 544

Qy        157 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 176
                    ||||||:::

Db        545 ------TCTTATGTA---------------------------------------------- 553

Qy        177 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 196

Db        553 ------------------------------------------------------------ 553

Qy        197 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 216

Db        553 ------------------------------------------------------------ 553

Qy        217 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 236
                                                                      |||

Db        554 ----------------------------------------------------TATACGGAA 562

Qy        237 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 CGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACA 622

Qy        257 ValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr---ArgCysGlyGly 275
              |||||||||||||||||||||||||||||||||||||||||||:::   :::
Db        623 GTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTGAGTGACAAAGTCAGTTTG 682

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              :::   :::  :::  :::|||    :::     |||     |||||||||
Db        683 ATGTTTGAGGACAGATTATTGCTATGGAGACTAATTTAGGTTCCT---TTTTGCCAGTAT 739

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 799

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGAC 859

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 919

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        920 TGCAACTTAGATATAATATTG 940

RESULT 16
US-10-066-543-1084
; Sequence 1084, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1084
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3
;   OTHER INFORMATION: n = A,T,C or G
```

US-10-066-543-1084

Alignment Scores:

Pred. No.:	5.92e-125	Length:	633
Score:	1054.00	Matches:	200
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	55.4%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2 (1-362) x US-10-066-543-1084 (1-633)

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 633
```

RESULT 17
US-10-066-543-548
; Sequence 548, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 19, 71
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-548

Alignment Scores:

Pred. No.:	5.31e-122	Length:	625
Score:	1031.00	Matches:	198
Percent Similarity:	99.0%	Conservative:	0
Best Local Similarity:	99.0%	Mismatches:	1
Query Match:	54.2%	Indels:	1
DB:	5	Gaps:	0

US-10-757-745-2 (1-362) x US-10-066-543-548 (1-625)

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy        173 -GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 192
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy        192 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy        212 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        232 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy        252 oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        272 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy        292 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy        312 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 623
```

RESULT 18
US-10-066-543-1031
; Sequence 1031, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather

```
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1031
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 10
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1031

Alignment Scores:
Pred. No.:                1.39e-120       Length:          625
Score:                    1020.00         Matches:         194
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              53.6%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2 (1-362) x US-10-066-543-1031 (1-625)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 581

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 326
              ||||||||||||||||||||||||||||||||||||||||||
Db        582 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAG 623

RESULT 19
```

```
US-10-066-543-983
; Sequence 983, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 983
;   LENGTH: 553
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 6, 35
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-983

Alignment Scores:
Pred. No.:                1.79e-105       Length:          553
Score:                    902.00          Matches:         171
Percent Similarity:       98.3%           Conservative:    0
Best Local Similarity:    98.3%           Mismatches:      3
Query Match:              47.4%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2 (1-362) x US-10-066-543-983 (1-553)

Qy        129 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 148
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 CTCTANATGCATGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATA 89

Qy        149 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 TTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTAT 149

Qy        169 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 GAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 209

Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 GTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAAACCTT 269

Qy        209 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 TTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAG 329

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 AGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATG 389

Qy        249 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 CAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGA 449

Qy        269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 GAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 509
```

```
Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 302
              ||||||||||||||||||||||||||||||||||||||||||
Db        510 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 551

RESULT 20
US-10-066-543-1363
; Sequence 1363, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1363
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-066-543-1363

Alignment Scores:
Pred. No.:              2.25e-73      Length:         401
Score:                  651.00        Matches:        127
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            34.2%         Indels:         0
DB:                     5             Gaps:           0

US-10-757-745-2 (1-362) x US-10-066-543-1363 (1-401)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 78

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 138

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 198

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 318

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 378

Qy        253 GluSerAlaThrValIlePhe 259
              |||||||||||||||||||||
Db        379 GAGTCAGCTACAGTTATATTT 399

RESULT 21
US-10-242-535A-8082
; Sequence 8082, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8082
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-8082

Alignment Scores:
Pred. No.:              1.61e-65      Length:         400
Score:                  590.00        Matches:        111
Percent Similarity:     100.0%        Conservative:   1
Best Local Similarity:  99.1%         Mismatches:     0
Query Match:            31.0%         Indels:         0
DB:                     7             Gaps:           0

US-10-757-745-2 (1-362) x US-10-242-535A-8082 (1-400)

Qy        223 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 242
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GTGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAA 124

Qy        243 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGAT 184

Qy        263 ThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGAT 244

Qy        283 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 GTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 304

Qy        303 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAla 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCA 364

Qy        323 AlaAlaGluGluGlyHisIleIleProArgSerLeu 334
              |||||||||||||||||||||||||||||||||||||
Db        365 GCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 400

RESULT 22
US-10-085-783A-8082
; Sequence 8082, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
```

```
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8082
;    LENGTH: 400
;    TYPE: DNA
;    ORGANISM: Human
US-10-085-783A-8082

Alignment Scores:
Pred. No.:                1.61e-65        Length:          400
Score:                    590.00          Matches:         111
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    99.1%           Mismatches:      0
Query Match:              31.0%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-2 (1-362) x US-10-085-783A-8082 (1-400)

Qy         223 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 242
               :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          65 GTGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAA 124

Qy         243 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 262
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         125 ATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGAT 184

Qy         263 ThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 282
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         185 ACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGAT 244

Qy         283 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 302
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         245 GTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 304

Qy         303 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAla 322
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         305 TCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCA 364

Qy         323 AlaAlaGluGluGlyHisIleIleProArgSerLeu 334
               ||||||||||||||||||||||||||||||||||||
Db         365 GCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 400

RESULT 23
US-10-066-543-826/c
; Sequence 826, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 826
;    LENGTH: 391
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: 373
;    OTHER INFORMATION: n = A,T,C or G
US-10-066-543-826

Alignment Scores:
Pred. No.:                3.12e-55        Length:          391
Score:                    510.00          Matches:         98
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              26.8%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2 (1-362) x US-10-066-543-826 (1-391)

Qy          37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         337 AGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGG 278

Qy          57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         277 GCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAA 218

Qy          77 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 96
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         217 ACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACC 158

Qy          97 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 116
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         157 ACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTC 98

Qy         117 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          97 ATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 44

RESULT 24
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
;    LENGTH: 634
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-783536

Alignment Scores:
Pred. No.:                8.83e-28        Length:          634
Score:                    299.50          Matches:         63
Percent Similarity:       83.3%           Conservative:    2
Best Local Similarity:    80.8%           Mismatches:      10
Query Match:              15.7%           Indels:          3
DB:                       4               Gaps:            1
```

```
US-10-757-745-2 (1-362) x US-09-925-065A-783536 (1-634)

Qy          85 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 104
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         265 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 324

Qy         105 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         325 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 384

Qy         125 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         385 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTAAGTA 444

Qy         145 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 162
                  ::::::|||         |||           |||   |||   |||
Db         445 TTATCACTTCTATTTAATAGTAATGTG---------TCTTATGTATATACAAAG 489

RESULT 25
US-09-878-722-121
; Sequence 121, Application US/09878722
; Publication No. US20020040127A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Clapper, Jonathan
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.524
;  CURRENT APPLICATION NUMBER: US/09/878,722
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 245
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-878-722-121

Alignment Scores:
Pred. No.:              1.72e-28        Length:         176
Score:                  298.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            15.7%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2 (1-362) x US-09-878-722-121 (1-176)

Qy         144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           3 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 62

Qy         164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 122

Qy         184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176

RESULT 26
US-09-878-178-50/c
; Sequence 50, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527
;  CURRENT APPLICATION NUMBER: US/09/878,178
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 2237
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-878-178-50

Alignment Scores:
Pred. No.:              1.72e-28        Length:         176
Score:                  298.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            15.7%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2 (1-362) x US-09-878-178-50 (1-176)

Qy         144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy         164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy         184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 27
US-09-904-456-121
; Sequence 121, Application US/09904456
; Publication No. US20030017167A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.524C1
;  CURRENT APPLICATION NUMBER: US/09/904,456
;  CURRENT FILING DATE:  2001-07-11
;  NUMBER OF SEQ ID NOS: 247
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-904-456-121

Alignment Scores:
Pred. No.:              1.72e-28        Length:         176
Score:                  298.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            15.7%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2 (1-362) x US-09-904-456-121 (1-176)

Qy         144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           3 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 62

Qy         164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 122

Qy         184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176
```

```
RESULT 28
US-10-046-935-50/c
; Sequence 50, Application US/10046935
; Publication No. US20020156011A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C1
;  CURRENT APPLICATION NUMBER: US/10/046,935
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2239
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-046-935-50

Alignment Scores:
Pred. No.:              1.72e-28        Length:         176
Score:                  298.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            15.7%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2 (1-362) x US-10-046-935-50 (1-176)

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy        164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 29
US-10-146-502-50/c
; Sequence 50, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C2
;  CURRENT APPLICATION NUMBER: US/10/146,502
;  CURRENT FILING DATE:  2002-05-14
;  NUMBER OF SEQ ID NOS: 2241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-146-502-50

Alignment Scores:
Pred. No.:              1.72e-28        Length:         176
Score:                  298.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            15.7%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2 (1-362) x US-10-146-502-50 (1-176)

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy        164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 30
US-10-060-036-1367/c
; Sequence 1367, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
;  APPLICANT: Benson, Darin R.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Persing, David H.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF PANCREATIC CANCER
;  FILE REFERENCE: 210121.566
;  CURRENT APPLICATION NUMBER: US/10/060,036
;  CURRENT FILING DATE:  2002-01-30
;  NUMBER OF SEQ ID NOS: 4560
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1367
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-060-036-1367

Alignment Scores:
Pred. No.:              1.72e-28        Length:         176
Score:                  298.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            15.7%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2 (1-362) x US-10-060-036-1367 (1-176)

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy        164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 31
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
```

```
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783537

Alignment Scores:
Pred. No.:              2.32e-27        Length:         606
Score:                  296.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            15.6%           Indels:         0
DB:                     4               Gaps:           0

US-10-757-745-2 (1-362) x US-09-925-065A-783537 (1-606)

Qy         85 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 473

Qy        105 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 533

Qy        125 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTG 587

RESULT 32
US-10-363-345A-21053
; Sequence 21053, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21053
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21053
US-10-363-345A-21053

Alignment Scores:
Pred. No.:              1.68e-23        Length:         725
Score:                  267.00          Matches:        63
Percent Similarity:     56.8%           Conservative:   0
Best Local Similarity:  56.8%           Mismatches:     19
Query Match:            14.0%           Indels:         30
DB:                     8               Gaps:           1

US-10-757-745-2 (1-362) x US-10-363-345A-21053 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 ATGGAGTTGGGGAGTTGTTTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGTTT 382

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||  ||||||||||||||  ||||||  |||||||||
Db        383 GAGGTGAAAAAGCGGCGATTTTTGTGTGTGGAGTTTGTTTCGGTCGTAAGTTGCGATGTC 442

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu---------------- 55
                 |||      |||||||||  ||||||||||||||||||
Db        443 GTAGTGGTTTAGTGTTTTTTGGTCGAGAACGATTGGGAGATGGA-AGTACGTTTTCGTTG 501

Qy         55 ---------------------------------------------------------- 55

Db        502 TTTTCGTTTTCGGTTTTATTATTTAGGTCGGTGGAGCGTATAAATTTTATTTTTTGTGTT 561

Qy         56 ------------ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                        |||   ||||||  |||||||||      ||||||||||||   |||
Db        562 TTTGTTTTTTAGAGGGTTTTGAATTTTTATTTCGAGTTTTCGGTGGAGGAGAGCGTTTTG 621

Qy         72 GluArgArgProGluThrIleSerGluProLys 82
              |||||||||   |||   |||   |||   |||
Db        622 GAACGTCGATTTGAAATTATTTTTGAGTTTAAG 654

RESULT 33
US-10-363-345A-21054/c
; Sequence 21054, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21054
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21054
US-10-363-345A-21054

Alignment Scores:
Pred. No.:              1.68e-23        Length:         725
Score:                  267.00          Matches:        63
Percent Similarity:     56.8%           Conservative:   0
Best Local Similarity:  56.8%           Mismatches:     19
Query Match:            14.0%           Indels:         30
DB:                     8               Gaps:           1

US-10-757-745-2 (1-362) x US-10-363-345A-21054 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 ATGGAGTTGGGGAGTTGTTTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGTTT 344

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||  ||||||||||||||  ||||||  |||||||||
Db        343 GAGGTGAAAAAGCGGCGATTTTTGTGTGTGGAGTTTGTTTCGGTCGTAAGTTGCGATGTC 284

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu---------------- 55
                 |||      |||||||||  ||||||||||||||||||
Db        283 GTAGTGGTTTAGTGTTTTTTGGTCGAGAACGATTGGGAGATGGA-AGTACGTTTTCGTTG 225

Qy         55 ---------------------------------------------------------- 55
```

```
Db        224 TTTTCGTTTTCGGTTTTATTATTTAGGTCGGTGGAGCGTATAAATTTTATTTTTTGTGTT 165

Qy         56 ------------ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                          ||| |||||| |||||||||     ||||||||||| |||
Db        164 TTTGTTTTTTAGAGGGTTTTGAATTTTTATTTCGAGTTTTCGGTGGAGGAGAGCGTTTTG 105

Qy         72 GluArgArgProGluThrIleSerGluProLys 82
              ||||||||| |||  |||  |||  |||
Db        104 GAACGTCGATTTGAAATTATTTTTGAGTTTAAG 72

RESULT 34
US-10-363-483A-21053
; Sequence 21053, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21053
;    LENGTH: 725
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;    OTHER INFORMATION: CpG-island No: 21053
US-10-363-483A-21053

Alignment Scores:
Pred. No.:                1.68e-23        Length:          725
Score:                    267.00          Matches:         63
Percent Similarity:       56.8%           Conservative:    0
Best Local Similarity:    56.8%           Mismatches:      19
Query Match:              14.0%           Indels:          30
DB:                       9               Gaps:            1

US-10-757-745-2 (1-362) x US-10-363-483A-21053 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 ATGGAGTTGGGGAGTTGTTTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGTTT 382

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              |||||||||||||||||  ||||||||||||||  ||||||  ||||||||
Db        383 GAGGTGAAAAAGCGGCGATTTTTGTGTGTGGAGTTTGTTTCGGTCGTAAGTTGCGATGTC 442

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu--------------- 55
                 |||      |||||||||   ||||||||||||||||||
Db        443 GTAGTGGTTTAGTGTTTTTTGGTCGAGAACGATTGGGAGATGGA-AGTACGTTTTCGTTG 501

Qy         55 ------------------------------------------------------------ 55

Db        502 TTTTCGTTTTCGGTTTTATTATTTAGGTCGGTGGAGCGTATAAATTTTATTTTTTGTGTT 561

Qy         56 ------------ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                          ||| |||||| |||||||||     ||||||||||| |||
Db        562 TTTGTTTTTTAGAGGGTTTTGAATTTTTATTTCGAGTTTTCGGTGGAGGAGAGCGTTTTG 621

Qy         72 GluArgArgProGluThrIleSerGluProLys 82
              ||||||||| |||  |||  |||  |||
Db        622 GAACGTCGATTTGAAATTATTTTTGAGTTTAAG 654

RESULT 35
US-10-363-483A-21054/c
; Sequence 21054, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21054
;    LENGTH: 725
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;    OTHER INFORMATION: CpG-island No: 21054
US-10-363-483A-21054

Alignment Scores:
Pred. No.:                1.68e-23        Length:          725
Score:                    267.00          Matches:         63
Percent Similarity:       56.8%           Conservative:    0
Best Local Similarity:    56.8%           Mismatches:      19
Query Match:              14.0%           Indels:          30
DB:                       9               Gaps:            1

US-10-757-745-2 (1-362) x US-10-363-483A-21054 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 ATGGAGTTGGGGAGTTGTTTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGTTT 344

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              |||||||||||||||||  ||||||||||||||  ||||||  ||||||||
Db        343 GAGGTGAAAAAGCGGCGATTTTTGTGTGTGGAGTTTGTTTCGGTCGTAAGTTGCGATGTC 284

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu--------------- 55
                 |||      |||||||||   ||||||||||||||||||
Db        283 GTAGTGGTTTAGTGTTTTTTGGTCGAGAACGATTGGGAGATGGA-AGTACGTTTTCGTTG 225

Qy         55 ------------------------------------------------------------ 55

Db        224 TTTTCGTTTTCGGTTTTATTATTTAGGTCGGTGGAGCGTATAAATTTTATTTTTTGTGTT 165

Qy         56 ------------ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                          ||| |||||| |||||||||     ||||||||||| |||
Db        164 TTTGTTTTTTAGAGGGTTTTGAATTTTTATTTCGAGTTTTCGGTGGAGGAGAGCGTTTTG 105

Qy         72 GluArgArgProGluThrIleSerGluProLys 82
              ||||||||| |||  |||  |||  |||
Db        104 GAACGTCGATTTGAAATTATTTTTGAGTTTAAG 72

RESULT 36
US-10-424-599-132360
; Sequence 132360, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132360
;    LENGTH: 1686
;    TYPE: DNA
;    ORGANISM: Glycine max
;    FEATURE:
;    OTHER INFORMATION: Clone ID: PAT_MRT3847_90528C.1
US-10-424-599-132360
```

```
Alignment Scores:
Pred. No.:              3.36e-20        Length:         1686
Score:                  246.00          Matches:        91
Percent Similarity:     42.2%           Conservative:   54
Best Local Similarity:  26.5%           Mismatches:     121
Query Match:            12.9%           Indels:         79
DB:                     7               Gaps:           14

US-10-757-745-2 (1-362) x US-10-424-599-132360 (1-1686)

Qy         59 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle 78
              :::|||   |||||||||               |||         ||||||   |||
Db        .482 TCTTCCCTCTTCGAACCT--------------GCAAAAAGAGAAAGGCCATCGACG--- 523

Qy         79 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 98
                                                   |||   ||||||:::
Db        524 ----------------------------------ACGACGAC-GATTCTTCTGAGGTC 546

Qy         99 LysIleSerProSerGlu--------------------AspThrGlnGlnGluAsnGly 111
              |||:::   ||||||:::                    |||          ::::::|||
Db        547 AAGGTCAAGCCATCCGATAAAGCTACCGATATCACCGGTGATGAGAATATTGACTCAGGA 606

Qy        112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131
                  |||   ::::::::::::|||:::         :::          |||   :::
Db        607 AAAGCCTTCAAGATATTGAGTTACAATGTTTGGTTCCGGGAAGACTTGGAGTTGCACAAG 666

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
              |||   :::   :::         :::      :::|||||||||   |||       |||
Db        667 AGGATGAAGGCTATTGGCGACCTTGTTCAGTTCCATTCCCCTGACTTTATCTGTTTCCAG 726

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSer------SerAsnTyrGlu 169
              ||||||   |||       |||        |||:::   :::       |||   |||
Db        727 GAGGTTACTCCCAATATATATGACATTTTCAAGCGATCCACTTGGTGGAGTGGGTATCGT 786

Qy        170 IleIleThrGlyHisGluGluGlyTyr---------PheThrAlaIleMetLeuLysLys 186
                       |||       |||          :::      :::::|||   |||
Db        787 TGCTCAGTTTCTTCTGAGATGGATTATTCAAGACCCTACTTTTGTATGGTGTTAAGCAAA 846

Qy        187 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 206
              |||||:|       |||::::::       ||||||   ::::::   |||   |||
Db        847 CTGCCTGTGAAATCATTCAGCAAGAAG------CCCTTCAGCAATTCTATAATGGGAAGA 900

Qy        207 AsnLeuLeuCysValHisValAsn---ValSerGlyAsnGluLeuCysLeuMetThrSer 225
                 |||         |||          ||||||       |||   :::   ||||||
Db        901 GAACTTTGCATTGCGGAGGTTGAAGCTGCAAGTGGCAAGCCATTGGTTATTGCCACTAGC 960

Qy        226 HisLeuGluSer-----------------------ThrArgGlyHisAlaAlaGluArg 237
              ||||||||||||                      :::   ::::::   ||||||
Db        961 CATCTTGAGAGTCCCTGTCCAGCTCCTCCAAAATGGGATCAGATGTACAGCAAGGAAAGA 1020

Qy        238 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
              :::   |||             :::     :::::::::   |||        :::|||
Db       1021 GTAGTGCAGGCCAATGAGGCTATAAACCTTCTCAAGAAACAGCCA---------AGTGTT 1071

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro 277
              :::|||   ||||||   |||      |||::::::      :::         |||
Db       1072 GTTTTTGGGGGAGACATGAACTGGAATGACCAACAGGATGGTCAATAT---CCTCTACAA 1128

Qy        278 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
              :::      ||||||   |||      |||   :::|||            :::|||:::
Db       1129 GATGGATGGGTTGATGCCTGGTCTCAGCTA---AGACCAAATGAAAGTGGTTGGACTTAT 1185

Qy        298 AspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArg 317
              ||||||:::   |||      |||                ::   |||   ||||||
Db       1186 GACACCAAGTCAAACCAGATGTTGACAGGCAACCGTACTCTCCAAAAGCGATTAGATCGC 1245

Qy        318 IlePhe--------------PheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
                               |||:::
Db       1246 TTTATTTGCCAATTCACTGATTTTAAGATAACC------------------------- 1278

Qy        333 SerLeuAspLeuLeuGlyLeuGlu---------------------------------- 340
              |||:::|||::::::|||:::|||
```

```
Db       1279 AGTGTTGACATGATTGGGATGGAAGCAATACCTGGTGTTTCATACAACAAAGAAAAGAAA 1338

Qy        341 -----------------LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeu 354
                              :::|||              ||||||||||||::|||||||
Db       1339 GTAAGAAAGGAGATAAAACAACTGGTACTCCCAGTTTTGCCTAGTGATCATTATGGCCTG 1398

Qy        355 LeuCysAsnLeu 358
              |||   :::
Db       1399 CTTTTGACAATT 1410

RESULT 37
US-09-876-889-312/c
; Sequence 312, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
;  APPLICANT: Benson, Darin R.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
;  TITLE OF INVENTION:  CANCER THERAPY AND DIAGNOSIS
;  FILE REFERENCE: 210121.466C3
;  CURRENT APPLICATION NUMBER: US/09/876,889
;  CURRENT FILING DATE:  2001-06-06
;  NUMBER OF SEQ ID NOS: 353
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 312
;   LENGTH: 268
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-876-889-312

Alignment Scores:
Pred. No.:              2.81e-18        Length:         268
Score:                  221.00          Matches:        41
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            11.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2 (1-362) x US-09-876-889-312 (1-268)

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAA 208

Qy        342 LeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 CTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATA 148

Qy        362 Leu 362
              |||
Db        147 TTG 145

RESULT 38
US-10-437-963-42449
; Sequence 42449, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
```

```
; SEQ ID NO 42449
;   LENGTH: 1497
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_456C.1
US-10-437-963-42449

Alignment Scores:
Pred. No.:                8.83e-14        Length:         1497
Score:                    195.50          Matches:        74
Percent Similarity:       42.3%           Conservative:   46
Best Local Similarity:    26.1%           Mismatches:     111
Query Match:              10.3%           Indels:         53
DB:                       7               Gaps:           11

US-10-757-745-2 (1-362) x US-10-437-963-42449 (1-1497)

Qy        114 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 133
              |||   ::::::|||:::|||:::      :::          |||       |||
Db        685 TTCAAAATCATGACATACAATGTCTGGTTTCATGAGGATTTGGAATTGTGTAGAAGGATG 744

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
              :::          :::      ::::::||||||:::|||       |||||||||
Db        745 GATGCTCTTGGGGATCTTATTAAGAACCACAACCCTGACCTTATATGCTTCCAGGAGGTT 804

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSer------SerAsnTyrGluIleIle 171
              |||      |||      |||:::|||             |||
Db        805 ACACCAAACATATATCTGCTTCTACAAAAATCTGACTGGTGGCAAGAATACCGATGCTCG 864

Qy        172 ThrGlyHis-----------GluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 187
              :::          :::      |||:::          ::::::   |||
Db        865 CTGTCGAATAGCATGGCTATGCAGCGGAAATATTAC---------TGCATGCAGATGAGC 915

Qy        188 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 207
              ::::::   :::::::|||   :::      ||||||   ::::::   |||   |||
Db        916 AAATTGCCTGTTGAGTCTTTTGACTGCACTCCATTTTCCAACTCAATAATGGGAAGAGAA 975

Qy        208 LeuLeuCysValHisValAsnValSerGlyAsn---GluLeuCysLeuMetThrSerHis 226
              |||          ||||||          |||      :::||||   |||  |||||||||
Db        976 CTGTGTGTAGCACATGTCAAAACTGGAGGTGCAGTTAAGCTAGTGCTGGCCACTAGCCAC 1035

Qy        227 LeuGluSer-----------------------ThrArgGlyHisAlaAlaGluArgMet 238
              |||||||||                     :::   ::::::   ||||||:::
Db       1036 CTTGAGAGTCCAATGCCTGGGCCTCCTACTTGGGATCAGATGTACAGCACGGAACGAGTT 1095

Qy        239 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258
              |||          ||||||   :::                          ||||||
Db       1096 GCTCAGGCAAACAAATCTCTCAAAATTTTG---------GGATCCTTCCGCAATGTAATA 1146

Qy        259 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly--------- 275
              |||   ||||||   |||      |||:::          |||
Db       1147 TTTTGTGGCGATATGAACTGGGATGACAAA--------------GGAGATGGGCCCTTC 1191

Qy        276 ---LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294
                 ||||||          |||   |||      |||   ||||||
Db       1192 CCCCTACCAGCTGGCTGGACCGATGCTTGGATTGAGTTG---AAGCCGGGTGAAGATGGC 1248

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314
              :::|||:::||||||:::   ||||||   |||                :::    |||
Db       1249 TGGACTTATGACACAAAAGCTAACAGTATGCTATCAGCCAACCGTAAGTTACAGAAGAGA 1308

Qy        315 PheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeu 334
              ||||||          :::  |||                              ::::::
Db       1309 TTGGACCGATTCGTGTGCAAGCTGGCA------------GATTTCAAGATTAACAACATC 1356

Qy        335 AspLeuLeuGly---------------------------------------------- 338
              ::::::|||
Db       1357 CAGATGATCGGGAAGAATGCGATACCTGGTCTCTCATATGTCAAGGAGAAGAAGGTCCGC 1416

Qy        339 ------LeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356
                    :::   ||||||:::                ||||||||||||:::||||||:::
Db       1417 AAGGAAGTTCGCAAACTGGAACTTCCTGTCTTACCCAGTGACCACTTTGGACTTGTTTTG 1476

Qy        357 AsnLeuAspIle 360
              ::::::   :::
Db       1477 AGCATTACTCTT 1488

RESULT 39
US-10-363-345A-21055/c
; Sequence 21055, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21055
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21055
US-10-363-345A-21055

Alignment Scores:
Pred. No.:                7.1e-13         Length:         725
Score:                    184.50          Matches:        40
Percent Similarity:       58.9%           Conservative:   26
Best Local Similarity:    35.7%           Mismatches:     17
Query Match:              9.7%            Indels:         30
DB:                       8               Gaps:           1

US-10-757-745-2 (1-362) x US-10-363-345A-21055 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ::::::|||   :::   |||:::      ::::::      :::::::::   ||||||
Db        403 ATAAAATTAAAAAATTACCTAAAAAACGAAAAAAAAAACGACGAAAAAAAAAAACGAACCT 344

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ::::::||||||||||||||||||   :::::::|||   |||:::|||:::   |||
Db        343 AAAATAAAAAAACGACGACTTCTATATATAAAATTTACCTCGATCGCAAACTACGATACC 284

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGlu---------------------- 53
              |||:::   |||   ||||||   |||||||||   :::
Db        283 GCAATAACTCAATACTTCCTAACCGAAAACGACTA-AAAAATAAAAATACGCTTTCGCTA 225

Qy         53 ------------------------------------------------------------ 53

Db        224 TCCTCGCTTCCGACTCCATTACCTAAATCGATAAAACGCACAAACCTCATCCCTTATACC 165

Qy         54 ------MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                    ::::::   ||||||||||||||||||||||:::::::::::::||||||
Db        164 TTTATTTTTCAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAACGCCTTA 105

Qy         72 GluArgArgProGluThrIleSerGluProLysThr 83
              :::|||||||||:::|||||||||:::||||||||
Db        104 AAACGCCGACCTAAAACCATCTCTAAACCCAAAACC 69

RESULT 40
US-10-363-345A-21056
; Sequence 21056, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
```

```
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21056
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21056
US-10-363-345A-21056

Alignment Scores:
Pred. No.:              7.1e-13       Length:          725
Score:                  184.50        Matches:         40
Percent Similarity:     58.9%         Conservative:    26
Best Local Similarity:  35.7%         Mismatches:      17
Query Match:            9.7%          Indels:          30
DB:                     8             Gaps:            1

US-10-757-745-2 (1-362) x US-10-363-345A-21056 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ::::::|||   :::   |||:::     ::::::     :::::::::   ||||||
Db        323 ATAAAATTAAAAAATTACCTAAAAAACGAAAAAAAAAACGACGAAAAAAAAAAAACGAACCT 382

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ::::::||||||||||||||||||   ::::::|||   |||:::|||:::   |||
Db        383 AAAATAAAAAAACGACGACTTCTATATATAAAATTTACCTCGATCGCAAACTACGATACC 442

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGlu---------------------- 53
              |||:::   |||   ||||||   |||||||||   :::
Db        443 GCAATAACTCAATACTTCCTAACCGAAAACGACTA-AAAAATAAAAATACGCTTTCGCTA 501

Qy         53 ------------------------------------------------------------ 53

Db        502 TCCTCGCTTCCGACTCCATTACCTAAATCGATAAAACGCACAAACCTCATCCCTTATACC 561

Qy         54 ------MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                    ::::::   ||||||||||||||||||||||||:::::::::::||||||
Db        562 TTTATTTTTCAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAACGCCTTA 621

Qy         72 GluArgArgProGluThrIleSerGluProLysThr 83
              :::|||||||||:::|||||||||:::||||||||
Db        622 AAACGCCGACCTAAAACCATCTCTAAACCCAAAACC 657

RESULT 41
US-10-363-483A-21055/c
; Sequence 21055, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21055
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21055
US-10-363-483A-21055

Alignment Scores:
Pred. No.:              7.1e-13       Length:          725
Score:                  184.50        Matches:         40
Percent Similarity:     58.9%         Conservative:    26
Best Local Similarity:  35.7%         Mismatches:      17
Query Match:            9.7%          Indels:          30
DB:                     9             Gaps:            1

US-10-757-745-2 (1-362) x US-10-363-483A-21055 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ::::::|||   :::   |||:::     ::::::     :::::::::   ||||||
Db        403 ATAAAATTAAAAAATTACCTAAAAAACGAAAAAAAAAACGACGAAAAAAAAAAAACGAACCT 344

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ::::::||||||||||||||||||   ::::::|||   |||:::|||:::   |||
Db        343 AAAATAAAAAAACGACGACTTCTATATATAAAATTTACCTCGATCGCAAACTACGATACC 284

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGlu---------------------- 53
              |||:::   |||   ||||||   |||||||||   :::
Db        283 GCAATAACTCAATACTTCCTAACCGAAAACGACTA-AAAAATAAAAATACGCTTTCGCTA 225

Qy         53 ------------------------------------------------------------ 53

Db        224 TCCTCGCTTCCGACTCCATTACCTAAATCGATAAAACGCACAAACCTCATCCCTTATACC 165

Qy         54 ------MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                    ::::::   ||||||||||||||||||||||||:::::::::::||||||
Db        164 TTTATTTTTCAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAACGCCTTA 105

Qy         72 GluArgArgProGluThrIleSerGluProLysThr 83
              :::|||||||||:::|||||||||:::||||||||
Db        104 AAACGCCGACCTAAAACCATCTCTAAACCCAAAACC 69

RESULT 42
US-10-363-483A-21056
; Sequence 21056, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21056
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21056
US-10-363-483A-21056

Alignment Scores:
Pred. No.:              7.1e-13       Length:          725
Score:                  184.50        Matches:         40
Percent Similarity:     58.9%         Conservative:    26
Best Local Similarity:  35.7%         Mismatches:      17
Query Match:            9.7%          Indels:          30
DB:                     9             Gaps:            1

US-10-757-745-2 (1-362) x US-10-363-483A-21056 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ::::::|||   :::   |||:::     ::::::     :::::::::   ||||||
Db        323 ATAAAATTAAAAAATTACCTAAAAAACGAAAAAAAAAACGACGAAAAAAAAAAAACGAACCT 382

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ::::::||||||||||||||||||   ::::::|||   |||:::|||:::   |||
Db        383 AAAATAAAAAAACGACGACTTCTATATATAAAATTTACCTCGATCGCAAACTACGATACC 442

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGlu---------------------- 53
              |||:::   |||   ||||||   |||||||||   :::
```

```
Db        443 GCAATAACTCAATACTTCCTAACCGAAAACGACTA-AAAAATAAAAATACGCTTTCGCTA 501

Qy         53 ------------------------------------------------------------ 53

Db        502 TCCTCGCTTCCGACTCCATTACCTAAATCGATAAAACGCACAAACCTCATCCCTTATACC 561

Qy         54 ------MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                   ::::::   ||||||||||||||||||||||||||:::::::::::|||||||
Db        562 TTTATTTTTCAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAAACGCCTTA 621

Qy         72 GluArgArgProGluThrIleSerGluProLysThr 83
              :::|||||||||:::|||||||||:::|||||||||
Db        622 AAACGCCGACCTAAAACCATCTCTAAACCCAAAACC 657

RESULT 43
US-09-864-761-16762
; Sequence 16762, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16762
;   LENGTH: 444
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9

US-09-864-761-16762

Alignment Scores:
Pred. No.:               5.8e-13        Length:          444
Score:                   182.50         Matches:         37
Percent Similarity:      74.5%          Conservative:    4
Best Local Similarity:   67.3%          Mismatches:      11
Query Match:             9.6%           Indels:          3
DB:                      3              Gaps:            1

US-10-757-745-2 (1-362) x US-09-864-761-16762 (1-444)

Qy        135 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 151
              |||:::|||       |||                |||||||||||||||||||||||
Db        221 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 280

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 340

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186
              ||||||:::       |||   |||   :::      :::|||
Db        341 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 385

RESULT 44
US-09-864-761-33229
; Sequence 33229, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
```

```
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33229
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
;   OTHER INFORMATION: SWISSPROT HIT: P18004, EVALUE 7.00e-01
;   OTHER INFORMATION: NT HIT: AF201687.1, EVALUE 3.00e-45
;   OTHER INFORMATION: EST_HUMAN HIT: AV717253.1, EVALUE 5.00e-45
US-09-864-761-33229

Alignment Scores:
Pred. No.:                1.19e-12        Length:         691
Score:                    182.50          Matches:        37
Percent Similarity:       74.5%           Conservative:   4
Best Local Similarity:    67.3%           Mismatches:     11
Query Match:              9.6%            Indels:         3
DB:                       3               Gaps:           1

US-10-757-745-2 (1-362) x US-09-864-761-33229 (1-691)

Qy        135 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 151
              |||:::|||       |||                |||||||||||||||||||||||||
Db         83 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 142

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 202

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186
              ||||||:::         |||   |||   :::      :::|||
Db        203 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 247

RESULT 45
US-10-450-763-5709
; Sequence 5709, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 5709
;   LENGTH: 267
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (182)..(205)
;   OTHER INFORMATION: 87% homologous to Homo sapiens AD022 protein,accession number
;   OTHER INFORMATION: AF223469,Smith-Waterman Score=38.
US-10-450-763-5709

Alignment Scores:
Pred. No.:                1.39e-11        Length:         267
Score:                    169.00          Matches:        32
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              8.9%            Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2 (1-362) x US-10-450-763-5709 (1-267)

Qy         24 LysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaValAla 43
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCT 62

Qy         44 GlnCysPheLeuAlaGluAsnAspTrpGluMetGlu 55
              ||||||||||||||||||||||||||||||||||||
Db         63 CAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAA 98

RESULT 46
US-10-021-323-14543
; Sequence 14543, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
;  APPLICANT: Deikman, Jill
;  APPLICANT:  Feng, Paul C.C.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Ziegler, Todd E.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(52274)B
;  CURRENT APPLICATION NUMBER: US/10/021,323
;  CURRENT FILING DATE:  2001-12-12
;  PRIOR APPLICATION NUMBER: US 60/255, 619
;  PRIOR FILING DATE: 2000-12-14
;  NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 14543
;   LENGTH: 525
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3829-034-Q1-K6-G12
US-10-021-323-14543

Alignment Scores:
Pred. No.:                9.39e-10        Length:         525
Score:                    158.50          Matches:        53
Percent Similarity:       46.1%           Conservative:   29
Best Local Similarity:    29.8%           Mismatches:     63
Query Match:              8.3%            Indels:         33
DB:                       7               Gaps:           7

US-10-757-745-2 (1-362) x US-10-021-323-14543 (1-525)

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   :::   :::          :::   |||:::|||||||||:::|||
Db         19 AAGAGGATGAAAGCTATTGGTGACCTTATACAACTTCATTCCCCTGATATCATTTGTTTT 78

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              |||||||||   |||       |||                |||   ||||||:::
Db         79 CAAGAGGTTACTCCGATGATTTATGACATTTTT------CGTGGGTCTAATTGG------ 126

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIle---------------------- 182
                          :::||||||      :::|||
Db        127 -----------TGGAAAGGGTACCGGTGCTCTATCTCGGACAACACGGCGAGTTTGAGA 174

Qy        183 ---------MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 199
                      ::::::   |||::::::   :::||||||                ||||||
Db        175 GCTTACTTTTGTATACAGTTAAGCAAATTGCCTGTGAAATCCTTTCGTCGTGAACCTTTT 234

Qy        200 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 219
              ::::::   |||   |||   |||             |||   |||||||||   |||
Db        235 GATAATTCGATAATGGGGCGAGAACTTTGCATGACTGAGGTTGAAGTTTCTGGGGGTGAA 294

Qy        220 ------LeuCysLeuMetThrSerHisLeuGluSer---------------------- 229
                    |||   :::   |||:::||||||||||||
Db        295 ACAACGCTTGTTGTTGCTACTACCCATCTCGAGAGTCCGTGTCCTGCTCCGCCGAAATGG 354

Qy        230 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 249
              :::   ::::::   |||||||:::   |||   |||:::   :::      ::::::
Db        355 GATCAAATGTATAGTAAGGAACGTGTAGAACAAGCGAAAGTAGCGATCAATATTCTCGAG 414
```

```
Qy        250 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
Db        415 AATAACCCG---------AATGTTATTTTTGGGGGTGATATGAATTGGGATGACAAATTA 465

Qy        270 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
Db        466 GATGGTCGATTT---CCTTTAACTGATGGATGGATTGATGCTTGGAACGAGTTA 516

RESULT 47
US-10-437-963-47256
; Sequence 47256, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 47256
;   LENGTH: 1236
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_50042C.1
US-10-437-963-47256

Alignment Scores:
Pred. No.:              1.93e-08      Length:          1236
Score:                  153.00        Matches:         86
Percent Similarity:     40.2%         Conservative:    65
Best Local Similarity:  22.9%         Mismatches:      142
Query Match:            8.0%          Indels:          84
DB:                     7             Gaps:            15

US-10-757-745-2 (1-362) x US-10-437-963-47256 (1-1236)

Qy         36 AlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu 55
Db        190 TCCGCCTGCGAGGCCGCGCGCCCCGTGGAGGTCGACGCCGAAAACGACGGCGAC------ 243

Qy         56 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 75
Db        244 ---------GATCCCGCTTCCTCGCCCCCTCCTCCGAGGGCGAGGAAGAAGCGGGTGCGC 294

Qy         76 GluThrIleSerGluProLysThrTyrValAsp-----------------LeuThrAsn 89
Db        295 GAGGCGTGCGCCGACGAGGAGGAGGAGGAGGAGGGCGAGGGGGCAGATGGGGCGGGTTCG 354

Qy         90 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 109
Db        355 CCGCGTCCCTCCGACGCCGCCGCCGCCAAGAAAAAAAAATTTGAGAATAATTTA------ 408

Qy        110 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 129
Db        409 AACAAGAAGACCTTCAAAATCATGACATACAATGTCTGGATCAGAGAGGATATTGAACTG 468

Qy        130 SerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePhe 149
Db        469 CACAGAAGGCTGGGCGCTCTTGGAGATCTTATTCAACTTCACAACCCAGACTTCATATGC 528

Qy        150 LeuGln--------------------GluValIlePro-ProTyr------TyrSerTy 160
Db        529 TTCCAGGTGCCAGGTAACAAGGAGAACGAGAAAATCGAACCCCTTCCTGTTGTGGTTCTA 588

Qy        160 rLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheTh 180
Db        589 CACCCTCCTGCGAGAGGACTCCCACCAAATCTTGAAG----------CG 627

Qy        180 rAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePr 200
Db        628 AAGCTTGCTGGC-AGGATGAGCAAATTACCGGTGAGCGAATCGAACCCCATCCCATTT-- 684

Qy        200 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn---Gl 219
Db        685 -TCGAAATCAATAATGAAGAGGGAGCTATGCGTAGCGGTCGTCAAAACGGGAGAGATCCA 743

Qy        219 uLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAs 239
Db        744 TTTGGCTGTGGGCACAAGTCACCTAGAGAGCCCC------TGCCCTCTACCTCCTCTGTG 797

Qy        239 nGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal----- 257
Db        798 GGATCTGAAGTACAGCGAGAAGCGAGTAGCTCAGGCAAAGCAGTCTCTGGAGATTCTTGG 857

Qy        258 ----------------IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 272
Db        858 ACAATTGCGCAATGCAATCTTCTGCGGCGACATGAACTGGGAAGACAAA---GTAGATGG 914

Qy        272 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 292
Db        915 TCCATTCCCTCTACCTGACGGCTGGATTGATGCTTGGGTGGAGCTG---AAGCCAGGTGA 971

Qy        292 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr---------- 308
Db        972 CAATGGTTGGACCTATGATACGAAAGCCAATGCCATGCTGTCAGCCAACTTCAAGCAGCA 1031

Qy        309 ----------------AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAl 323
Db       1032 GAAGAGGCCAGACCAGTTCGTGTGCAAGTTGTCGGATTTCAAGATC---------- 1077

Qy        323 aAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu-------- 340
Db       1078 -------------------------GACGACATTGAGATGATCGGAAAAGAGGCGATTCC 1112

Qy        341 ------------------------------------------------LysLeuAspCysGl 345
Db       1113 GGGTGTCGTGTACTACAAGGAGAAGATAGTTCGTAAGGAGTTTCATAAGTTGGAACTGCC 1172

Qy        345 yArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
Db       1173 TGTCCTACCCAGCAAACATTTTGGGCTTGTTCTGACAATTACTCTT 1218

RESULT 48
US-10-764-420-1725
; Sequence 1725, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
;  APPLICANT: Lum, Pek Yee
;  APPLICANT:  Tan, Yejun
;  APPLICANT:  Dai, Hongyue
;  TITLE OF INVENTION: Methods For Determining Whether An Agent
;  TITLE OF INVENTION:  Possesses A Defined Biological Activity
;  FILE REFERENCE: ROSA122057
;  CURRENT APPLICATION NUMBER: US/10/764,420
;  CURRENT FILING DATE:  2004-01-23
;  PRIOR APPLICATION NUMBER: US 60/442,797
;  PRIOR FILING DATE: 2003-01-24
;  PRIOR APPLICATION NUMBER: US 60/474,413
;  PRIOR FILING DATE: 2003-05-30
;  NUMBER OF SEQ ID NOS: 3683
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1725
;   LENGTH: 9737
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: 2002, 2030, 2068, 2082, 2164, 2240, 4178, 4212, 4214, 4385,
;   LOCATION: 9487
;   OTHER INFORMATION: n = A,T,C or G
US-10-764-420-1725

Alignment Scores:
Pred. No.:                3.25e-06        Length:          9737
Score:                    147.00          Matches:         86
Percent Similarity:       33.7%           Conservative:    53
Best Local Similarity:    20.8%           Mismatches:      161
Query Match:              7.7%            Indels:          113
DB:                       9               Gaps:            16

US-10-757-745-2 (1-362) x US-10-764-420-1725 (1-9737)

Qy         35 ValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp--------- 51
Db       7429 ATGGCGTGATGCCCACAGGCAGTGGGGCCCTGGCTCATACCGGAGAGGGACAGTGAAGGC 7488

Qy         52 ----------------------------------------------------TrpGluMetGlu 55
Db       7489 CCTGACTGACAGCTGGCTATGGGACTGTTTCTCCTGCAGTGAGTTCCATGGGAA------ 7542

Qy         56 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 75
Db       7543 ----------------------ACGGCACCAGTCGACTCTACAGTGCTCTCGCCAAGACCG 7581

Qy         76 GluThr------------IleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91
Db       7582 TCAACAGCAGTGCAGCTGCTCAGCACCCGGAGTACTTGGTGTCAACTGACCCCGAACATC 7641

Qy         92 ThrThrAspSer---------------------------------------------- 95
Db       7642 TGGAGCCCATCGATCCTAAGGAACTTCTGGAGGAATGCAGGGCTGTTCTGCACACTCGGC 7701

Qy         96 ------ThrThrSerLysIleSerProSerGluAspThrGln-GlnGluAsnGlySerMe 113
Db       7702 CACCCCGCTACCAGCGGATTTTGTGGACCTGAGGACAGATTGCTCCAGCAGCCACTCTCC 7761

Qy        113 tPheSerLeuIleThrTrpAsnIle------------------AspGlyLeuAsp----- 125
Db       7762 CATTCGCGTCATGCAGTGGAACATCCTCGCCCAAGCTCTCGGAGAAGGCAAAGACAACTT 7821

Qy        126 -------------------LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
Db       7822 TGTGCAGTGCCCCGTGGAAGCGCTCAAATGGGAAGAGAGGAAGTGCCTGATCCTGGAGGA 7881

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
Db       7882 GATCCTGGCTTACCAGCCAGACATACTGTGCCTCCAGGAAGTG---GACCACTACTTTGA 7938

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
Db       7939 CACCTTCCAGCCACTCCTCAGTAGA-----------CTGGGCTACCAAGGCACGTTTTT 7986

Qy        179 e-------------------------------------------------------ThrAl 181
Db       7987 CCCCAAGCCCTGGTCACCATGTCTAGATGTGGAACACAACAACGGTCCAGATGGCTGTGC 8046

Qy        181 aIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSe 201
Db       8047 CTTATTTTTTCTCCAAAACCGATTCAAGCTTATCAGCAGCACCAATATTAGGCTGACAGC 8106

Qy        201 rThrLysMetMetArgAsnLeuLeuCysVal-----------HisValAsnValSerGl 217
Db       8107 CATGACCCTAAAAACCAATCAGGTGGCCATCGCACAGACCCTGGAGTGCAAGGAGTCTGG 8166

Qy        217 yAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluAr 237
Db       8167 GCGACAGTTCTGCATTGCTGTCACCCACTTAAAAGCCCGCACTGGCTGGGAGCGGTTCCG 8226

Qy        237 gMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVa 257
Db       8227 GTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTGCAGAACATCACCCAGGGAGCAAAGAT 8286

Qy        257 l------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu------ValThrArgCy 273
Db       8287 CCCCCTGATCGTCTGCGGGGACTTCAACGCAGAGCCAACCGAAGAGGTCTACAAACACTT 8346

Qy        273 sGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------GlyLysPr 290
Db       8347 TGCGTCCTCCAGCCTCAACCTCAACAGCGCCTACAAGCTGCTGAGTCCCGATGGACAGTC 8406

Qy        290 oLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAl 310
Db       8407 GGAGCCTCCCTACACAACCTGG-----------------AAGATCCGGACCTCAGGCGA 8448

Qy        310 aCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIl 330
Db       8449 GTGTCGCCACACGCTGGACTATATCTGGTACTCCAGACATGCTCTGAGT---------GT 8499

Qy        330 eProArgSerLeuAspLeuLeu--------------GlyLeuGluLysLeuAspCysGl 345
Db       8500 CACGTCCGCCCTGGATCTACTCACTGAAGAACAGATTGGGCCCAACCGGCTCCCATCCTT 8559

Qy        345 yArgProSerAspHisTrpGlyLeuLeuCysAsn 357
Db       8560 CCATTACCCCTCGGACCACCTGTCCCTGGTGTGTGAC 8596

RESULT 49
US-10-425-115-66072/c
; Sequence 66072, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 66072
;   LENGTH: 932
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_160258C.1
US-10-425-115-66072

Alignment Scores:
Pred. No.:                6.17e-06        Length:          932
Score:                    132.00          Matches:         65
Percent Similarity:       42.0%           Conservative:    37
Best Local Similarity:    26.7%           Mismatches:      85
Query Match:              6.9%            Indels:          58
DB:                       8               Gaps:            11

US-10-757-745-2 (1-362) x US-10-425-115-66072 (1-932)

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSer---------SerAsnTyrGluIleIleThr 172
Db        928 CCATACATATATATGCTTATGCAAAAATCTGACTGGTGGCAACAATACACATGCTTGTTG 869

Qy        173 GlyHisGluGluGly-----------TyrPheThrAlaIleMetLeuLysLysSerArg 188
Db        868 TCACATGAGAATGCTATTCAGATGCCGTATTAC---------TGCATGCAGTTGAGCAAA 818

Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
Db        817 ATGCCGACGAAGCCATCTGACTGCATCCCATTTTCGAATTCAACAATGGGAAGAGAACTA 758

Qy        209 LeuCysValHisValAsnValSerGly---AsnGluLeuCysLeuMetThrSerHisLeu 227
```

```
Db        757 TGCATCGCAAGTGTTAGGACAGGAGAGATCACAAAGTTGATTTTGGCCACAACCCACCTC 698

Qy        228 GluSerThr-ArgGlyHis-----------------------AlaAlaGluArgMetAs 239

Db        697 GAGAGCCCATGCCCTGCACCTCCCAAGTGGGGATCAGATGTATAGCAAAGAACGTGTGGC 638

Qy        239 nGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePh 259

Db        637 TCAGGCAAAGCAATCCCTGGAATTCTGGGAGGCTTCCCCAAT-GCC---------ATACT 588

Qy        259 eAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly---------- 275

Db        587 CTGTGGGGACATGAACTGGGACGACAAG--------------GGAGACGGCCCTTTCCC 543

Qy        276 -LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTy 295

Db        542 TCTGCAGGATGGGTGGACTGATGCTTGGGTTGAGCTG---AAACCTGGAGAGGATGGCTG 486

Qy        295 rThrTrpAspThrGlnMetAsnSerAsnLeuGly-IleThrAlaAlaCysLysLeuArgP 315

Db        485 GACCTACGACACCAAGGCCAATGGCATGCTGTCAGGCAACCGCAAGCTGCAGAAG-CGAT 427

Qy        315 heAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuA 335

Db        426 TGGACCGGTTTGTGTGCAAGCTAGCA------------GACTTCAAGATGGACAGCATTG 379

Qy        335 spLeuLeuGlyLeuGlu------------------------------------------ 340

Db        378 AGATGATCGGGAAGGAGGCAATACCTGGTGTCTCCTACTTCAAGGAGAAGAAGATTCGGA 319

Qy        341 -----------LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysA 357

Db        318 AAGAGAACCAAAGGATTGAACTGCCTGTCTTTCCAAGTGACCACTTTGGGCTTGTTTTGA 259

Qy        357 snLeu 358

Db        258 CCATC 254
```

RESULT 50
US-10-425-115-105489
; Sequence 105489, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 105489
; LENGTH: 898
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(898)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_27708C.1
US-10-425-115-105489

Alignment Scores:

Pred. No.:	1.41e-05	Length:	898
Score:	129.00	Matches:	49
Percent Similarity:	41.9%	Conservative:	29
Best Local Similarity:	26.3%	Mismatches:	72
Query Match:	6.8%	Indels:	36
DB:	8	Gaps:	7

US-10-757-745-2 (1-362) x US-10-425-115-105489 (1-898)

```
Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        376 TCTGGAAATTCAAATAATCATTCGGATAATAAGACTGTC---AAAATAATGACATACAAT 432

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        433 GTATGGTTCCGAGAGGAATTGGAACTGATTAGAAGGATGAATGCTATTGGAGATCTTATC 492

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        493 CAGCATCACAGCCCAGATCTCATATGCTTCCAGGAGGTCACTCCAAACATTTACCTGCTT 552

Qy        161 LeuLysLysArgSer------SerAsnTyrGluIleIleThrGlyHis------------ 174

Db        553 TTCGAGAAATCAGACTGGTGGCAAGCATACAAATGCTCCTTGCCGCATAAGATGGCCATG 612

Qy        175 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194

Db        613 GAAAGATCATACTAC---------TGCATGCAGATGAGCAGGTTGCCGGTGAAGTCGTTT 663

Qy        195 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 214

Db        664 GACCGCAAACCATTCTTGAACTCAAAGATGGGACGGGAGCTGTGCGTAGCCGATGTGACC 723

Qy        215 ValSerGlyAsn---GluLeuCysLeuMetThrSerHisLeuGluSer------------ 229

Db        724 GTTGGAGGTTCGATCAAGCTGGTGCTGGCTACAAGCCACCTGGAGAGCCCATGCCCCGGA 783

Qy        230 --------------------------------ThrArgGlyHisAlaAlaGluArgMet 238

Db        784 CCCTCGACGTGNGACCAGATGTTCAGCAAGGAAAGAGTAGCTCAGGCAAGCGAGTCTTTG 843

Qy        239 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258

Db        844 AACACTCTGAGAGCTTTC----------------------------CGCAACGNCATA 873

Qy        259 PheAlaGlyAspThrAsn 264

Db        874 TTGTGCGGNGACATGAAC 891
```

RESULT 51
US-10-998-197-33
; Sequence 33, Application US/10998197
; Publication No. US20050222070A1
; GENERAL INFORMATION:
; APPLICANT: Dohrmann, Cord
; APPLICANT: Austen, Matthias
; TITLE OF INVENTION: Pancreas-Specific Proteins
; FILE REFERENCE: 2923-669
; CURRENT APPLICATION NUMBER: US/10/998,197
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: PCT/EP03/05700
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: EP 02020829.4
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: EP 02011963.2
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-998-197-33

Alignment Scores:

Pred. No.:	5.7e-05	Length:	1767
Score:	128.00	Matches:	83
Percent Similarity:	33.0%	Conservative:	55
Best Local Similarity:	19.9%	Mismatches:	164

```
Query Match:          6.7%      Indels:       116
DB:                   9         Gaps:         18

US-10-757-745-2 (1-362) x US-10-998-197-33 (1-1767)

Qy         20 ProGluValLysLysArgArgLeuLeuCysValGluPheAla---------SerValAla 36
              |||              |||||||||          |||           :::
Db        308 CCCAGGCCCGCATCCCCCCGGCTGCTGGCGGCGGCCTCGGCGGCCTCGGGCGCCGCGAGG 367

Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
              ||||||          |||      |||   :::                   :::
Db        368 TCGTGTTCCCGAACAGTG------TGTTCCATGGGAACCGGTACAAGCAGACTCTATAGT 421

Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
              ||||||                 :::       ||||||     :::   ||||||
Db        422 GCTCTC------------GCCAAGACACTGAACAGCAGCGCTGCCTCCCAGCACCCAGAG 469

Qy         77 ThrIleSerGluPro---------------------------------------------- 81
              :::     |||
Db        470 TATTTGGTGTCACCTGACCCAGAGCATCTGGAGCCCATTGATCCTAAAGAGCTTCTTGAG 529

Qy         82 ----------------------------------------LysThrTyrValAspLeu 87
                                                     :::   :::||||||||||
Db        530 GAATGCAGGGCCGTCCTGCACACCCGACCTCCCCGGTTCCAGAGGGATTTTGTGGATCTG 589

Qy         88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 107
                 :::              ||||||      |||
Db        590 AGGACAGAT------TGCCCTAGTACCCACCCACCTATC---------------------- 622

Qy        108 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle------------------ 121
                                  ::::::   |||||||||
Db        623 --------------------AGGGTTATGCAATGGAACATCCTCGCCCAAGCTCTTGGA 661

Qy        122 AspGlyLeuAsp----------------------LeuAsnAsnLeuSerGluArgAla 133
              :::|||   |||                                 ||||||
Db        662 GAAGGCAAAGACAACTTTGTACAGTGCCCTGTTGAAGCACTCAAATGGGAAGAAAGGAAA 721

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
                 :::      :::      |||   ||||||::::::   ||||||||||||
Db        722 TGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTCCAAGAGGTG 781

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
                    |||:::          :::          |||                 |||
Db        782 ---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA------------CTAGGC 826

Qy        174 HisGluGluGlyTyrPhe------------------------------------------ 179
              ::::::      :::|||
Db        827 TATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAACACAACAAT 886

Qy        180 ------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
                        |||:::        ::::::|||   ||||||   :::
Db        887 GGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTCAACAGTGCC 946

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal------------ 211
                 |||          :::      :::      |||   :::   :::
Db        947 AATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCACAGACCCTG 1006

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
                       ||||||   :::   |||:::      :::||||||::::::
Db       1007 GAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAAGCACGCACT 1066

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
              |||      |||      |||         :::|||:::   :::|||
Db       1067 GGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTGCAAAACATC 1126

Qy        252 ProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
                 :::  |||   :::      |||      ||||||   |||           |||
Db       1127 ACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAGCCAACAGAA 1186

Qy        270 ------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
                          |||:::          ::::::   |||
Db       1187 GAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTACAAGCTGCTG 1246

Qy        288 ---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304
                        |||:::   :::          ||||||
Db       1247 AGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG-----------------AAG 1288

Qy        305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 324
                 :::      :::      |||:::          |||   |||::::::          |||
Db       1289 ATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCTAAACATGCT 1348

Qy        325 GluGluGlyHisIleIleProArgSerLeuAspLeuLeu--------------GlyLeu 339
                              :::      :::||||||||||||                |||
Db       1349 CTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAGATTGGACCC 1399

Qy        340 GluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsn 357
                 :::|||          :::||||||||||||     |||:::|||:::
Db       1400 AACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGTGAC 1453

RESULT 52
US-10-887-553A-839
; Sequence 839, Application US/10887553A
; Publication No. US20050085436A1
; GENERAL INFORMATION:
;  APPLICANT: Garza, Dan
;  APPLICANT:  Li, Hao
;  TITLE OF INVENTION: Method to treat conditions associated
;  TITLE OF INVENTION:  with insulin signalling dysregulation
;  FILE REFERENCE: 4-33262
;  CURRENT APPLICATION NUMBER: US/10/887,553A
;  CURRENT FILING DATE:  2004-07-08
;  PRIOR APPLICATION NUMBER: 60/485,883
;  PRIOR FILING DATE: 2003-08-07
;  NUMBER OF SEQ ID NOS: 1208
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 839
;   LENGTH: 1971
;   TYPE: DNA
;   ORGANISM: human
US-10-887-553A-839

Alignment Scores:
Pred. No.:              6.81e-05      Length:         1971
Score:                  128.00        Matches:        83
Percent Similarity:     33.0%         Conservative:   55
Best Local Similarity:  19.9%         Mismatches:     164
Query Match:            6.7%          Indels:         116
DB:                     9             Gaps:           18

US-10-757-745-2 (1-362) x US-10-887-553A-839 (1-1971)

Qy         20 ProGluValLysLysArgArgLeuLeuCysValGluPheAla---------SerValAla 36
              |||              |||||||||          |||           :::
Db        308 CCCAGGCCCGCATCCCCCCGGCTGCTGGCGGCGGCCTCGGCGGCCTCGGGCGCCGCGAGG 367

Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
              ||||||          |||      |||   :::                   :::
Db        368 TCGTGTTCCCGAACAGTG------TGTTCCATGGGAACCGGTACAAGCAGACTCTATAGT 421

Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
              ||||||                 :::       ||||||     :::   ||||||
Db        422 GCTCTC------------GCCAAGACACTGAACAGCAGCGCTGCCTCCCAGCACCCAGAG 469

Qy         77 ThrIleSerGluPro---------------------------------------------- 81
              :::     |||
Db        470 TATTTGGTGTCACCTGACCCAGAGCATCTGGAGCCCATTGATCCTAAAGAGCTTCTTGAG 529

Qy         82 ----------------------------------------LysThrTyrValAspLeu 87
                                                     :::   :::||||||||||
Db        530 GAATGCAGGGCCGTCCTGCACACCCGACCTCCCCGGTTCCAGAGGGATTTTGTGGATCTG 589

Qy         88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 107
                 :::              ||||||      |||
Db        590 AGGACAGAT------TGCCCTAGTACCCACCCACCTATC---------------------- 622

Qy        108 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle------------------ 121
```

```
Db        623 ----------------------AGGGTTATGCAATGGAACATCCTCGCCCAAGCTCTTGGA 661

Qy        122 AspGlyLeuAsp-----------------------LeuAsnAsnLeuSerGluArgAla 133

Db        662 GAAGGCAAAGACAACTTTGTACAGTGCCCTGTTGAAGCACTCAAATGGGAAGAAAGGAAA 721

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153

Db        722 TGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTCCAAGAGGTG 781

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173

Db        782 ---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA------------CTAGGC 826

Qy        174 HisGluGluGlyTyrPhe------------------------------------------ 179

Db        827 TATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAACACAACAAT 886

Qy        180 ------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195

Db        887 GGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTCAACAGTGCC 946

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal------------ 211

Db        947 AATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCACAGACCCTG 1006

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231

Db       1007 GAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAAGCACGCACT 1066

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251

Db       1067 GGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTGCAAAACATC 1126

Qy        252 ProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269

Db       1127 ACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAGCCAACAGAA 1186

Qy        270 ------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287

Db       1187 GAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTACAAGCTGCTG 1246

Qy        288 ---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304

Db       1247 AGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------------AAG 1288

Qy        305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 324

Db       1289 ATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCTAAACATGCT 1348

Qy        325 GluGluGlyHisIleIleProArgSerLeuAspLeuLeu--------------GlyLeu 339

Db       1349 CTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAGATTGGACCC 1399

Qy        340 GluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsn 357

Db       1400 AACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGTGAC 1453
```

```
RESULT 53
US-11-097-143-9713
; Sequence 9713, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9713
;   LENGTH: 1350
;   TYPE: DNA
;   ORGANISM: DROSOPHILA
US-11-097-143-9713

Alignment Scores:
Pred. No.:                4.27e-05        Length:          1350
Score:                    127.50          Matches:         74
Percent Similarity:       34.3%           Conservative:    53
Best Local Similarity:    20.0%           Mismatches:      108
Query Match:              6.7%            Indels:          135
DB:                       10              Gaps:            19

US-10-757-745-2 (1-362) x US-11-097-143-9713 (1-1350)

Qy         87 LeuThrAsnGluGluThrThrAsp----------------------SerThrThrSerLys 99

Db        262 CTGGCGGACGAGGATGTAGCCGACGGACTGCGATGCCTCAAGCTGAATTCCGTGTCGAGG 321

Qy        100 Ile------SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 117

Db        322 GTTTGCAGTGCGCCAGTCGAG-----------------GGTGATGACATTCGATTGCTG 363

Qy        118 ThrTrpAsnIle---------------------------AspGlyLeu------------ 124

Db        364 CAGTGGAACATCCTATCACAAACTCTCGGGCAGCACAACGATGGATTTGTGCGCTGTCCC 423

Qy        125 ---AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyr 143

Db        424 GAGGAGGCTCTCACCTGGGAGCACCGCAAGTACCTGATTGTCCAGGAGATTCTGCAGAAC 483

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163

Db        484 CAGCCCGATGTGATTTGTCTGCAAGAGGTA------GACCACTTCAAGTTCCTGCAGACC 537

Qy        164 ------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe------ 179

Db        538 GTGTTGGGTAGCCAGAACTAC--------------------GCAGGCATCTTCTTCCCG 576

Qy        180 ------------------------------------------------------ThrAlaIle 182

Db        577 AAACCGGATTCCCCGTGCCTGTACATCGAGCAGAACAATGGACCCGATGGCTGTGCCATC 636

Qy        183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProProSerThr 202

Db        637 TTCTACAAGAGGGACAAACTGCAGCTGCAGGGATACGACACGAGGATCCTGGAAGTGTGG 696

Qy        203 LysMetMetArgAsnLeuLeuCysVal------------HisValAsnValSerGlyAsn 218

Db        697 CGGGTGCAGAGCAACCAGGTGGCCATCGCGGCCCGGCTGCGTATGCGATCCAGTGGTCGG 756

Qy        219 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis------AlaAlaGlu 236

Db        757 GAATTCTGTGTGGCCACCACCCACTTAAAGGCCCGCCATGGCGCCCTTTTGGCGAAATTA 816
```

```
Qy        237 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 256
Db        817 CGGAACGAACAGGGACGTGATCTCATTAGGTTTGTGAAGCAGTTTGCCGGAGATACGCCG 876

Qy        257 ValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu---------------------- 269
Db        877 CTGCTGCTCTGCGGCGATTTCAATGCCGAACCGGTGGAGCCCATCTATGCCACGATCTTG 936

Qy        270 ---------ValThrArgCysGlyGlyLeu------------------------------ 276
Db        937 GGATGTGATCTCCTCCGCTTGGGCAGCGCCTATGCCGACGTGAAGCTAGACCGCGAGGAG 996

Qy        277 ---------ProAsnAsnIleValAspValTrpGluPheLeuGlyLys------------ 289
Db        997 ATACTACATCCCAAC------GCGGATGTCGGTGAGTTCGTCGCGAAATCGATGAAACGC 1050

Qy        290 ---ProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
Db       1051 GAGCCACCCTACACC------ACCTGGAAA---------------------------- 1074

Qy        309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328
Db       1075 -------------------------------ATTCGCGAAGAGGGCGAGGAATGTCAC 1101

Qy        329 IleIle-----------------ProArgSerLeuAspLeu----------------- 336
Db       1102 ACCATCGACTACGTCTTTTACACACCCGATCGCCTCAAGATCAAAAACTGTCTAGACTTT 1161

Qy        337 --------------LeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
Db       1162 CCGGCCGGAGAGCAAATTGGCAAGAATCGCACACCCAGCTTCCAATATCCATCGGATCAC 1221

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIle 361
Db       1222 TTTTCCCTGGTGTGTGACTTTGAGCTACTG 1251
```

RESULT 54
US-10-087-192-1141/c
; Sequence 1141, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1141
; LENGTH: 90043
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(90043)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1141

Alignment Scores:

Pred. No.:	0.061	Length:	90043
Score:	126.00	Matches:	57
Percent Similarity:	37.9%	Conservative:	42
Best Local Similarity:	21.8%	Mismatches:	114
Query Match:	6.6%	Indels:	48
DB:	5	Gaps:	10

US-10-757-745-2 (1-362) x US-10-087-192-1141 (1-90043)

```
Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      85949 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 85890

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db      85889 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 85842

Qy        171 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 179
Db      85841 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 85785

Qy        180 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
Db      85784 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 85725

Qy        193 SerGlnGluIleIleProProProSerThrLysMetMetArgAsnLeuLeuCysVal--- 211
Db      85724 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 85665

Qy        212 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
Db      85664 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 85605

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
Db      85604 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 85545

Qy        249 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 266
Db      85544 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 85485

Qy        267 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
Db      85484 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 85425

Qy        285 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
Db      85424 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 85377

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321
Db      85376 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 85323

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu------------ 337
Db      85322 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 85272

Qy        338 ---GlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356
Db      85271 ATTGGGCCCAACCGGCTCCCATCCTTCCATTACCCCTCGGACCACCTGTCCCTGGTGTGT 85212

Qy        357 Asn 357
Db      85211 GAC 85209
```

RESULT 55
US-10-369-493-24191
; Sequence 24191, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039

```
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24191
;   LENGTH: 1644
;   TYPE: DNA
;   ORGANISM: Xenorhabdus nematophilus
US-10-369-493-24191

Alignment Scores:
Pred. No.:                0.000544        Length:          1644
Score:                    120.00          Matches:         73
Percent Similarity:       40.7%           Conservative:    65
Best Local Similarity:    21.5%           Mismatches:      123
Query Match:              6.3%            Indels:          78
DB:                       6               Gaps:            18

US-10-757-745-2 (1-362) x US-10-369-493-24191 (1-1644)

Qy         53 GluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu--- 71
Db         28 GAAGCCTGGAAAGCATTGCAGCAACATTTTGAG---CAGATAAAAGATACTCACCTGCGA 84

Qy         72 ------------------GluArgArgProGluThrIleSer---------GluProLys 82
Db         85 GATCTGTTTGAACAGGACAAGGCGCGTTTTGCGACATTCTCTGCCACATTTGACGATCAG 144

Qy         83 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 102
Db        145 ATGCTGGTGGATTATTCCAAAAACCGTATTACCGCAGAAACACTGGAGAAATTACAGGCT 204

Qy        103 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 122
Db        205 CTGGCGAAAGAAACTGATTTATCCGGGGCTATTCATAGCATGTTTTCC------------ 252

Qy        123 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 142
Db        253 GGTGAAAAGATTAACCGTACTGAAGATCGCGCT---------GTATTGCATATTGCTTTG 303

Qy        143 ------------TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 158
Db        304 CGTAACCGCAGTAACTCTCCGATTAAGGTTGATGAGGAAGATGTTATGCCGCAGGTCAAT 363

Qy        159 SerTyrLeu---LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGly 177
Db        364 GCGGTACTGGCAAAAATGAAATCATTCAGTGAGCGCATTATTAGTGGGGAGTGGAAAGGT 423

Qy        178 TyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle 197
Db        424 TATACCGGTAAAGCTATTACTGATGTGGTAAACATCGGCATTGGTGGTTCAGATCTTGGC 483

Qy        198 ProPheProSerThr---------LysMetMetArgAsnLeuLeuCysValHisVal--- 213
Db        484 CCATATATGGTGACAGAGGCACTGAAAGCCTATAAGAACCACCTTAATATGCACTTTGTA 543

Qy        214 ---AsnValSerGlyAsnGluLeuCys-------------------------------- 221
Db        544 TCCAATGTGGATGGCACTCATATCGCGGAAACACTGAAGTCTCTGGATCCGGAAACCACG 603

Qy        222 ---------------------LeuMetThrSerHisLeuGluSerThrArgGlyHisAla 234
Db        604 CTGTTCCTGATTGCATCGAAGACTTTTACCACGCAGGAAACGATGACTAACGCACATTCT 663

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
Db        664 GCC---CGTGATTGGTTCTTGCAAAGTGCAGTTGATGAAGTACATGTTGCCAAACATTTT 720

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
Db        721 GCGGCACTT------------TCAACGAAT------GAACAGGAAGTCAGCAAGTTCGGT 762

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294
Db        763 ATTGATCCGCAAAACATGTTTGAGTTCTGGGACTGGGTCGGCGGC------------CGT 810

Qy        295 TyrThr---TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeu 313
Db        811 TATTCTCTTTGG--------------TCTGCGATTGGTTTATCCATTGTCCTGTCCATT 855

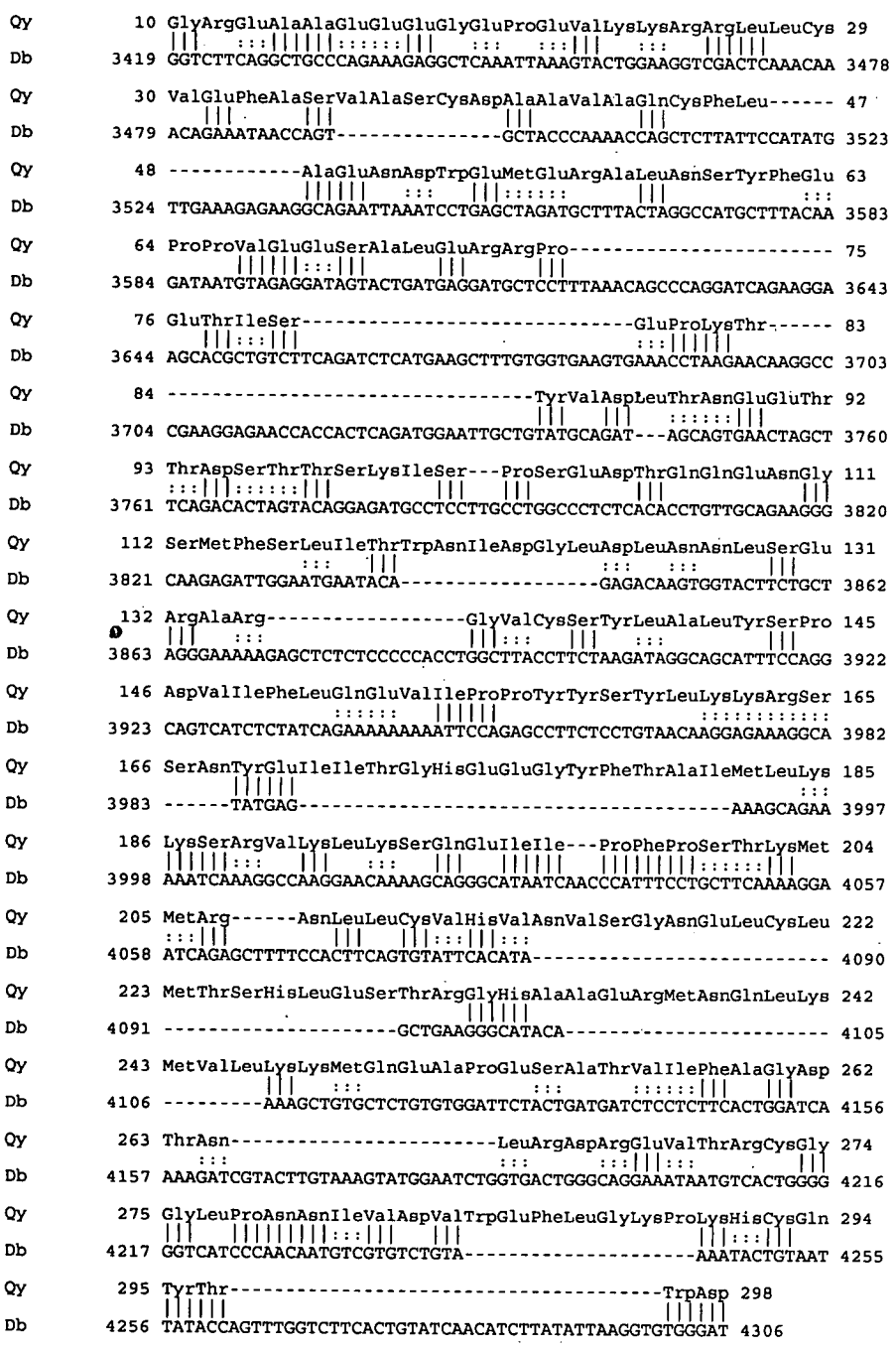
Qy        314 ArgPheAspArgIle---------------------------------------PhePhe 320
Db        856 GGATATGATAATTTCGAGCAATTGCTGAATGGTGCCCATGCTATGGATAAGCACTTTGCA 915

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeu 339
Db        916 CAAACTGCTCCTGAACAGAAT------ATTCCGGTTATTCTGGCGCTGATTGGCATC 966

RESULT 56
US-10-473-574-33
; Sequence 33, Application US/10473574
; Publication No. US20040116670A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE CORPORATION; HAFALIA, April J.A.;
;  APPLICANT:  TANG, Y. Tom; YUE, Henry;
;  APPLICANT:  KHAN, Farrah A.; ISON, Craig H.;
;  APPLICANT:  BAUGHN, Mariah R.; WARREN, Bridget A.;
;  APPLICANT:  DUGGAN, Brendan M.; THANGAVELU, Kavitha;
;  APPLICANT:  HONCHELL, Cynthia D.; AZIMZAI, Yalda;
;  APPLICANT:  ELLIOTT, Vicki S.; BURFORD, Neil;
;  APPLICANT:  DING, Li; YUE, Huibin;
;  APPLICANT:  BECHA, Shanya; EMERLING, Brooke M.;
;  APPLICANT:  RICHARDSON, Thomas W.; LEE, Soo Yeun;
;  APPLICANT:  BANDMAN, Olga; LAL, Preeti G.;
;  APPLICANT:  LEE, Sally; GIETZEN, Kimberly J.;
;  APPLICANT:  CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;
;  APPLICANT:  LEE, Ernestine A.; SWARNAKAR, Anita;
;  APPLICANT:  RING, Huijun Z.; JONES, Karen Anne
;  TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
;  FILE REFERENCE: PF-0918 USN
;  CURRENT APPLICATION NUMBER: US/10/473,574
;  CURRENT FILING DATE:  2003-09-29
;  PRIOR APPLICATION NUMBER: PCT/US02/09288
;  PRIOR FILING DATE: 2002-03-25
;  PRIOR APPLICATION NUMBER: US 60/294,451
;  PRIOR FILING DATE: 2001-05-29
;  PRIOR APPLICATION NUMBER: US 60/291,870
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: US 60/290,518
;  PRIOR FILING DATE: 2001-05-10
;  PRIOR APPLICATION NUMBER: US 60/288,609
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: US 60/283,769
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: US 60/281,323
;  PRIOR FILING DATE: 2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/280,508
;  PRIOR FILING DATE: 2001-03-29
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PERL Program
; SEQ ID NO 33
;   LENGTH: 5299
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7948948CB1
US-10-473-574-33

Alignment Scores:
Pred. No.:                0.00763         Length:          5299
Score:                    117.50          Matches:         76
Percent Similarity:       34.7%           Conservative:    48
Best Local Similarity:    21.3%           Mismatches:      104
Query Match:              6.2%            Indels:          129
DB:                       7               Gaps:            18

US-10-757-745-2 (1-362) x US-10-473-574-33 (1-5299)
```

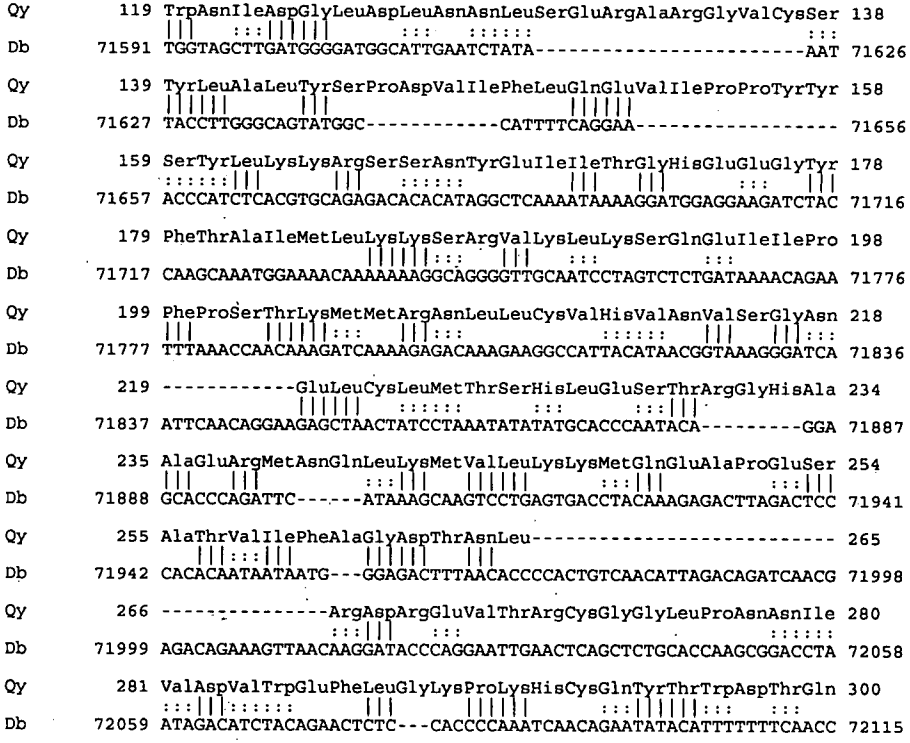


RESULT 57
US-10-939-321-1
; Sequence 1, Application US/10939321
; Publication No. US20050221336A1
; GENERAL INFORMATION:
; APPLICANT: MARTIGNETTI, John
; APPLICANT: DOWLING, Oonagh
; APPLICANT: DIFEO, Analisa
; APPLICANT: RAMIREZ, Maria Celeste M.
; APPLICANT: TUKEL, Turgut
; APPLICANT: GLUCKSMAN, Marc J.
; TITLE OF INVENTION: Mutations in Capillary Morphogenesis Gene-2 (Cmg-2)
; FILE REFERENCE: 02420/100N211-US1
; CURRENT APPLICATION NUMBER: US/10/939,321
; CURRENT FILING DATE: 2004-09-09
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 168050
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-939-321-1

Alignment Scores:

Pred. No.:	2.09	Length:	168050
Score:	117.50	Matches:	63
Percent Similarity:	42.9%	Conservative:	48
Best Local Similarity:	24.3%	Mismatches:	100
Query Match:	6.2%	Indels:	49
DB:	9	Gaps:	11

US-10-757-745-2 (1-362) x US-10-939-321-1 (1-168050)



```
Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              :::|||          |||   |||:::  ||||||:::|||   :::   |||||||||
Db      72116 CTGAACTTTTTATAGGGT---ACAAGCAATTGCAAAATAAGAGGCAATGAAATATTCTTT 72172

Qy        321 ArgAlaAlaAlaGluGluGlyHisIle------------IleProArgSerLeuAspLeu 336
              :::      |||      |||   :::             :::::::::   :::
Db      72173 AAAATATTTGCGGTGTTTGGTTTTTTGTTCTTGTGATAGTTTACTCAGAATGTGGGAATT 72232

Qy        337 LeuGlyLeuGluLysLeuAspCysGlyArgProProSerAspHisTrpGlyLeuLeu 355
              :::   :::|||  ||||||     |||     ||||||||||||
Db      72233 GAACAATGAGATCACATGGACACAGGAAGGGGAATATCACACTC-TGGGGACTGTTG 72288

RESULT 58
US-10-484-577-662/c
; Sequence 662, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
;  APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft
;  TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGT1A
;  FILE REFERENCE: F2285PCT-1
;  CURRENT APPLICATION NUMBER: US/10/484,577
;  CURRENT FILING DATE:  2004-01-22
;  PRIOR APPLICATION NUMBER: PCT/EP 02/08220
;  PRIOR FILING DATE: 2002-07-23
;  PRIOR APPLICATION NUMBER: EP 01 11 7608.8
;  PRIOR FILING DATE: 2001-07-23
;  PRIOR APPLICATION NUMBER: EP 02011710.7
;  PRIOR FILING DATE: 2002-05-24
;  NUMBER OF SEQ ID NOS: 683
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 662
;   LENGTH: 96960
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-484-577-662

Alignment Scores:
Pred. No.:                0.992          Length:          96960
Score:                    117.00         Matches:         59
Percent Similarity:       44.4%          Conservative:    41
Best Local Similarity:    26.2%          Mismatches:      91
Query Match:              6.2%           Indels:          35
DB:                       8              Gaps:            9

US-10-757-745-2 (1-362) x US-10-484-577-662 (1-96960)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::           |||:::
Db      34966 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGAATCA 34908

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||      :::::::::|||   ||||||:::|||      ||||||:::
Db      34907 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 34854

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::           |||   |||      :::
Db      34853 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 34797

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||           ::::::|||      |||   ::::::
Db      34796 CAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 34755

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                    ||||||:::   |||
Db      34754 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 34695

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::      :::         |||       ||||||:::   |||:::
Db      34694 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 34635

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::  |||   |||:::              ||||||  ::::::      :::
Db      34634 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 34575

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          ||||   |||      ::::|||   ||||||
Db      34574 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 34530

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266
                 :::|||       :::|||   |||:::|||      ||||||   |||
Db      34529 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 34473

Qy        267 AspArgGluValThrArgCysGlyGlyLeuProAsnAsn---------------------- 279
                    :::   |||      |||      |||
Db      34472 CTGTCAACATTAGACAGATCAACGGGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 34413

Qy        280 ----------------------IleValAspValTrpGluPheLeuGlyLysProLysHis 292
                                   :::::::|||::::::      |||      ||||||
Db      34412 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 34356

Qy        293 CysGlnTyrThrTrp 297
                :::|||||:::
Db      34355 ACAGAATATACATTT 34341

RESULT 59
US-10-473-574-31
; Sequence 31, Application US/10473574
; Publication No. US20040116670A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE CORPORATION; HAFALIA, April J.A.;
;  APPLICANT:  TANG, Y. Tom; YUE, Henry;
;  APPLICANT:  KHAN, Farrah A.; ISON, Craig H.;
;  APPLICANT:  BAUGHN, Mariah R.; WARREN, Bridget A.;
;  APPLICANT:  DUGGAN, Brendan M.; THANGAVELU, Kavitha;
;  APPLICANT:  HONCHELL, Cynthia D.; AZIMZAI, Yalda;
;  APPLICANT:  ELLIOTT, Vicki S.; BURFORD, Neil;
;  APPLICANT:  DING, Li; YUE, Huibin;
;  APPLICANT:  BECHA, Shanya; EMERLING, Brooke M.;
;  APPLICANT:  RICHARDSON, Thomas W.; LEE, Soo Yeun;
;  APPLICANT:  BANDMAN, Olga; LAL, Preeti G.;
;  APPLICANT:  LEE, Sally; GIETZEN, Kimberly J.;
;  APPLICANT:  CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;
;  APPLICANT:  LEE, Ernestine A.; SWARNAKAR, Anita;
;  APPLICANT:  RING, Huijun Z.; JONES, Karen Anne
;  TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
;  FILE REFERENCE: PF-0918 USN
;  CURRENT APPLICATION NUMBER: US/10/473,574
;  CURRENT FILING DATE:  2003-09-29
;  PRIOR APPLICATION NUMBER: PCT/US02/09288
;  PRIOR FILING DATE: 2002-03-25
;  PRIOR APPLICATION NUMBER: US 60/294,451
;  PRIOR FILING DATE: 2001-05-29
;  PRIOR APPLICATION NUMBER: US 60/291,870
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: US 60/290,518
;  PRIOR FILING DATE: 2001-05-10
;  PRIOR APPLICATION NUMBER: US 60/288,609
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: US 60/283,769
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: US 60/281,323
;  PRIOR FILING DATE: 2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/280,508
;  PRIOR FILING DATE: 2001-03-29
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PERL Program
; SEQ ID NO 31
;   LENGTH: 5322
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 6260407CB1
US-10-473-574-31

Alignment Scores:
```

```
Pred. No.:              0.0393          Length:         5322
Score:                  112.00          Matches:        74
Percent Similarity:     32.9%           Conservative:   45
Best Local Similarity:  20.4%           Mismatches:     111
Query Match:            5.9%            Indels:         132
DB:                     7               Gaps:           17

US-10-757-745-2 (1-362) x US-10-473-574-31 (1-5322)

Qy         10 GlyArgGluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCys 29
Db       3480 GGTCTTCAGGCTGCCCAGAAAGAGGCTCAAATTAAAGTACTGGAAGGTCGACTCAAACAA 3539

Qy         30 ValGluPheAlaSerValAla---------------------------------------- 36
Db       3540 ACAGAAATAACCAGTGCTACCCAAAACCAGCTCTTATTCCATATGTTGAAAGAGAAGGCA 3599

Qy         37 ------------SerCysAspAlaAlaValAlaGlnCysPheLeu--------------- 47
Db       3600 GAATTAAATCCTGAGCTAGATGCTTTACTAGGCCATGCTTTACAAGATCTAGATAGCGTA 3659

Qy         48 ------------AlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGlu 63
Db       3660 CCATTAGAAAATGTAGAGGATAGTACTGATGAGGATGCTCCTTTAAACAGC--------- 3710

Qy         64 ProProValGluGluSerAlaLeuGlu------------------------------- 72
Db       3711 CCAGGATCAGAAGGAAGCACGCTGTCTTCAGATCTCATGAAGCTTTGTGGTGAAGTGAAA 3770

Qy         73 ---------------ArgArgProGluThrIleSerGluProLysThrTyrValAspLeu 87
Db       3771 CCTAAGAACAAGGCCCGAAGGAGAACCACCACTCAGATGGAATTGCTGTATGCAGAT--- 3827

Qy         88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSer---ProSerGluAspThr 106
Db       3828 AGCAGTGAACTAGCTTCAGACACTAGTACAGGAGATGCCTCCTTGCCTGGCCCTCTCACA 3887

Qy        107 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 126
Db       3888 CCTGTTGCAGAAGGGCAAGAGATTGGAATGAATACA----------------GAGACA 3929

Qy        127 AsnAsnLeuSerGluArgAlaArg----------------GlyValCysSerTyrLeu 140
Db       3930 AGTGGTACTTCTGCTAGGGAAAAAGAGCTCTCTCCCCCACCTGGCTTACCTTCTAAGATA 3989

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
Db       3990 GGCAGCATTTCCAGGCAGTCATCTCTATCAGAAAAAAAAATTCCAGAGCCTTCTCCTGTA 4049

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
Db       4050 ACAAGGAGAAAGGCA------TATGAG---------------------------------- 4070

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle---ProPhe 199
Db       4071 ------AAAGCAGAAAAATCAAAGGCCAAGGAACAAAAGCAGGGCATAATCAACCCATTT 4124

Qy        200 ProSerThrLysMetMetArg------AsnLeuLeuCysValHisValAsnValSerGly 217
Db       4125 CCTGCTTCAAAAGGAATCAGAGCTTTTCCACTTCAGTGTATTCACATA------------ 4172

Qy        218 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 237
Db       4173 ----------------------------------GCTGAAGGGCATACA--------- 4187

Qy        238 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
Db       4188 ----------------------AAAGCTGTGCTCTGTGTGGATTCTACTGATGATCTC 4223

Qy        258 IlePheAlaGlyAspThrAsn----------------------LeuArgAspArgGlu 269
Db       4224 CTCTTCACTGGATCAAAAGATCGTACTTGTAAAGTATGGAATCTGGTGACTGGGCAGGAA 4283

Qy        270 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLys 289
Db       4284 ATAATGTCACTGGGGGGTCATCCCAACAATGTCGTGTCTGTA---------------- 4325

Qy        290 ProLysHisCysGlnTyrThr---------------------------------------- 296
Db       4326 ---AAATACTGTAATTATACCAGTTTGGTCTTCACTGTATCAACATCTTATATTAAGGTG 4382

Qy        297 TrpAsp 298
Db       4383 TGGGAT 4388

RESULT 60
US-09-738-626-2561
; Sequence 2561, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 2561
;   LENGTH: 1608
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-2561

Alignment Scores:
Pred. No.:              0.00757         Length:         1608
Score:                  111.00          Matches:        99
Percent Similarity:     36.2%           Conservative:   55
Best Local Similarity:  23.2%           Mismatches:     135
Query Match:            5.8%            Indels:         138
DB:                     3               Gaps:           24

US-10-757-745-2 (1-362) x US-09-738-626-2561 (1-1608)

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPhe---AlaSerValAlaSerCysAsp 39
Db         13 GAAGTGAAGAAGCGTAATCTCCTCGTAGCTCCCCTCACTGCTTCCCTGGTGTTCTGTAAC 72

Qy         40 AlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsn 59
Db         73 TTGGCTGTT---------------GCAGCAAACGCCGTTGAAGTTGAGGCC--------- 108

Qy         60 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 79
Db        109 ---------GAATCACCAGTTGTCATCAATGAAGTTGAATCCAACAGCGACCCAGTTGGT 159

Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr------ 97
Db        160 GAC---------TGGGTGGAGTTGGCTAACACCGACAACAACAACTCCATCGACATTTCC 210

Qy         98 ------SerLysIleSerProSerGluAsp---------------------------- 105
```

```
Db        211 GGTTGGTCCTTAGTCGATGACAAGGAAGACCTGGAAAATGCCCTCGTCCTTCCTGAAGGC 270
Qy        106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
Db        271 ACTGAGATTGAGTCCGGTGGATACTTTGTTATCTACACCGACTCTGCTGATTACGTGCCT 330
Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
Db        331 ACCAACAAC------------ACCTTTGGTGGCCAGGAATACTTCGGCCTCGGCAAAGAT 378
Qy        146 AspValIlePheLeuGln------------GluValIleProProTyr---TyrSerTyr 160
Db        379 GACACTGTTACTCTGCGCAACGCTGAAGGCGAAGTAGTTGCTACCTATTCCTGGAAGGAT 438
Qy        161 LeuLysLysArgSerSerAsn---TyrGluIleIle---------------------- 171
Db        439 CTGGGCGAGCACGCAGAAAACACCTATGGTCGCATCCCAGATATGACTGGTGATTTCGCA 498
Qy        172 ----------------------------------------------------ThrGlyHisGlu 175
Db        499 AACACCGGCGTTCCAACCCCAGGTGCAAAGAATGTTGCTGCTGAAGGCTCCGGCGAAGAA 558
Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
Db        559 GAAGGC--------------GTTGTTGCAAACGCCCAGCTTCCATTCCACAACGTTGAA 603
Qy        196 IleIlePro-PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVa 215
Db        604 ATCACCCCAATTCACCTCG------------------------------GTGGAGAT 630
Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
Db        631 TTCACCGGTGAAGATATGTCCGGCGTTGATTTCGATGCAA---ACAGCACCGCATGGATC 687
Qy        235 aGluArgMetAsn---------------------------------------------- 239
Db        688 GCCAACAATGACATTGGAAAGATCTACTCCCTCGCCCACGACATAGCTAACAACACCTAC 747
Qy        240 ------------GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
Db        748 AAGCTGACTGGCGAATGGGAAACCGGCTACCCAGAAGGCGGCGGAGAG-CCAGACGCTGA 806
Qy        255 aThrValIlePheAla------GlyAsp------------------ThrAsnLeuArgAs 267
Db        807 AGGCATCGTCGCAGCTACCAACGGTGACATCTACCTGTCCACCGAGCGCAACAACGCTGA 866
Qy        267 pArgGluValThrArgCysGlyGlyLeu------------------------------Pr 277
Db        867 CAAGAACGTCTCTCGCCCATCCATCCTGCGTTTTGCTACCCCAACTGGCAAGACTGGCGT 926
Qy        277 oAsnAsnIleValAspValTrp---------GluPheLeuGlyLysProLysHisCysGl 294
Db        927 ACAAAACGCAGTTCAGGAATGGGACCTGTCTGAGTTCGTCGGC---------------- 969
Qy        294 nTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuAr 314
Db        970 ----------GACATTCAGCCCAATGGT------GGTCTTGAGGCAATCGCGCAGCTCGA 1013
Qy        314 gPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLe 334
Db       1014 G---GACAACATCTTCGTTGTCGGTGTCGAAGAGACAGGTGATGTCATC------GTTGT 1064
Qy        334 uAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLe 354
Db       1065 TGATCTTTCCGCTGACCAGCCAGTTCTGGTTCAAAGGTACGAATCTTCCTTCGACGGTGT 1124
Qy        354 uLeuCysAsnLeuAsp 359
Db       1125 CATG---TCGCTTGAT 1137

RESULT 61
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
;   LENGTH: 3309400
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Alignment Scores:
Pred. No.:                1.81e+03        Length:          3309400
Score:                    111.00          Matches:         99
Percent Similarity:       36.2%           Conservative:    55
Best Local Similarity:    23.2%           Mismatches:      135
Query Match:              5.8%            Indels:          138
DB:                       3               Gaps:            24

US-10-757-745-2 (1-362) x US-09-738-626-1 (1-3309400)

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPhe---AlaSerValAlaSerCysAsp 39
Db    2477656 GAAGTGAAGAAGCGTAATCTCCTCGTAGCTCCCCTCACTGCTTCCCTGGTGTTCTGTAAC 2477715
Qy         40 AlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsn 59
Db    2477716 TTGGCTGTT--------------GCAGCAAACGCCGTTGAAGTTGAGGCC--------- 2477751
Qy         60 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 79
Db    2477752 ---------GAATCACCAGTTGTCATCAATGAAGTTGAATCCAACAGCGACCCAGTTGGT 2477802
Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr------ 97
Db    2477803 GAC---------TGGGTGGAGTTGGCTAACACCGACAACAACAACTCCATCGACATTTCC 2477853
Qy         98 ------SerLysIleSerProSerGluAsp------------------------------ 105
Db    2477854 GGTTGGTCCTTAGTCGATGACAAGGAAGACCTGGAAAATGCCCTCGTCCTTCCTGAAGGC 2477913
Qy        106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
Db    2477914 ACTGAGATTGAGTCCGGTGGATACTTTGTTATCTACACCGACTCTGCTGATTACGTGCCT 2477973
Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
Db    2477974 ACCAACAAC------------ACCTTTGGTGGCCAGGAATACTTCGGCCTCGGCAAAGAT 2478021
Qy        146 AspValIlePheLeuGln------------GluValIleProProTyr---TyrSerTyr 160
Db    2478022 GACACTGTTACTCTGCGCAACGCTGAAGGCGAAGTAGTTGCTACCTATTCCTGGAAGGAT 2478081
Qy        161 LeuLysLysArgSerSerAsn---TyrGluIleIle---------------------- 171
```

```
Db    2478082 CTGGGCGAGCACGCAGAAAACACCTATGGTCGCATCCCAGATATGACTGGTGATTTCGCA 2478141

Qy        172 ----------------------------------------------ThrGlyHisGlu 175

Db    2478142 AACACCGGCGTTCCAACCCCAGGTGCAAAGAATGTTGCTGCTGAAGGCTCCGGCGAAGAA 2478201

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195

Db    2478202 GAAGGC--------------GTTGTTGCAAACGCCCAGCTTCCATTCCACAACGTTGAA 2478246

Qy        196 IleIlePro-PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVa 215

Db    2478247 ATCACCCCAATTCACCTCG------------------------------GTGGAGAT 2478273

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235

Db    2478274 TTCACCGGTGAAGATATGTCCGGCGTTGATTTCGATGCAA---ACAGCACCGCATGGATC 2478330

Qy        235 aGluArgMetAsn---------------------------------------------- 239

Db    2478331 GCCAACAATGACATTGGAAAGATCTACTCCCTCGCCCACGACATAGCTAACAACACCTAC 2478390

Qy        240 ------------GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255

Db    2478391 AAGCTGACTGGCGAATGGGAAACCGGCTACCCAGAAGGCGGCGGAGAG-CCAGACGCTGA 2478449

Qy        255 aThrValIlePheAla------GlyAsp-----------------ThrAsnLeuArgAs 267

Db    2478450 AGGCATCGTCGCAGCTACCAACGGTGACATCTACCTGTCCACCGAGCGCAACAACGCTGA 2478509

Qy        267 pArgGluValThrArgCysGlyGlyLeu----------------------------Pr 277

Db    2478510 CAAGAACGTCTCTCGCCCATCCATCCTGCGTTTTGCTACCCCAACTGGCAAGACTGGCGT 2478569

Qy        277 oAsnAsnIleValAspValTrp---------GluPheLeuGlyLysProLysHisCysGl 294

Db    2478570 ACAAAACGCAGTTCAGGAATGGGACCTGTCTGAGTTCGTCGGC---------------- 2478612

Qy        294 nTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuAr 314

Db    2478613 ----------GACATTCAGCCCAATGGT------GGTCTTGAGGCAATCGCGCAGCTCGA 2478656

Qy        314 gPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLe 334

Db    2478657 G---GACAACATCTTCGTTGTCGGTGTCGAAGAGACAGGTGATGTCATC------GTTGT 2478707

Qy        334 uAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLe 354

Db    2478708 TGATCTTTCCGCTGACCAGCCAGTTCTGGTTCAAAGGTACGAATCTTCCTTCGACGGTGT 2478767

Qy        354 uLeuCysAsnLeuAsp 359

Db    2478768 CATG---TCGCTTGAT 2478780

RESULT 62
US-10-450-763-20786
; Sequence 20786, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom

; SEQ ID NO 20786
;   LENGTH: 2277
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2226)..(2276)
;   OTHER INFORMATION: 64% homologous to Homo sapiens putative p150,accession number
;   OTHER INFORMATION: U93569,Smith-Waterman Score=68.
US-10-450-763-20786

Alignment Scores:
Pred. No.:              0.0154      Length:          2277
Score:                  110.50      Matches:         73
Percent Similarity:     43.0%       Conservative:    55
Best Local Similarity:  24.5%       Mismatches:      112
Query Match:            5.8%        Indels:          59
DB:                     9           Gaps:            15

US-10-757-745-2 (1-362) x US-10-450-763-20786 (1-2277)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41

Db        186 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 233

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61

Db        234 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 269

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81

Db        270 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 326

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100

Db        327 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 368

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119

Db        369 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 425

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139

Db        426 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 476

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159

Db        477 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 518

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179

Db        519 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 578

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199

Db        579 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 638

Qy        199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215

Db        639 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 698

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235

Db        699 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 749

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255

Db        750 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 803

Qy        255 aThrValIlePheAlaGlyAsp-----------------------ThrAsnLeuArgA 267

Db        804 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 860
```

```
Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
Db        861 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 919

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db        920 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 966
```

RESULT 63
US-10-450-763-11636
; Sequence 11636, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 11636
; LENGTH: 2606
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (265)..(726)
; OTHER INFORMATION: 60% homologous to Cloning vector pTM30 repressor LacI-
; OTHER INFORMATION: q,accession number AF023620,Smith-Waterman Score=361.
US-10-450-763-11636

Alignment Scores:

Pred. No.:	0.0192	Length:	2606
Score:	110.50	Matches:	73
Percent Similarity:	42.6%	Conservative:	54
Best Local Similarity:	24.5%	Mismatches:	113
Query Match:	5.8%	Indels:	59
DB:	9	Gaps:	15

US-10-757-745-2 (1-362) x US-10-450-763-11636 (1-2606)

```
Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
Db        506 ATCAGGCAGCAACATTTGCTG---------TTCACCAATATCTGC---TGTTCTGCAGCC 553

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
Db        554 TCAGCTGCTGATACCCAGGCGAACAGGGTCTGGAGT----------------------- 589

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
Db        590 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 646

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
Db        647 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAAACCCATCTG 688

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
Db        689 TATGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 745

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
Db        746 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 796

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
```

```
Db        797 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 838

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
Db        839 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 898

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
Db        899 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 958

Qy        199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215
Db        959 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1018

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
Db       1019 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1069

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
Db       1070 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1123

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267
Db       1124 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1180

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
Db       1181 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1239

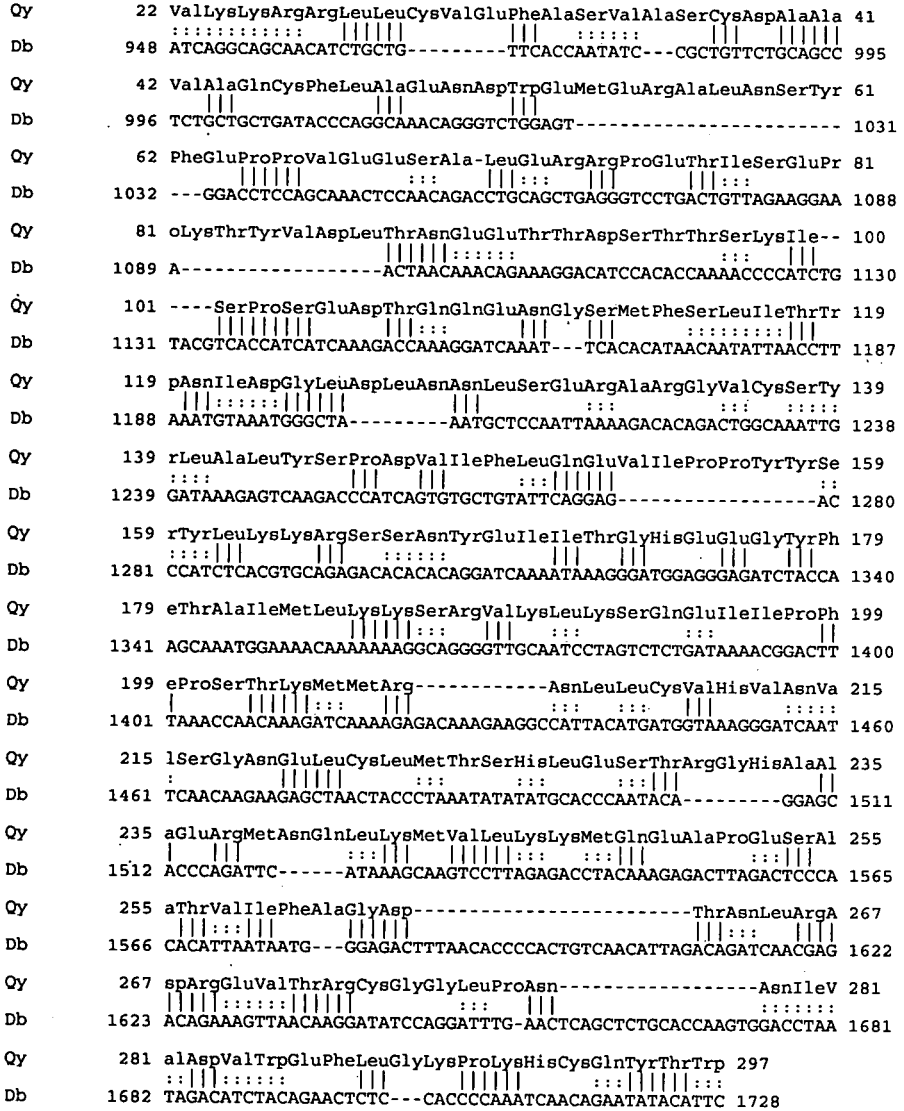
Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db       1240 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1286
```

RESULT 64
US-10-450-763-2332
; Sequence 2332, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 2332
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2614)..(4473)
; OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
; OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-2332

Alignment Scores:

Pred. No.:	0.0577	Length:	5130
Score:	110.50	Matches:	73
Percent Similarity:	43.0%	Conservative:	55
Best Local Similarity:	24.5%	Mismatches:	112
Query Match:	5.8%	Indels:	59
DB:	9	Gaps:	15

US-10-757-745-2 (1-362) x US-10-450-763-2332 (1-5130)

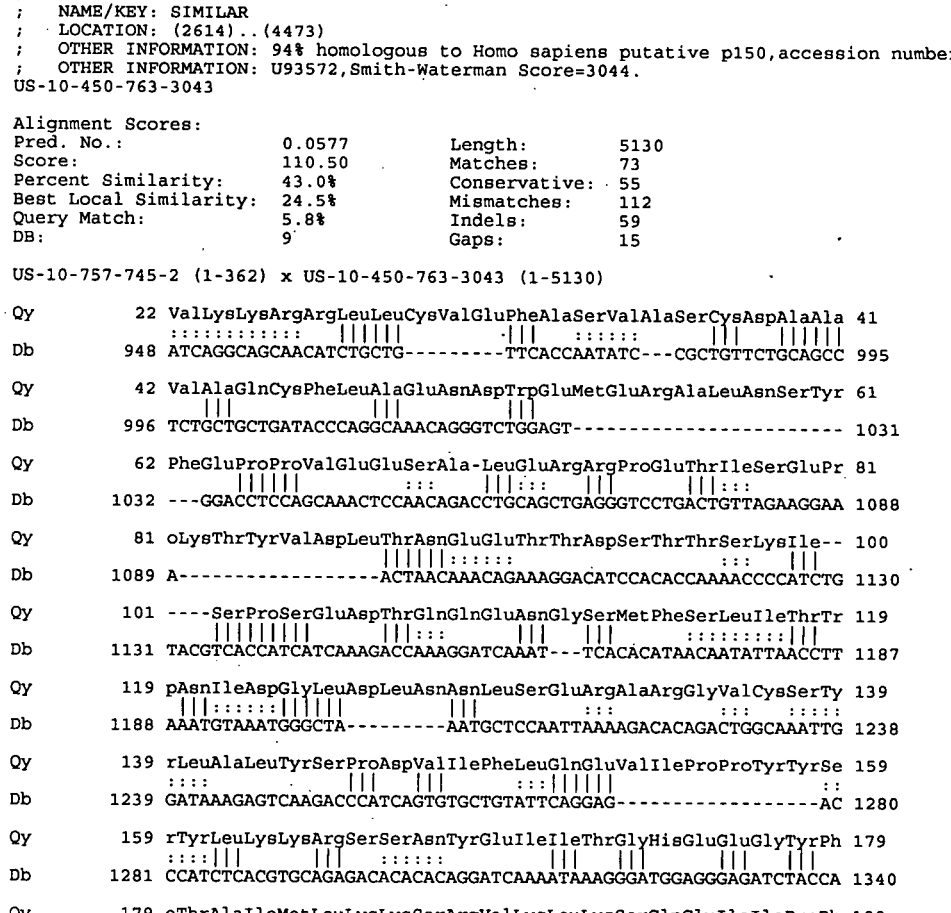


RESULT 65
US-10-450-763-3043
; Sequence 3043, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 3043
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2614)..(4473)
; OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
; OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-3043

Alignment Scores:

Pred. No.:	0.0577	Length:	5130
Score:	110.50	Matches:	73
Percent Similarity:	43.0%	Conservative:	55
Best Local Similarity:	24.5%	Mismatches:	112
Query Match:	5.8%	Indels:	59
DB:	9	Gaps:	15

US-10-757-745-2 (1-362) x US-10-450-763-3043 (1-5130)



```
Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267
Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728
```

RESULT 66
US-10-450-763-3970
; Sequence 3970, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 3970
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2614)..(4473)
; OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
; OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-3970

Alignment Scores:

Pred. No.:	0.0577	Length:	5130
Score:	110.50	Matches:	73
Percent Similarity:	43.0%	Conservative:	55
Best Local Similarity:	24.5%	Mismatches:	112
Query Match:	5.8%	Indels:	59
DB:	9	Gaps:	15

US-10-757-745-2 (1-362) x US-10-450-763-3970 (1-5130)

```
Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT----------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy        199 eProSerThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnVa 215
Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267
Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728
```

RESULT 67
US-10-450-763-6046
; Sequence 6046, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 6046
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2614)..(4473)
; OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
; OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-6046

Alignment Scores:

Pred. No.:	0.0577	Length:	5130
Score:	110.50	Matches:	73
Percent Similarity:	43.0%	Conservative:	55
Best Local Similarity:	24.5%	Mismatches:	112

```
Query Match:            5.8%        Indels:        59
DB:                     9           Gaps:          15

US-10-757-745-2 (1-362) x US-10-450-763-6046 (1-5130)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy        199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215
Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp-----------------------ThrAsnLeuArgA 267
Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728

RESULT 68
US-10-450-763-6593
; Sequence 6593, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 6593
;   LENGTH: 5130
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2614)..(4473)
;   OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
;   OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-6593

Alignment Scores:
Pred. No.:              0.0577      Length:        5130
Score:                  110.50      Matches:       73
Percent Similarity:     43.0%       Conservative:  55
Best Local Similarity:  24.5%       Mismatches:    112
Query Match:            5.8%        Indels:        59
DB:                     9           Gaps:          15

US-10-757-745-2 (1-362) x US-10-450-763-6593 (1-5130)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy        199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215
Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
```

```
Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255

Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267

Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn------------------AsnIleV 281

Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297

Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728
```

RESULT 69
US-10-450-763-8626
; Sequence 8626, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 8626
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2614)..(4473)
; OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
; OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-8626

Alignment Scores:

Pred. No.:	0.0577	Length:	5130
Score:	110.50	Matches:	73
Percent Similarity:	43.0%	Conservative:	55
Best Local Similarity:	24.5%	Mismatches:	112
Query Match:	5.8%	Indels:	59
DB:	9	Gaps:	15

US-10-757-745-2 (1-362) x US-10-450-763-8626 (1-5130)

```
Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41

Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61

Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT----------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81

Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100

Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130
```

```
Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119

Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139

Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159

Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179

Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199

Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy        199 eProSerThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnVa 215

Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235

Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255

Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267

Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn------------------AsnIleV 281

Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297

Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728
```

RESULT 70
US-10-450-763-9629
; Sequence 9629, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 9629
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2614)..(4473)
; OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
; OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-9629

```
Alignment Scores:
Pred. No.:                0.0577          Length:          5130
Score:                    110.50          Matches:         73
Percent Similarity:       43.0%           Conservative:    55
Best Local Similarity:    24.5%           Mismatches:      112
Query Match:              5.8%            Indels:          59
DB:                       9               Gaps:            15

US-10-757-745-2 (1-362) x US-10-450-763-9629 (1-5130)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
              :::::::::::   ||||||         |||   ::::::      |||   ||||||
Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
              |||             |||         |||
Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
              ||||||         :::    |||:::   |||      |||:::
Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
                          ||||||::::::                :::   |||
Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy         101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
               |||||||||      |||:::      |||   |||      :::::::::|||
Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy         119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
               |||::::::||||||          |||         :::         :::   :::::
Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy         139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
               ::::      |||   |||      :::||||||                      ::
Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy         159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
               ::::|||      |||   ::::::          |||   |||       |||   |||
Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy         179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
                        ||||||:::   |||   :::         :::          ||
Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy         199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215
               |      ||||||:::   |||         :::   :::   |||      :::::
Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy         215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
               :         ||||||      :::       :::       :::|||            ||
Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy         235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
               |   |||          :::|||   ||||||:::   :::|||          :::|||
Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy         255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267
               |||:::|||      ||||||                        |||:::   ||||
Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy         267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
               |||||::::::||||||          :::   |||                :::::::
Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy         281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
               ::|||::::::       |||       ||||||     :::|||||:::
Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728

RESULT 71
US-10-450-763-10788
; Sequence 10788, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 10788
;   LENGTH: 5130
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (6928)..(6113)
;   OTHER INFORMATION: 92% homologous to Homo sapiens protein kinase,accession
;   OTHER INFORMATION: number X85545,Smith-Waterman Score=1338.
US-10-450-763-10788

Alignment Scores:
Pred. No.:                0.0577          Length:          5130
Score:                    110.50          Matches:         73
Percent Similarity:       43.0%           Conservative:    55
Best Local Similarity:    24.5%           Mismatches:      112
Query Match:              5.8%            Indels:          59
DB:                       9               Gaps:            15

US-10-757-745-2 (1-362) x US-10-450-763-10788 (1-5130)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
              :::::::::::   ||||||         |||   ::::::      |||   ||||||
Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
              |||             |||         |||
Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
              ||||||         :::    |||:::   |||      |||:::
Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
                          ||||||::::::                :::   |||
Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy         101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
               |||||||||      |||:::      |||   |||      :::::::::|||
Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy         119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
               |||::::::||||||          |||         :::         :::   :::::
Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy         139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
               ::::      |||   |||      :::||||||                      ::
Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy         159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
               ::::|||      |||   ::::::          |||   |||       |||   |||
Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy         179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
                        ||||||:::   |||   :::         :::          ||
Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy         199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215
```

```
Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235

Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255

Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267

Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn------------------AsnIleV 281

Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297

Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728

RESULT 72
US-10-450-763-19864
; Sequence 19864, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 19864
;   LENGTH: 5130
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2614)..(4473)
;   OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
;   OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-19864

Alignment Scores:
Pred. No.:              0.0577          Length:         5130
Score:                  110.50          Matches:        73
Percent Similarity:     43.0%           Conservative:   55
Best Local Similarity:  24.5%           Mismatches:     112
Query Match:            5.8%            Indels:         59
DB:                     9               Gaps:           15

US-10-757-745-2 (1-362) x US-10-450-763-19864 (1-5130)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41

Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61

Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81

Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100

Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119

Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139

Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159

Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179

Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199

Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy        199 eProSerThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnVa 215

Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235

Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255

Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267

Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn------------------AsnIleV 281

Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297

Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728

RESULT 73
US-10-450-763-20770
; Sequence 20770, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 20770
;   LENGTH: 5130
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
```

```
;   LOCATION: (2614)..(4473)
;   OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
;   OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-20770

Alignment Scores:
Pred. No.:              0.0577          Length:         5130
Score:                  110.50          Matches:        73
Percent Similarity:     43.0%           Conservative:   55
Best Local Similarity:  24.5%           Mismatches:     112
Query Match:            5.8%            Indels:         59
DB:                     9               Gaps:           15

US-10-757-745-2 (1-362) x US-10-450-763-20770 (1-5130)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
              :::::::::::::   ||||||          |||   ::::::      |||   ||||||
Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
                 |||             |||         |||
Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
                 ||||||          :::    |||:::   |||       |||:::
Db        1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
                           ||||||::::::                          :::   |||
Db        1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
                 |||||||||       |||:::     |||   |||       :::::::::|||
Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
                |||::::::||||||          |||          :::          :::   :::::
Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
                ::::            |||   |||       :::||||||                ::
Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
                ::::|||     |||   ::::::          |||   |||       |||   |||
Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
                          ||||||:::   |||   :::          :::           ||
Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy        199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215
                |      ||||||:::   |||             :::   :::   |||      :::::
Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
                :           ||||||      :::       :::       :::|||           ||
Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
                |   |||         :::|||   ||||||:::   :::|||          :::||||
Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267
                 |||:::|||       ||||||                          |||:::   ||||
Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
                |||||::::::||||||          :::   |||                  :::::::
Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
                ::|||::::::       |||       ||||||      :::||||||:::
```

```
Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728

RESULT 74
US-10-450-763-20902
; Sequence 20902, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 20902
;   LENGTH: 5130
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2614)..(4473)
;   OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
;   OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-20902

Alignment Scores:
Pred. No.:              0.0577          Length:         5130
Score:                  110.50          Matches:        73
Percent Similarity:     43.0%           Conservative:   55
Best Local Similarity:  24.5%           Mismatches:     112
Query Match:            5.8%            Indels:         59
DB:                     9               Gaps:           15

US-10-757-745-2 (1-362) x US-10-450-763-20902 (1-5130)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
              :::::::::::::   ||||||          |||   ::::::      |||   ||||||
Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
                 |||             |||         |||
Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
                 ||||||          :::    |||:::   |||       |||:::
Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
                           ||||||::::::                          :::   |||
Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
                 |||||||||       |||:::     |||   |||       :::::::::|||
Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
                |||::::::||||||          |||          :::          :::   :::::
Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
                ::::            |||   |||       :::||||||                ::
Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
                ::::|||     |||   ::::::          |||   |||       |||   |||
Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340
```

```
Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy        199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215
Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp-----------------------ThrAsnLeuArgA 267
Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728
```

```
RESULT 75
US-10-450-763-25259
; Sequence 25259, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 25259
;   LENGTH: 5130
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2614)..(4473)
;   OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
;   OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-450-763-25259
```

```
Alignment Scores:
Pred. No.:                0.0577          Length:          5130
Score:                    110.50          Matches:         73
Percent Similarity:       43.0%           Conservative:    55
Best Local Similarity:    24.5%           Mismatches:      112
Query Match:              5.8%            Indels:          59
DB:                       9               Gaps:            15

US-10-757-745-2 (1-362) x US-10-450-763-25259 (1-5130)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
```

```
Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy        199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215
Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp-----------------------ThrAsnLeuArgA 267
Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728
```

```
Search completed: March 11, 2006, 05:51:31
Job time : 2414.27 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 20:06:44 ; Search time 411.677 Seconds
(without alignments)
1563.065 Million cell updates/sec

Title: US-10-757-745-2
Perfect score: 1902
Sequence: 1 MELGSCLEGGREAAEEEGEP.....DCGRFPSDHWGLLCNLDIIL 362

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170331_4111/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_417_@runat_09032006_170331_4111 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1902	100.0	1920	3	US-09-697-863A-1	Sequence 1, Appli
	2	1902	100.0	1948	3	US-09-620-312D-889	Sequence 889, App
	3	1258	66.1	1312	3	US-09-697-863A-3	Sequence 3, Appli
c	4	1177	61.9	1079	3	US-09-118-554-63	Sequence 63, Appl
c	5	1177	61.9	1079	3	US-09-118-627-63	Sequence 63, Appl
c	6	1177	61.9	1079	3	US-09-602-877A-63	Sequence 63, Appl
c	7	221	11.6	268	3	US-09-397-787-312	Sequence 312, App
	8	145	7.6	379	3	US-09-621-976-8403	Sequence 8403, Ap
c	9	114.5	6.0	232547	3	US-09-949-016-16603	Sequence 16603, A
	10	111	5.8	1630	4	US-09-605-703B-1143	Sequence 1143, Ap
	11	111	5.8	1725	4	US-09-605-703B-1141	Sequence 1141, Ap
c	12	109	5.7	106380	3	US-09-949-016-17553	Sequence 17553, A
c	13	108.5	5.7	1830121	3	US-09-557-884-1	Sequence 1, Appli
c	14	108.5	5.7	1830121	3	US-09-643-990A-1	Sequence 1, Appli
c	15	108.5	5.7	1830121	3	US-10-158-865-1	Sequence 1, Appli
	16	108	5.7	3640	3	US-09-976-594-888	Sequence 888, App
	17	108	5.7	4061	3	US-09-620-312D-363	Sequence 363, App
c	18	108	5.7	86439	3	US-09-949-016-11945	Sequence 11945, A
c	19	108	5.7	86440	3	US-09-949-016-16990	Sequence 16990, A
c	20	108	5.7	818128	3	US-09-949-016-14546	Sequence 14546, A
c	21	108	5.7	818128	3	US-09-949-016-14547	Sequence 14547, A
c	22	108	5.7	818128	3	US-09-949-016-14548	Sequence 14548, A
c	23	108	5.7	818128	3	US-09-949-016-14549	Sequence 14549, A
c	24	108	5.7	818128	3	US-09-949-016-14550	Sequence 14550, A
c	25	108	5.7	818128	3	US-09-949-016-14551	Sequence 14551, A
c	26	108	5.7	818128	3	US-09-949-016-14552	Sequence 14552, A
c	27	108	5.7	818128	3	US-09-949-016-14553	Sequence 14553, A
c	28	108	5.7	818128	3	US-09-949-016-14554	Sequence 14554, A
c	29	108	5.7	818128	3	US-09-949-016-14555	Sequence 14555, A
c	30	108	5.7	818128	3	US-09-949-016-14556	Sequence 14556, A
c	31	108	5.7	818128	3	US-09-949-016-14557	Sequence 14557, A
c	32	108	5.7	818128	3	US-09-949-016-14558	Sequence 14558, A
c	33	108	5.7	818128	3	US-09-949-016-14559	Sequence 14559, A
c	34	108	5.7	818128	3	US-09-949-016-14560	Sequence 14560, A
c	35	108	5.7	818128	3	US-09-949-016-14561	Sequence 14561, A
c	36	108	5.7	818128	3	US-09-949-016-14562	Sequence 14562, A
c	37	108	5.7	818128	3	US-09-949-016-14564	Sequence 14564, A
c	38	108	5.7	818128	3	US-09-949-016-14565	Sequence 14565, A
c	39	108	5.7	818128	3	US-09-949-016-14566	Sequence 14566, A
c	40	108	5.7	818128	3	US-09-949-016-14567	Sequence 14567, A
c	41	106	5.6	32042	3	US-09-245-281-44	Sequence 44, Appl
c	42	106	5.6	32042	3	US-09-340-620A-63	Sequence 63, Appl
c	43	106	5.6	32042	3	US-09-728-721-63	Sequence 63, Appl
c	44	106	5.6	39553	3	US-09-949-002-810	Sequence 810, App
c	45	106	5.6	58111	3	US-09-949-002-673	Sequence 673, App
c	46	106	5.6	102520	3	US-09-949-016-17367	Sequence 17367, A
c	47	106	5.6	102526	3	US-09-949-016-12448	Sequence 12448, A
c	48	106	5.6	175236	3	US-09-949-016-14353	Sequence 14353, A
c	49	106	5.6	294836	3	US-09-949-016-15974	Sequence 15974, A
	50	106	5.6	331814	3	US-09-949-016-12008	Sequence 12008, A
	51	106	5.6	331814	3	US-09-949-016-17056	Sequence 17056, A
c	52	105	5.5	206433	3	US-09-949-016-13527	Sequence 13527, A
c	53	105	5.5	254778	3	US-09-949-016-12417	Sequence 12417, A
c	54	105	5.5	421491	3	US-09-949-016-12805	Sequence 12805, A
c	55	105	5.5	421494	3	US-09-949-016-14060	Sequence 14060, A
c	56	104.5	5.5	61462	3	US-09-949-016-17522	Sequence 17522, A
c	57	104	5.5	131631	3	US-09-949-016-11757	Sequence 11757, A
	58	103.5	5.4	1627	3	US-09-485-473-4	Sequence 4, Appli
	59	103	5.4	152655	3	US-09-949-016-16246	Sequence 16246, A
	60	102	5.4	26711	3	US-09-949-016-14214	Sequence 14214, A
	61	102	5.4	35837	3	US-09-949-016-15232	Sequence 15232, A
	62	102	5.4	168394	3	US-09-949-016-13002	Sequence 13002, A
	63	101.5	5.3	72843	3	US-09-949-016-12574	Sequence 12574, A
	64	101	5.3	179904	3	US-09-949-002-577	Sequence 577, App
	65	101	5.3	179905	3	US-09-949-002-705	Sequence 705, App
	66	100	5.3	879	3	US-09-270-767-11336	Sequence 11336, A
	67	100	5.3	5238	3	US-09-080-855-1	Sequence 1, Appli
	68	100	5.3	5238	3	US-09-566-076-1	Sequence 1, Appli
	69	100	5.3	140725	3	US-09-949-016-17074	Sequence 17074, A
c	70	100	5.3	192700	3	US-09-949-016-11820	Sequence 11820, A
c	71	100	5.3	192704	3	US-09-949-016-17182	Sequence 17182, A
c	72	100	5.3	212139	3	US-09-949-016-16065	Sequence 16065, A
	73	99.5	5.2	840	3	US-09-252-991A-14967	Sequence 14967, A
c	74	99.5	5.2	1149	3	US-09-252-991A-14783	Sequence 14783, A
	75	99.5	5.2	1461	3	US-09-252-991A-14833	Sequence 14833, A

ALIGNMENTS

RESULT 1
US-09-697-863A-1
; Sequence 1, Application US/09697863A

; Patent No. 6812203
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/09/697,863A
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1627)..(1627)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1108)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-09-697-863A-1

Alignment Scores:

Pred. No.:	5.07e-223	Length:	1920
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2 (1-362) x US-09-697-863A-1 (1-1920)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 79

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 139

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 199

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 259

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 319

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 379

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1099

Qy        361 IleLeu 362
              ||||||
Db       1100 ATATTG 1105
```

RESULT 2
US-09-620-312D-889
; Sequence 889, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25

```
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-09-620-312D-889

Alignment Scores:
Pred. No.:                5.19e-223       Length:          1948
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2 (1-362) x US-09-620-312D-889 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124

RESULT 3
US-09-697-863A-3
; Sequence 3, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/09/697,863A
;  CURRENT FILING DATE:  2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (122)..(1234)
;   OTHER INFORMATION:
US-09-697-863A-3

Alignment Scores:
Pred. No.:                3.43e-144       Length:          1312
Score:                    1258.00         Matches:         240
Percent Similarity:       80.6%           Conservative:    43
Best Local Similarity:    68.4%           Mismatches:      64
Query Match:              66.1%           Indels:          4
DB:                       3               Gaps:            3

US-10-757-745-2 (1-362) x US-09-697-863A-3 (1-1312)

Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31
              |||||||||:::   |||   |||       ||||||:::|||||||||   |||:::
Db        191 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 247

Qy         32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51
              ||||||   |||       ||||||       :::             |||   |||||||||
Db        248 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 307

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
              |||:::   ::::::||||||::::::|||||||||   |||       :::
Db        308 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 367
```

```
Qy          72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91
Db         368 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 427

Qy          92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111
Db         428 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 484

Qy         112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131
Db         485 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 544

Qy         132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
Db         545 AGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 604

Qy         152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
Db         605 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 664

Qy         172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191
Db         665 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTT 724

Qy         192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211
Db         725 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 784

Qy         212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
Db         785 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 844

Qy         232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
Db         845 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 904

Qy         252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271
Db         905 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 964

Qy         272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
Db         965 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 1024

Qy         292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
Db        1025 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1084

Qy         312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
Db        1085 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1138

Qy         332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
Db        1139 CAAAGTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1198

Qy         352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
Db        1199 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1231

RESULT 4
US-09-118-554-63/c
; Sequence 63, Application US/09118554A
; Patent No. 6365348
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND
;  TITLE OF INVENTION: METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.450C1
;  CURRENT APPLICATION NUMBER: US/09/118,554A
;  CURRENT FILING DATE: 1998-07-17
;  EARLIER APPLICATION NUMBER: 08/988,255
;  EARLIER FILING DATE: 1997-12-24
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-118-554-63

Alignment Scores:
Pred. No.:                2.22e-134       Length:          1079
Score:                    1177.00         Matches:         226
Percent Similarity:       99.6%           Conservative:    0
Best Local Similarity:    99.6%           Mismatches:      1
Query Match:              61.9%           Indels:          1
DB:                       3               Gaps:            0

US-10-757-745-2 (1-362) x US-09-118-554-63 (1-1079)

Qy         136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
Db        1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
Db        1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
Db         957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
Db         897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
Db         837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy         236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
Db         777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy         256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
Db         717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy         276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
Db         657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy         296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
Db         597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy         316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
Db         537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy         336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
Db         478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy         356 CysAsnLeuAspIleIleLeu 362
Db         418 TGCAACTTAGATATAATATTG 398

RESULT 5
US-09-118-627-63/c
; Sequence 63, Application US/09118627A
; Patent No. 6379951
; GENERAL INFORMATION:
```

```
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER
;  TITLE OF INVENTION: AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.446C1
;  CURRENT APPLICATION NUMBER: US/09/118,627A
;  CURRENT FILING DATE: 1998-07-17
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-118-627-63

Alignment Scores:
Pred. No.:                2.22e-134      Length:          1079
Score:                    1177.00        Matches:         226
Percent Similarity:       99.6%          Conservative:    0
Best Local Similarity:    99.6%          Mismatches:      1
Query Match:              61.9%          Indels:          1
DB:                       3              Gaps:            0

US-10-757-745-2 (1-362) x US-09-118-627-63 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398
```

```
RESULT 6
US-09-602-877A-63/c
; Sequence 63, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE:  210121.446C5
;  CURRENT APPLICATION NUMBER: US/09/602,877A
;  CURRENT FILING DATE:  2000-06-22
;  NUMBER OF SEQ ID NOS:  107
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-602-877A-63

Alignment Scores:
Pred. No.:                2.22e-134      Length:          1079
Score:                    1177.00        Matches:         226
Percent Similarity:       99.6%          Conservative:    0
Best Local Similarity:    99.6%          Mismatches:      1
Query Match:              61.9%          Indels:          1
DB:                       3              Gaps:            0

US-10-757-745-2 (1-362) x US-09-602-877A-63 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398
```

RESULT 7
US-09-397-787-312/c
; Sequence 312, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-312

Alignment Scores:

Pred. No.:	6.13e-18	Length:	268
Score:	221.00	Matches:	41
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	11.6%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2 (1-362) x US-09-397-787-312 (1-268)

```
Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 341
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAA 208

Qy        342 LeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 CTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATA 148

Qy        362 Leu 362
              |||
Db        147 TTG 145
```

RESULT 8
US-09-621-976-8403
; Sequence 8403, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8403
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8403

Alignment Scores:

Pred. No.:	2.32e-08	Length:	379
Score:	145.00	Matches:	28
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	96.6%	Mismatches:	0
Query Match:	7.6%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2 (1-362) x US-09-621-976-8403 (1-379)

```
Qy         55 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 74
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 78

Qy         75 ProGluThrIleSerGluProLysThr 83
              |||||||||||||||||||||||||||
Db         79 CCTGAAACCATCTCTGAGCCCAAGACC 105
```

RESULT 9
US-09-949-016-16603/c
; Sequence 16603, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16603
; LENGTH: 232547
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16603

Alignment Scores:

Pred. No.:	4.11	Length:	232547
Score:	114.50	Matches:	58
Percent Similarity:	45.4%	Conservative:	45
Best Local Similarity:	25.6%	Mismatches:	87
Query Match:	6.0%	Indels:	37
DB:	3	Gaps:	10

US-10-757-745-2 (1-362) x US-09-949-016-16603 (1-232547)

```
Qy          91 GluThrThrAspSerThr-ThrSerLysIleSerProSerGluAspThrGlnGln-GluA 110
               ||| ||| |||||| ||||||||||||:::                   ::: |
Db       72006 GAAGAAACTCCATCAACTAACGAGCAAAATAACCAGCCAACATCATAATGACAGGATCAA 71947

Qy         110 snGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuS 130
               ||  |||     :::::::::|||  |||:::::|||      ||||||:::
Db       71946 AT---TCACACATAACAATATTAACCTTAAATGTAAATGGA------CTAAATTCTCCAA 71896

Qy         130 erGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheL 150
               :::|||  |||  :::  :::::::::           |||  |||     :
Db       71895 TTAAAAGAAACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTA 71839

Qy         150 euGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluI 170
               ::||||||                 ::::::|||      |||   ::::::
Db       71838 TTCAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCA 71797

Qy         170 leIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValL 190
               |||  |||      :::  |||                ||||||:::  |||
Db       71796 AAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTG 71737

Qy         190 ysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuC 210
               :::        :::         |||      ||||||:::   |||:::      |
Db       71736 CAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAT 71677
```

```
Qy        210 ysValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerH 226
Db      71676 GCCATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAATTATCCTAAATATAT 71617

Qy        226 isLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuL 246
Db      71616 ATGCACCCAATACA---------GGAGCACCCAGATTG------ATAAAGCAAGTCCTGA 71572

Qy        246 ysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu- 265
Db      71571 GTGACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCC 71515

Qy        266 ----------------------------------------ArgAspArgGluValThrA 272
Db      71514 CACTGTCAACATTAGACAGATCAATGAGACAGAAAGTTAATAAGGATACCCAGGAATTGA 71455

Qy        272 rgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysH 292
Db      71454 ACTCAGCTCTGAACCAAGCAGACCTAATAGATATCTACAGAACTCTC---CACCCCAAAT 71398

Qy        292 isCysGlnTyrThrTrp 297
Db      71397 CAACAGAATATACATTT 71381
```

RESULT 10
US-09-605-703B-1143
; Sequence 1143, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1143
; LENGTH: 1630
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1630)
; OTHER INFORMATION: FRXA02398
US-09-605-703B-1143

Alignment Scores:

Pred. No.:	0.00367	Length:	1630
Score:	111.00	Matches:	99
Percent Similarity:	36.2%	Conservative:	55
Best Local Similarity:	23.2%	Mismatches:	135
Query Match:	5.8%	Indels:	138
DB:	4	Gaps:	24

US-10-757-745-2 (1-362) x US-09-605-703B-1143 (1-1630)

```
Qy         21 GluValLysLysArgArgLeuLeuCysValGluPhe---AlaSerValAlaSerCysAsp 39
Db        107 GAAGTGAAGAAGCGTAATCTCCTCGTAGCTCCCCTCACTGCTTCCCTGGTGTTCTGTAAC 166

Qy         40 AlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsn 59
Db        167 TTGGCTGTT--------------GCAGCAAACGCCGTTGAAGTTGAGGCC--------- 202

Qy         60 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 79
Db        203 ---------GAATCACCAGTTGTCATCAATGAAGTTGAATCCAACAGCGACCCAGTTGGT 253

Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr------ 97
Db        254 GAC---------TGGGTGGAGTTGGCTAACACCGACAACAACAACTCCATCGACATTTCC 304

Qy         98 ------SerLysIleSerProSerGluAsp---------------------------- 105
Db        305 GGTTGGTCCTTAGTCGATGACAAGGAAGACCTGGAAAATGCCCTCGTCCTTCCTGAAGGC 364

Qy        106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
Db        365 ACTGAGATTGAGTCCGGTGGATACTTTGTTATCTACACCGACTCTGCTGATTACGTGCCT 424

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
Db        425 ACCAACAAC------------ACCTTTGGTGGCCAGGAATACTTCGGCCTCGGCAAAGAT 472

Qy        146 AspValIlePheLeuGln------------GluValIleProProTyr---TyrSerTyr 160
Db        473 GACACTGTTACTCTGCGCAACGCTGAAGGCGAAGTAGTTGCTACCTATTCCTGGAAGGAT 532

Qy        161 LeuLysLysArgSerSerAsn---TyrGluIleIle---------------------- 171
Db        533 CTGGGCGAGCACGCAGAAAACACCTATGGTCGCATCCCAGATATGACTGGTGATTTCGCA 592

Qy        172 ------------------------------------------------ThrGlyHisGlu 175
Db        593 AACACCGGCGTTCCAACCCCAGGTGCAAAGAATGTTGCTGCTGAAGGCTCCGGCGAAGAA 652

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
Db        653 GAAGGC--------------GTTGTTGCAAACGCCCAGCTTCCATTCCACAACGTTGAA 697

Qy        196 IleIlePro-PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVa 215
Db        698 ATCACCCCAATTCACCTCG------------------------------GTGGAGAT 724

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
Db        725 TTCACCGGTGAAGATATGTCCGGCGTTGATTTCGATGCAA---ACAGCACCGCATGGATC 781

Qy        235 aGluArgMetAsn---------------------------------------------- 239
Db        782 GCCAACAATGACATTGGAAAGATCTACTCCCTCGCCCACGACATAGCTAACAACACCTAC 841

Qy        240 -------------GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
Db        842 AAGCTGACTGGCGAATGGGAAACCGGCTACCCAGAAGGCGGCGGAGAG-CCAGACGCTGA 900

Qy        255 aThrValIlePheAla------GlyAsp-----------------ThrAsnLeuArgAs 267
Db        901 AGGCATCGTCGCAGCTACCAACGGTGACATCTACCTGTCCACCGAGCGCAACAACGCTGA 960

Qy        267 pArgGluValThrArgCysGlyGlyLeu----------------------------Pr 277
Db        961 CAAGAACGTCTCTCGCCCATCCATCCTGCGTTTTGCTACCCCAACTGGCAAGACTGGCGT 1020

Qy        277 oAsnAsnIleValAspValTrp---------GluPheLeuGlyLysProLysHisCysGl 294
Db       1021 ACAAAACGCAGTTCAGGAATGGGACCTGTCTGAGTTCGTCGGC---------------- 1063

Qy        294 nTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuAr 314
Db       1064 ----------GACATTCAGCCCAATGGT------GGTCTTGAGGCAATCGCGCAGCTCGA 1107

Qy        314 gPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLe 334
Db       1108 G---GACAACATCTTCGTTGTCGGTGTCGAAGAGACAGGTGATGTCATC------GTTGT 1158

Qy        334 uAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLe 354
```

```
Db       1159 TGATCTTTCCGCTGACCAGCCAGTTCTGGTTCAAAGGTACGAATCTTCCTTCGACGGTGT 1218

Qy        354 uLeuCysAsnLeuAsp 359
Db       1219 CATG---TCGCTTGAT 1231
```

RESULT 11
US-09-605-703B-1141
; Sequence 1141, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1141
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1702)
; OTHER INFORMATION: RXN02398
US-09-605-703B-1141

Alignment Scores:

Pred. No.:	0.00402	Length:	1725
Score:	111.00	Matches:	99
Percent Similarity:	36.2%	Conservative:	55
Best Local Similarity:	23.2%	Mismatches:	135
Query Match:	5.8%	Indels:	138
DB:	4	Gaps:	24

US-10-757-745-2 (1-362) x US-09-605-703B-1141 (1-1725)

```
Qy         21 GluValLysLysArgArgLeuLeuCysValGluPhe---AlaSerValAlaSerCysAsp 39
Db        107 GAAGTGAAGAAGCGTAATCTCCTCGTAGCTCCCCTCACTGCTTCCCTGGTGTTCTGTAAC 166

Qy         40 AlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsn 59
Db        167 TTGGCTGTT--------------GCAGCAAACGCCGTTGAAGTTGAGGCC--------- 202

Qy         60 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 79
Db        203 ---------GAATCACCAGTTGTCATCAATGAAGTTGAATCCAACAGCGACCCAGTTGGT 253

Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr------ 97
Db        254 GAC---------TGGGTGGAGTTGGCTAACACCGACAACAACAACTCCATCGACATTTCC 304

Qy         98 ------SerLysIleSerProSerGluAsp---------------------------- 105
Db        305 GGTTGGTCCTTAGTCGATGACAAGGAAGACCTGGAAAATGCCCTCGTCCTTCCTGAAGGC 364

Qy        106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
Db        365 ACTGAGATTGAGTCCGGTGGATACTTTGTTATCTACACCGACTCTGCTGATTACGTGCCT 424

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
```

```
Db        425 ACCAACAAC------------ACCTTTGGTGGCCAGGAATACTTCGGCCTCGGCAAAGAT 472

Qy        146 AspValIlePheLeuGln------------GluValIleProProTyr---TyrSerTyr 160
Db        473 GACACTGTTACTCTGCGCAACGCTGAAGGCGAAGTAGTTGCTACCTATTCCTGGAAGGAT 532

Qy        161 LeuLysLysArgSerSerAsn---TyrGluIleIle------------------------ 171
Db        533 CTGGGCGAGCACGCAGAAAACACCTATGGTCGCATCCCAGATATGACTGGTGATTTCGCA 592

Qy        172 ----------------------------------------------------ThrGlyHisGlu 175
Db        593 AACACCGGCGTTCCAACCCCAGGTGCAAAGAATGTTGCTGCTGAAGGCTCCGGCGAAGAA 652

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
Db        653 GAAGGC--------------GTTGTTGCAAACGCCCAGCTTCCATTCCACAACGTTGAA 697

Qy        196 IleIlePro-PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVa 215
Db        698 ATCACCCCAATTCACCTCG------------------------------GTGGAGAT 724

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
Db        725 TTCACCGGTGAAGATATGTCCGGCGTTGATTTCGATGCAA---ACAGCACCGCATGGATC 781

Qy        235 aGluArgMetAsn---------------------------------------------- 239
Db        782 GCCAACAATGACATTGGAAAGATCTACTCCCTCGCCCACGACATAGCTAACAACACCTAC 841

Qy        240 -------------GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
Db        842 AAGCTGACTGGCGAATGGGAAACCGGCTACCCAGAAGGCGGCGGAGAG-CCAGACGCTGA 900

Qy        255 aThrValIlePheAla------GlyAsp------------------ThrAsnLeuArgAs 267
Db        901 AGGCATCGTCGCAGCTACCAACGGTGACATCTACCTGTCCACCGAGCGCAACAACGCTGA 960

Qy        267 pArgGluValThrArgCysGlyGlyLeu------------------------------Pr 277
Db        961 CAAGAACGTCTCTCGCCCATCCATCCTGCGTTTTGCTACCCCAACTGGCAAGACTGGCGT 1020

Qy        277 oAsnAsnIleValAspValTrp---------GluPheLeuGlyLysProLysHisCysGl 294
Db       1021 ACAAAACGCAGTTCAGGAATGGGACCTGTCTGAGTTCGTCGGC----------------- 1063

Qy        294 nTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuAr 314
Db       1064 ----------GACATTCAGCCCAATGGT------GGTCTTGAGGCAATCGCGCAGCTCGA 1107

Qy        314 gPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLe 334
Db       1108 G---GACAACATCTTCGTTGTCGGTGTCGAAGAGACAGGTGATGTCATC------GTTGT 1158

Qy        334 uAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLe 354
Db       1159 TGATCTTTCCGCTGACCAGCCAGTTCTGGTTCAAAGGTACGAATCTTCCTTCGACGGTGT 1218

Qy        354 uLeuCysAsnLeuAsp 359
Db       1219 CATG---TCGCTTGAT 1231
```

RESULT 12
US-09-949-016-17553/c
; Sequence 17553, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

```
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17553
;   LENGTH: 106380
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17553

Alignment Scores:
Pred. No.:                5.5          Length:          106380
Score:                    109.00       Matches:         56
Percent Similarity:       44.0%        Conservative:    43
Best Local Similarity:    24.9%        Mismatches:      92
Query Match:              5.7%         Indels:          35
DB:                       3            Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-17553 (1-106380)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db      22502 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 22444

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db      22443 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 22390

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      22389 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 22333

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db      22332 CAGGAA----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 22291

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db      22290 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATAGAAAACAAAAAAAGGCAGGGGTTGCA 22231

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db      22230 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 22171

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
Db      22170 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 22111

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db      22110 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 22066

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db      22065 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 22009

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db      22008 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGACACCCAGGAATTGAAC 21949

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db      21948 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAATTCTC---CACCCCAAATCA 21892

Qy        293 CysGlnTyrThrTrp 297
Db      21891 ACAGAATATACATTT 21877

RESULT 13

US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
;  GENERAL INFORMATION:
;       APPLICANT: Fleischmann et al.
;       TITLE OF INVENTION: The Nucleotide sequence of
;                           the Haemophilus influenzae Rd Genome, Fragments
;                           Thereof, and Uses Thereof
;       NUMBER OF SEQUENCES: 1
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: MD
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3 1/2 inch diskette
;            COMPUTER: Dell Pentium
;            OPERATING SYSTEM: MS DOS v6.22
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,884
;            FILING DATE: 25-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/476,102
;            FILING DATE: JUN-5-1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Michelle S. Marks
;            REGISTRATION NUMBER: 41,971
;            REFERENCE/DOCKET NUMBER: PB186P3
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 301-309-8504
;            TELEFAX: 301-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1830121 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Alignment Scores:
Pred. No.:                627          Length:          1830121
Score:                    108.50       Matches:         60
Percent Similarity:       37.2%        Conservative:    46
Best Local Similarity:    21.1%        Mismatches:      112
Query Match:              5.7%         Indels:          67
DB:                       3            Gaps:            13

US-10-757-745-2 (1-362) x US-09-557-884-1 (1-1830121)

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db      43881 GGAACAAAAATGAAATTTATTTCTTTCAATATAAATGGG--------------TTACGT 43837

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      43836 GCTCGCCCACATCAACTTGAAGCCATTATTGAAAAATATCAGCCAGATGTGATTGGCTTA 43777

Qy        151 GlnGlu------ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 168
Db      43776 CAGGAAATTAAAGTTGCTGACGAGGCTTTTCCCTATGAAATTACAGAAAACTTAGGTTAT 43717

Qy        169 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 188
Db      43716 CACGTATTTCATCACGGTCAGAAAGGTCATTATGGCGTGGCATTATTGACTAAACAAGAA 43657

Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
Db      43656 CCTAAAGTG---ATTCGTCGTGGTTTTCCAACTGACAATGAAGATGCACAAAAACGCATT 43600
```

```
Qy        209 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr------------ 224
              :::       :::                |||     |||:::|||
Db      43599 ATTATGGCTGATCTGGAA---------ACTGAATTTGGCTTATTGACTGTCATTAATGGC 43549

Qy        225 ------------------SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
                               |||             |||:::      |||   |||:::
Db      43548 TATTTCCCACAGGGGGAAAGCCGTGCGCACGAGACTAAATTTCCAGCTAAAGAGAAATTT 43489

Qy        239 ---AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
                 |||:::      |||:::|||      |||      :::|||      :::
Db      43488 TATGCAGATCTTCAACAGTATTTAGAAAAA------GAACACGACAAATCGAATCCGATT 43435

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr------------------ 271
              :::     ||||||   |||:::       ::::::
Db      43434 TTAATTATGGGCGATATGAATATTAGCCCAAGTGATTTAGATATTGGTATCGGAGATGAA 43375

Qy        272 --------------------------ArgCysGlyGlyLeuProAsn------------ 278
                                       :::|||      ||||||
Db      43374 AATCGTAAACGTTGGTTACGCACAGGGAAATGTTCTTTCTTACCGGAAGAACGAGCGTGG 43315

Qy        279 --------------------AsnIleValAspValTrpGluPheLeuGlyLysProLys 291
                                    :::   |||   :::      |||
Db      43314 TATCAACGTTTGTATGATTACGGATTAGAAGACAGTTTTCGTAAGCTGAACCCAACAGCA 43255

Qy        292 HisCysGlnTyrThrTrp------------AspThrGlnMetAsnSerAsnLeuGlyIle 307
              :::   :::::::::|||                :::   |||   |||
Db      43254 AATGATAAATTCTCGTGGTTTGATTACCGTTCTAAAGGTTTTGATGATAATCGAGGT--- 43198

Qy        308 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe---ArgAlaAlaAlaGluGlu 326
                  ||||||   |||   |||           :::      ||||||
Db      43197 --------------TTACGAATTGATCATATTTTGGTAAGCCAAAAATTAGCTGAACGA 43153

Qy        327 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 346
                  :::      :::|||||:::      :::|||||
Db      43152 TGTGTCGATGTGGGCATTGCATTAGATATTCGAGCAATGGAAAAA--------------- 43108

Qy        347 PheProSerAspHis 351
                |||||||||||
Db      43107 ---CCGTCTGATCAC 43096
```

```
RESULT 14
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
;   GENERAL INFORMATION:
;        APPLICANT: Robert D. Fleischmann
;                   Mark D. Adams
;                   Owen White
;                   Hamilton O. Smith
;                   J. Craig Venter
;        TITLE OF INVENTION: The Nucleotide sequence of
;                            the Haemophilus influenzae Rd Genome, Fragments
;                            Thereof, and Uses Thereof
;        NUMBER OF SEQUENCES: 1
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville,
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3 1/2 inch diskette
;             COMPUTER: Dell Pentium
;             OPERATING SYSTEM: MS DOS v6.22
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/643,990A
;             FILING DATE: 23-Aug-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/487,429
;             FILING DATE: 1995-06-07
;             APPLICATION NUMBER: 08/426,787
;             FILING DATE: 1995-04-21
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB186P1C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-610-5790
;             TELEFAX: 310-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1830121 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Alignment Scores:
Pred. No.:                627              Length:          1830121
Score:                    108.50           Matches:         60
Percent Similarity:       37.2%            Conservative:    46
Best Local Similarity:    21.1%            Mismatches:      112
Query Match:              5.7%             Indels:          67
DB:                       3                Gaps:            13

US-10-757-745-2 (1-362) x US-09-643-990A-1 (1-1830121)

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||:::              |||:::::::||||||:::|||              |||
Db      43881 GGAACAAAAATGAAATTTATTTCTTTCAATATAAATGGG--------------TTACGT 43837

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                |||         :::   :::   :::      |||   ||||||||||||   |||
Db      43836 GCTCGCCCACATCAACTTGAAGCCATTATTGAAAAATATCAGCCAGATGTGATTGGCTTA 43777

Qy        151 GlnGlu------ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 168
              |||||||      |||             :::   |||             :::      |||
Db      43776 CAGGAAATTAAAGTTGCTGACGAGGCTTTTCCCTATGAAATTACAGAAAACTTAGGTTAT 43717

Qy        169 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 188
                :::                 :::::::|||:::::::          :::|||   |||
Db      43716 CACGTATTTCATCACGGTCAGAAAGGTCATTATGGCGTGGCATTATTGACTAAACAAGAA 43657

Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
                |||:::      :::         |||      :::         :::   :::
Db      43656 CCTAAAGTG---ATTCGTCGTGGTTTTCCAACTGACAATGAAGATGCACAAAAACGCATT 43600

Qy        209 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr------------ 224
              :::       :::                |||     |||:::|||
Db      43599 ATTATGGCTGATCTGGAA---------ACTGAATTTGGCTTATTGACTGTCATTAATGGC 43549

Qy        225 ------------------SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
                               |||             |||:::      |||   |||:::
Db      43548 TATTTCCCACAGGGGGAAAGCCGTGCGCACGAGACTAAATTTCCAGCTAAAGAGAAATTT 43489

Qy        239 ---AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
                 |||:::      |||:::|||      |||      :::|||      :::
Db      43488 TATGCAGATCTTCAACAGTATTTAGAAAAA------GAACACGACAAATCGAATCCGATT 43435

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr------------------ 271
              :::     ||||||   |||:::       ::::::
Db      43434 TTAATTATGGGCGATATGAATATTAGCCCAAGTGATTTAGATATTGGTATCGGAGATGAA 43375

Qy        272 --------------------------ArgCysGlyGlyLeuProAsn------------ 278
                                       :::|||      ||||||
Db      43374 AATCGTAAACGTTGGTTACGCACAGGGAAATGTTCTTTCTTACCGGAAGAACGAGCGTGG 43315

Qy        279 --------------------AsnIleValAspValTrpGluPheLeuGlyLysProLys 291
                                    :::   |||   :::      |||
Db      43314 TATCAACGTTTGTATGATTACGGATTAGAAGACAGTTTTCGTAAGCTGAACCCAACAGCA 43255
```

```
Qy        292 HisCysGlnTyrThrTrp-----------AspThrGlnMetAsnSerAsnLeuGlyIle 307
              :::   :::::::::|||                          :::   |||   |||
Db      43254 AATGATAAATTCTCGTGGTTTGATTACCGTTCTAAAGGTTTTGATGATAATCGAGGT--- 43198

Qy        308 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe---ArgAlaAlaAlaGluGlu 326
                          ||||||   |||  |||       :::      ||||||
Db      43197 --------------TTACGAATTGATCATATTTTGGTAAGCCAAAAATTAGCTGAACGA 43153

Qy        327 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 346
                      :::      :::||||||:::      :::||||||
Db      43152 TGTGTCGATGTGGGCATTGCATTAGATATTCGAGCAATGGAAAAA-------------- 43108

Qy        347 PheProSerAspHis 351
                ||||||||||||
Db      43107 ---CCGTCTGATCAC 43096

RESULT 15
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
;  TITLE OF INVENTION:  Thereof, and Uses Thereof
;  FILE REFERENCE: PB186P2C1D1
;  CURRENT APPLICATION NUMBER: US/10/158,865
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US 09/557,884
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US 08/476,102
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/426,787
;  PRIOR FILING DATE: 1995-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1830121
;   TYPE: DNA
;   ORGANISM: Haemophilus influenzae
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4747)..(4747)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9921)..(9921)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10150)..(10150)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (29298)..(29298)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36543)..(36543)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36551)..(36551)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36636)..(36636)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (40808)..(40810)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44416)..(44416)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44905)..(44905)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44975)..(44975)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45593)..(45593)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45732)..(45732)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (47036)..(47036)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51334)..(51334)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51602)..(51602)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51786)..(51786)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51805)..(51805)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (55369)..(55369)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65309)..(65309)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65313)..(65313)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (80024)..(80024)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (100091)..(100091)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (102696)..(102696)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (105121)..(105121)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (107248)..(107248)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (117136)..(117136)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (119750)..(119750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (119924)..(119924)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (120038)..(120038)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (121344)..(121344)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (122167)..(122167)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (122336)..(122336)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131340)..(131340)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131360)..(131360)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (139910)..(139910)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (140398)..(140398)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (142750)..(142750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145058)..(145058)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145171)..(145171)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145942)..(145942)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (147197)..(147197)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (150841)..(150841)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152500)..(152500)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
Alignment Scores:
Pred. No.:                627             Length:           1830121
Score:                    108.50          Matches:          60
Percent Similarity:       37.2%           Conservative:     46
Best Local Similarity:    21.1%           Mismatches:       112
Query Match:              5.7%            Indels:           67
DB:                       3               Gaps:             13

US-10-757-745-2 (1-362) x US-10-158-865-1 (1-1830121)

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||:::           |||::::::|||||::|||               |||
Db      43881 GGAACAAAAATGAAATTTATTTCTTTCAATATAAATGGG--------------TTACGT 43837

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                |||       :::   :::   :::      |||  ||||||||||||  |||
Db      43836 GCTCGCCCACATCAACTTGAAGCCATTATTGAAAAATATCAGCCAGATGTGATTGGCTTA 43777

Qy        151 GlnGlu------ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 168
              ||||||      |||          :::     |||          :::      |||
Db      43776 CAGGAAATTAAAGTTGCTGACGAGGCTTTTCCCTATGAAATTACAGAAAACTTAGGTTAT 43717

Qy        169 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 188
                :::           ::::::|||::::::          :::|||   |||
Db      43716 CACGTATTTCATCACGGTCAGAAAGGTCATTATGGCGTGGCATTATTGACTAAACAAGAA 43657

Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
                |||:::     :::          |||     :::          :::  :::
Db      43656 CCTAAAGTG---ATTCGTCGTGGTTTTCCAACTGACAATGAAGATGCACAAAAACGCATT 43600

Qy        209 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr---------- 224
              :::       :::                |||     |||:::|||
Db      43599 ATTATGGCTGATCTGGAA---------ACTGAATTTGGCTTATTGACTGTCATTAATGGC 43549

Qy        225 -----------------SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
                          |||             |||:::     |||   |||:::
Db      43548 TATTTCCCACAGGGGGAAAGCCGTGCGCACGAGACTAAATTTCCAGCTAAAGAGAAATTT 43489

Qy        239 ---AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
                  |||:::     |||:::|||     |||       :::|||      :::
Db      43488 TATGCAGATCTTCAACAGTATTTAGAAAAA------GAACACGACAAATCGAATCCGATT 43435

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr------------------ 271
              :::      ||||||   |||:::         ::::::
Db      43434 TTAATTATGGGCGATATGAATATTAGCCCAAGTGATTTAGATATTGGTATCGGAGATGAA 43375

Qy        272 --------------------------ArgCysGlyGlyLeuProAsn------------ 278
                                       :::|||     ||||||
Db      43374 AATCGTAAACGTTGGTTACGCACAGGGAAATGTTCTTTCTTACCGGAAGAACGAGCGTGG 43315

Qy        279 --------------------AsnIleValAspValTrpGluPheLeuGlyLysProLys 291
                                   :::   |||   :::      |||
Db      43314 TATCAACGTTTGTATGATTACGGATTAGAAGACAGTTTTCGTAAGCTGAACCCAACAGCA 43255

Qy        292 HisCysGlnTyrThrTrp------------AspThrGlnMetAsnSerAsnLeuGlyIle 307
              :::   :::::::::|||                  :::   |||   |||
Db      43254 AATGATAAATTCTCGTGGTTTGATTACCGTTCTAAAGGTTTTGATGATAATCGAGGT--- 43198

Qy        308 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe---ArgAlaAlaAlaGluGlu 326
                          ||||||  |||  |||      :::      ||||||
Db      43197 --------------TTACGAATTGATCATATTTTGGTAAGCCAAAAATTAGCTGAACGA 43153

Qy        327 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 346
                      :::      :::|||||:::      :::||||||
Db      43152 TGTGTCGATGTGGGCATTGCATTAGATATTCGAGCAATGGAAAAA--------------- 43108

Qy        347 PheProSerAspHis 351
                ||||||||||||
Db      43107 ---CCGTCTGATCAC 43096

RESULT 16
US-09-976-594-888
; Sequence 888, Application US/09976594
```

```
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 888
;   LENGTH: 3640
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 5202390CB1
US-09-976-594-888

Alignment Scores:
Pred. No.:                0.0313          Length:          3640
Score:                    108.00          Matches:         75
Percent Similarity:       42.0%           Conservative:    51
Best Local Similarity:    25.0%           Mismatches:      100
Query Match:              5.7%            Indels:          74
DB:                       3               Gaps:            19

US-10-757-745-2 (1-362) x US-09-976-594-888 (1-3640)

Qy          11 ArgGluAlaAlaGluGluGluGly------GluProGluValLysLysArgArgLeuLeu 28
               |||           :::  :::|||       |||     ||||||::::::
Db         697 AGAAGCATTTGTGACCATAAAGGATTCTCCTCCCCAGTCTACAAAAAGAAAAAATATAAG 756

Qy          29 CysValGluPheAlaSerValAlaSerCys-AspAlaAlaValAlaGlnCysPheLeuAl 48
                  :::     |||      ||| |||::::::     :::
Db         757 ACCTGTGGATATGAGCAAGTACCTCTATGCTGATAGTTCTAAAACTGAA---------- 805

Qy          48 aGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluPro---------- 64
                  ::   |||::::::   :::   |||   ::::::||||||
Db         806 -------GAAGAAGAATTGGATGAATCA---AATCCTTTTTATGAACCTAAATCAACTCC 855

Qy          65 ----------------------ProValGluGluSerAlaLeuGluArgArg-------- 74
                                    ||||||:::|||         |||||||||
Db         856 TCCTCCAAATAATTTGGTAAATCCTGTTCAAGAACTAGAAACTGAAAGGCGAGTGAAAAG 915

Qy          75 -------------ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGl 90
                           |||      :::|||   |||||||||   |||         |||||
Db         916 AAAGGCCCCGGCTCCACCAGTCCTCTCA---CCAAAAACAGGAGTA------TTAAATGA 966

Qy          90 uGluThrThrAsp---------------SerThrThrSerLysIleSerProSerGluAs 105
               |   |||                    ||||||:::   |||   ||||||
Db         967 AAACACAGTTTCTGCAGGAAAAGATCTCTCTACTTCTCCTAAGCCAAGCCCTATACCAAG 1026

Qy         105 pThr---------GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp-------- 119
                  ::::::   |||   |||      ||||||:::   |||
Db        1027 TCCTGTTTTGGGGCGAAAGCCAAATGCTAGTCAG---TCTTTGCTTGTATGGTGTAAAGA 1083

Qy         120 ----------AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVa 136
                        |||      |||:::   :::   |||   :::          |||
Db        1084 AGTTACAAAGAACTACCGAGGAGTAAAAATCACCAATTTTACTACATCGTGGAGAAATGG 1143

Qy         136 lCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGlnGluVa 153
                  |||:::   |||:::          :::   ||||||:::|||     :::   ::
Db        1144 TTTATCTTTTTGTGCAATATTACACCACTTTAGACCAGATTTAATTGACTACAAGTCTCT 1203

Qy         153 lIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGl 173
               :   |||                                          :::|||
Db        1204 GAATCCT-----------------------------------CAAGATATTAAAGAGAA 1227

Qy         173 yHisGluGluGlyTyr------PheThrAlaIle------------MetLeuLysLysSe 187
                :::::::::   |||      |||   :::|||            :::|||:::   ||

Db        1228 CAACAAAAAGGCATACGATGGATTTGCCAGCATAGGAATTTCCCGATTATTGGAACCTTC 1287

Qy         187 rArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAs 207
               |   :::   |||   :::          ||||||          |||   :::|||
Db        1288 TGATATGGTATTATTAGCA---------ATTCCTGATAAACTGACT---GTTATGACTTA 1335

Qy         207 nLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLe 227
                |||      :::      :::   ||||||  ||||||  ::::::
Db        1336 TCTCTATCAAATAAGGGCACATTTCAGTGGCCAAGAACTAAATGTCGTTCAGATAGAGGA 1395

Qy         227 uGluSerThr-ArgGlyHis---AlaAlaGluArgMetAsnGlnLeuLysMetValLeuL 246
                  |||::: :::   |||        |||   |||   |||::::::   ||||||
Db        1396 AAACAGCAGTAAAAGCACATATAAAGTTGGAAACTATGAAACAGATACAAACAGTTCTGT 1455

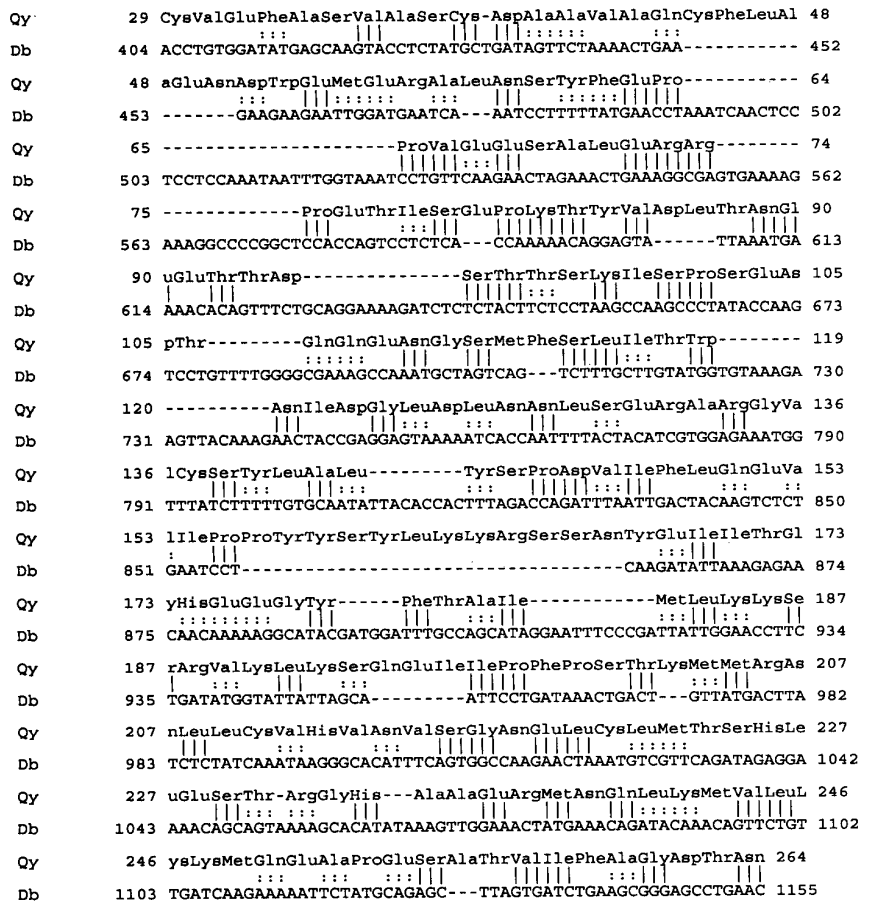
Qy         246 ysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 264
                        :::  :::   :::|||      ||||||   :::|||      |||
Db        1456 TGATCAAGAAAAATTCTATGCAGAGC---TTAGTGATCTGAAGCGGGAGCCTGAAC 1508

RESULT 17
US-09-620-312D-363
; Sequence 363, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 363
;   LENGTH: 4061
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (72)..(3062)
US-09-620-312D-363

Alignment Scores:
Pred. No.:                0.0374          Length:          4061
Score:                    108.00          Matches:         75
Percent Similarity:       42.0%           Conservative:    51
Best Local Similarity:    25.0%           Mismatches:      100
Query Match:              5.7%            Indels:          74
DB:                       3               Gaps:            19

US-10-757-745-2 (1-362) x US-09-620-312D-363 (1-4061)

Qy          11 ArgGluAlaAlaGluGluGluGly------GluProGluValLysLysArgArgLeuLeu 28
               |||           :::  :::|||       |||     ||||||::::::
Db         344 AGAAGCATTTGTGACCATAAAGGATTCTCCTCCCCAGTCTACAAAAAGAAAAAATATAAG 403
```



```
RESULT 18
US-09-949-016-11945/c
; Sequence 11945, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11945
;   LENGTH: 86439
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-11945

Alignment Scores:
Pred. No.:                5.22          Length:          86439
Score:                    108.00        Matches:         56
Percent Similarity:       44.0%         Conservative:    43
Best Local Similarity:    24.9%         Mismatches:      92
Query Match:              5.7%          Indels:          35
DB:                       3             Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-11945 (1-86439)
```

```
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16990
;   LENGTH: 86440
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16990

Alignment Scores:
Pred. No.:                5.22          Length:          86440
Score:                    108.00        Matches:         56
Percent Similarity:       44.0%         Conservative:    43
Best Local Similarity:    24.9%         Mismatches:      92
Query Match:              5.7%          Indels:          35
DB:                       3             Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-16990 (1-86440)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                  :::
Db      28322 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 28264

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||:::|||      ||||||:::
Db      28263 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 28210

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||  |||   :::   |||::::::          |||   |||      :::
Db      28209 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 28153

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  ::::::|||      |||   ::::::
Db      28152 CAGGAA-----------------ACTCATCTCACGTGCAGAGACACACATAGGCTCAAA 28111

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                |||||:::  |||
Db      28110 ATAAAAGGATGGAGAAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 28051

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::      :::       |||      ||||||:::   |||:::
Db      28050 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 27991

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::            ||||||   ::::::      :::
Db      27990 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 27931

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||   |||       :::|||   ||||||
Db      27930 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 27886

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db      27885 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 27829

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db      27828 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 27769

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db      27768 TCAGCTCTGCACCAAGGGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 27712

Qy        293 CysGlnTyrThrTrp 297
                 :::||||||:::
Db      27711 ACAGAATATACATTT 27697

RESULT 20
US-09-949-016-14546/c
; Sequence 14546, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14546
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14546

Alignment Scores:
Pred. No.:                197           Length:          818128
Score:                    108.00        Matches:         56
Percent Similarity:       44.0%         Conservative:    43
Best Local Similarity:    24.9%         Mismatches:      92
Query Match:              5.7%          Indels:          35
DB:                       3             Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-14546 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                  :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||:::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||  |||   :::   |||::::::          |||   |||      :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  ::::::|||      |||   ::::::
Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                |||||:::  |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::      :::       |||      ||||||:::   |||:::
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::            ||||||   ::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                 :::|||          |||   |||       :::|||   ||||||
```

```
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 -------------------------------------ArgAspArgGluValThrArg 272

Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297

Db     401503 ACAGAATATACATTT 401489

RESULT 21
US-09-949-016-14547/c
; Sequence 14547, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14547
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14547

Alignment Scores:
Pred. No.:                197             Length:         818128
Score:                    108.00          Matches:        56
Percent Similarity:       44.0%           Conservative:   43
Best Local Similarity:    24.9%           Mismatches:     92
Query Match:              5.7%            Indels:         35
DB:                       3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14547 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110

Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226

Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 -------------------------------------ArgAspArgGluValThrArg 272

Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297

Db     401503 ACAGAATATACATTT 401489

RESULT 22
US-09-949-016-14548/c
; Sequence 14548, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14548
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14548

Alignment Scores:
Pred. No.:                197             Length:         818128
Score:                    108.00          Matches:        56
Percent Similarity:       44.0%           Conservative:   43
Best Local Similarity:    24.9%           Mismatches:     92
Query Match:              5.7%            Indels:         35
DB:                       3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14548 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
```

```
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db     401944 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226

Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297

Db     401503 ACAGAATATACATTT 401489

RESULT 23
US-09-949-016-14549/c
; Sequence 14549, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14549
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G

US-09-949-016-14549

Alignment Scores:
Pred. No.:              197          Length:          818128
Score:                  108.00       Matches:         56
Percent Similarity:     44.0%        Conservative:    43
Best Local Similarity:  24.9%        Mismatches:      92
Query Match:            5.7%         Indels:          35
DB:                     3            Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-14549 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110

Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db     401944 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226

Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297

Db     401503 ACAGAATATACATTT 401489

RESULT 24
US-09-949-016-14550/c
; Sequence 14550, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
```

```
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14550
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14550

Alignment Scores:
Pred. No.:                197             Length:         818128
Score:                    108.00          Matches:        56
Percent Similarity:       44.0%           Conservative:   43
Best Local Similarity:    24.9%           Mismatches:     92
Query Match:              5.7%            Indels:         35
DB:                       3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14550 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                 :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||       ::::::::::|||  ||||||::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::            |||   |||       :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                :::::::|||       |||   ::::::
Db     401944 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                  ||||||:::   |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||       ||||||:::   |||:::
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::                ||||||   ::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||           |||   |||         :::|||   ||||||
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||       ||||||   |||
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::       |||     ||||||
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
                :::|||||||:::
Db     401503 ACAGAATATACATTT 401489
```

```
RESULT 25
US-09-949-016-14551/c
; Sequence 14551, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14551
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14551

Alignment Scores:
Pred. No.:                197             Length:         818128
Score:                    108.00          Matches:        56
Percent Similarity:       44.0%           Conservative:   43
Best Local Similarity:    24.9%           Mismatches:     92
Query Match:              5.7%            Indels:         35
DB:                       3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14551 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                 :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||       ::::::::::|||  ||||||::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::            |||   |||       :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                :::::::|||       |||   ::::::
Db     401944 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                  ||||||:::   |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||       ||||||:::   |||:::
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::                ||||||   ::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||           |||   |||         :::|||   ||||||
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
```

```
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297

Db     401503 ACAGAATATACATTT 401489

RESULT 26
US-09-949-016-14552/c
; Sequence 14552, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14552
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14552

Alignment Scores:
Pred. No.:                197             Length:         818128
Score:                    108.00          Matches:        56
Percent Similarity:       44.0%           Conservative:   43
Best Local Similarity:    24.9%           Mismatches:     92
Query Match:              5.7%            Indels:         35
DB:                       3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14552 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110

Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db     401944 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226

Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297

Db     401503 ACAGAATATACATTT 401489

RESULT 27
US-09-949-016-14553/c
; Sequence 14553, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14553
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14553

Alignment Scores:
Pred. No.:                197             Length:         818128
Score:                    108.00          Matches:        56
Percent Similarity:       44.0%           Conservative:   43
Best Local Similarity:    24.9%           Mismatches:     92
Query Match:              5.7%            Indels:         35
DB:                       3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14553 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110

Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
```

```
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db     401944 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
Db     401503 ACAGAATATACATTT 401489
```

RESULT 28
US-09-949-016-14554/c
; Sequence 14554, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14554
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14554

Alignment Scores:

Pred. No.:	197	Length:	818128
Score:	108.00	Matches:	56
Percent Similarity:	44.0%	Conservative:	43
Best Local Similarity:	24.9%	Mismatches:	92
Query Match:	5.7%	Indels:	35
DB:	3	Gaps:	9

US-10-757-745-2 (1-362) x US-09-949-016-14554 (1-818128)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db     401944 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
Db     401503 ACAGAATATACATTT 401489
```

RESULT 29
US-09-949-016-14555/c
; Sequence 14555, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14555
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14555

Alignment Scores:
Pred. No.:              197          Length:         818128
Score:                  108.00       Matches:        56
Percent Similarity:     44.0%        Conservative:   43
Best Local Similarity:  24.9%        Mismatches:     92
Query Match:            5.7%         Indels:         35
DB:                     3            Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14555 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||  |||  ||||||     :::  ||||||:::                    :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      :::::::::|||   ||||||:::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::  |||::::::              |||   |||      :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  :::::::|||      |||   ::::::
Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                  ||||||:::   |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::      :::          |||      ||||||:::   |||:::
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::            ||||||   ::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::|||           |||   |||          :::|||   ||||||
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||       :::|||   |||:::|||      ||||||   |||
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                        :::::::::|||::::::      |||      ||||||
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy         293 CysGlnTyrThrTrp 297
               :::|||||||:::
Db     401503 ACAGAATATACATTT 401489

RESULT 30
US-09-949-016-14556/c
; Sequence 14556, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14556
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14556

Alignment Scores:
Pred. No.:              197          Length:         818128
Score:                  108.00       Matches:        56
Percent Similarity:     44.0%        Conservative:   43
Best Local Similarity:  24.9%        Mismatches:     92
Query Match:            5.7%         Indels:         35
DB:                     3            Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14556 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||  |||  ||||||     :::  ||||||:::                    :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      :::::::::|||   ||||||:::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::  |||::::::              |||   |||      :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  :::::::|||      |||   ::::::
Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                  ||||||:::   |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::      :::          |||      ||||||:::   |||:::
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::            ||||||   ::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::|||           |||   |||          :::|||   ||||||
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||       :::|||   |||:::|||      ||||||   |||
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621
```

```
Qy        266 ----------------------------------------ArgAspArgGluValThrArg 272
                                                      :::|||   :::
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                               ::::::::::|||::::::      |||     ||||||
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
                ::::||||||:::
Db     401503 ACAGAATATACATTT 401489

RESULT 31
US-09-949-016-14557/c
; Sequence 14557, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14557
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14557

Alignment Scores:
Pred. No.:                197             Length:         818128
Score:                    108.00          Matches:        56
Percent Similarity:       44.0%           Conservative:   43
Best Local Similarity:    24.9%           Mismatches:     92
Query Match:              5.7%            Indels:         35
DB:                       3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14557 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||     :::   ||||||:::           :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||     :::::::::|||   ||||||:::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||  |||   :::   |||::::::           |||   |||      :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                 ::::::|||      |||   ::::::
Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                    ||||||:::   |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::          |||       ||||||:::   |||:::
```

```
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                  ::::::   |||   |||:::           ||||||   ::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                 :::|||            |||   |||          :::|||   ||||||
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                ::::||      :::|||   |||:::|||      ||||||   |||
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ----------------------------------------ArgAspArgGluValThrArg 272
                                                      :::|||   :::
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                               ::::::::::|||::::::      |||     ||||||
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
                ::::||||||:::
Db     401503 ACAGAATATACATTT 401489

RESULT 32
US-09-949-016-14558/c
; Sequence 14558, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14558
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14558

Alignment Scores:
Pred. No.:                197             Length:         818128
Score:                    108.00          Matches:        56
Percent Similarity:       44.0%           Conservative:   43
Best Local Similarity:    24.9%           Mismatches:     92
Query Match:              5.7%            Indels:         35
DB:                       3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14558 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||     :::   ||||||:::           :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||     :::::::::|||   ||||||:::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002
```

```
Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ::::|||   |||   :::   |||::::::            |||   |||      :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                 ::::::|||      |||   ::::::
Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                |||||:::   |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::          |||      |||||:::   |||:::
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::              ||||||   ::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||   |||          :::|||   ||||||
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||          :::|||   |||:::|||      ||||||   |||
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              :::::::::|||::::::      |||      ||||||
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db     401503 ACAGAATATACATTT 401489

RESULT 33
US-09-949-016-14559/c
; Sequence 14559, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14559
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14559

Alignment Scores:
Pred. No.:              197          Length:         818128
Score:                  108.00       Matches:        56
Percent Similarity:     44.0%        Conservative:   43
Best Local Similarity:  24.9%        Mismatches:     92
Query Match:            5.7%         Indels:         35
DB:                     3            Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14559 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                  :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||       :::::::::|||   ||||||:::|||       ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::            |||   |||      :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                 ::::::|||      |||   ::::::
Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                |||||:::   |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::          |||      |||||:::   |||:::
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::              ||||||   ::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||   |||          :::|||   ||||||
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||          :::|||   |||:::|||      ||||||   |||
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                                 :::::::::|||::::::      |||      ||||||
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db     401503 ACAGAATATACATTT 401489

RESULT 34
US-09-949-016-14560/c
; Sequence 14560, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949;016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 14560
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14560

Alignment Scores:
Pred. No.:                197             Length:         818128
Score:                    108.00          Matches:        56
Percent Similarity:       44.0%           Conservative:   43
Best Local Similarity:    24.9%           Mismatches:     92
Query Match:              5.7%            Indels:         35
DB:                       3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14560 (1-818128)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::                :::
Db      402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy          111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||       :::::::::|||   ||||||::|||      ||||||:::
Db      402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy          131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                :::|||   |||   :::   |||::::::              |||   |||      :::
Db      402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy          151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
                ||||||                  ::::::|||      |||   ::::::
Db      401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy          171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
                |||   |||      :::   |||                |||||:::   |||
Db      401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy          191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
                :::         :::         |||      ||||||:::   |||:::
Db      401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy          211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                  ::::::   |||   |||:::            ||||||   ::::::      :::
Db      401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy          227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                    :::|||            |||   |||          :::|||   ||||||
Db      401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy          247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                ::::|||          :::|||   |||:::|||      ||||||   |||
Db      401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy          266 ------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db      401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy          273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                         :::::::::|||:::::      |||      ||||||
Db      401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy          293 CysGlnTyrThrTrp 297
                ::::||||||:::
Db      401503 ACAGAATATACATTT 401489

RESULT 35
US-09-949-016-14561/c
; Sequence 14561, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14561
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14561

Alignment Scores:
Pred. No.:                197             Length:         818128
Score:                    108.00          Matches:        56
Percent Similarity:       44.0%           Conservative:   43
Best Local Similarity:    24.9%           Mismatches:     92
Query Match:              5.7%            Indels:         35
DB:                       3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14561 (1-818128)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::                :::
Db      402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy          111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||       :::::::::|||   ||||||::|||      ||||||:::
Db      402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy          131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                :::|||   |||   :::   |||::::::              |||   |||      :::
Db      402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy          151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
                ||||||                  ::::::|||      |||   ::::::
Db      401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy          171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
                |||   |||      :::   |||                |||||:::   |||
Db      401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy          191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
                :::         :::         |||      ||||||:::   |||:::
Db      401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy          211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                  ::::::   |||   |||:::            ||||||   ::::::      :::
Db      401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy          227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                    :::|||            |||   |||          :::|||   ||||||
Db      401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy          247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                ::::|||          :::|||   |||:::|||      ||||||   |||
Db      401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy          266 ------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db      401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561
```

```
Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||:::::::     |||      ||||||
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
                :::||||||:::
Db     401503 ACAGAATATACATTT 401489

RESULT 36
US-09-949-016-14562/c
; Sequence 14562, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14562
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14562

Alignment Scores:
Pred. No.:              197             Length:         818128
Score:                  108.00          Matches:        56
Percent Similarity:     44.0%           Conservative:   43
Best Local Similarity:  24.9%           Mismatches:     92
Query Match:            5.7%            Indels:         35
DB:                     3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14562 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                 :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||      :::::::::|||   ||||||:::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::             |||   |||      :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                :::::::|||      |||   ::::::
Db     401944 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                 ||||||:::   |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||      ||||||:::   |||:::
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
```

```
                :::::::   |||   |||:::            ||||||   :::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||   |||         :::|||   ||||||
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   ::::||   :::
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||:::::::     |||      ||||||
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
                :::||||||:::
Db     401503 ACAGAATATACATTT 401489

RESULT 37
US-09-949-016-14564/c
; Sequence 14564, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14564
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14564

Alignment Scores:
Pred. No.:              197             Length:         818128
Score:                  108.00          Matches:        56
Percent Similarity:     44.0%           Conservative:   43
Best Local Similarity:  24.9%           Mismatches:     92
Query Match:            5.7%            Indels:         35
DB:                     3               Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-14564 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                 :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||      :::::::::|||   ||||||:::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::             |||   |||      :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945
```

```
Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
Db     401503 ACAGAATATACATTT 401489
```

RESULT 38
US-09-949-016-14565/c
; Sequence 14565, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14565
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14565

Alignment Scores:

Pred. No.:	197	Length:	818128
Score:	108.00	Matches:	56
Percent Similarity:	44.0%	Conservative:	43
Best Local Similarity:	24.9%	Mismatches:	92
Query Match:	5.7%	Indels:	35
DB:	3	Gaps:	9

US-10-757-745-2 (1-362) x US-09-949-016-14565 (1-818128)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
Db     401503 ACAGAATATACATTT 401489
```

RESULT 39
US-09-949-016-14566/c
; Sequence 14566, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14566
; LENGTH: 818128
; TYPE: DNA

```
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14566

Alignment Scores:
Pred. No.:                197          Length:          818128
Score:                    108.00       Matches:         56
Percent Similarity:       44.0%        Conservative:    43
Best Local Similarity:    24.9%        Mismatches:      92
Query Match:              5.7%         Indels:          35
DB:                       3            Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-14566 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              ||| |||  ||||||     :::  ||||||:::                   :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      ::::::::::|||   ||||||:::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||:::::::           |||   |||      :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::|||      |||   ::::::
Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||  |||      ::   |||                 ||||||:::  |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::      :::      |||      ||||||:::   |||:::
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||  |||:::            ||||||   ::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                :::|||              |||   |||         :::|||   ||||||
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   ::::|||   :::
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
                ::::||||||:::
Db     401503 ACAGAATATACATTT 401489

RESULT 40
US-09-949-016-14567/c
; Sequence 14567, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14567
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14567

Alignment Scores:
Pred. No.:                197          Length:          818128
Score:                    108.00       Matches:         56
Percent Similarity:       44.0%        Conservative:    43
Best Local Similarity:    24.9%        Mismatches:      92
Query Match:              5.7%         Indels:          35
DB:                       3            Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-14567 (1-818128)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              ||| |||  ||||||     :::  ||||||:::                   :::
Db     402114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 402056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      ::::::::::|||   ||||||:::|||      ||||||:::
Db     402055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 402002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||:::::::           |||   |||      :::
Db     402001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 401945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::|||      |||   ::::::
Db     401944 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 401903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||  |||      ::   |||                 ||||||:::  |||
Db     401902 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 401843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::      :::      |||      ||||||:::   |||:::
Db     401842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 401783

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||  |||:::            ||||||   ::::::      :::
Db     401782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 401723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                :::|||              |||   |||         :::|||   ||||||
Db     401722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 401678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db     401677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 401621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   ::::|||   :::
Db     401620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 401561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
```

```
Db     401560 TCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 401504

Qy        293 CysGlnTyrThrTrp 297
                :::||||||:::
Db     401503 ACAGAATATACATTT 401489

RESULT 41
US-09-245-281-44/c
; Sequence 44, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
;  APPLICANT: Bertin, John
;  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
;  TITLE OF INVENTION: AND USES THEREOF
;  FILE REFERENCE: 07334/118001
;  CURRENT APPLICATION NUMBER: US/09/245,281
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: US 09/207,359
;  EARLIER FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 09/099,041
;  EARLIER FILING DATE: 1998-06-17
;  EARLIER APPLICATION NUMBER: US 09/019,942
;  EARLIER FILING DATE: 1998-02-06
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
;   LENGTH: 32042
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-245-281-44

Alignment Scores:
Pred. No.:                1.85          Length:          32042
Score:                    106.00        Matches:         56
Percent Similarity:       44.0%         Conservative:    43
Best Local Similarity:    24.9%         Mismatches:      92
Query Match:              5.6%          Indels:          35
DB:                       3             Gaps:            9

US-10-757-745-2 (1-362) x US-09-245-281-44 (1-32042)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||     :::   ||||||:::                  :::
Db      16224 GAAGAAACTGCATCAACTAATGAGCAAAA-TCGCCAGCTAACATCATAATGACAGGATCA 16166

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||     ::::::::::|||   ||||||:::|||      ||||||:::
Db      16165 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 16112

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||::::::              |||   |||     :::
Db      16111 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 16055

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  :::::::|||      |||   ::::::
Db      16054 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 16013

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||     :::   |||                  ||||||:::   |||
Db      16012 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 15953

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::       :::       |||      ||||||:::   |||:::
Db      15952 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 15893

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                 ::::::   |||   |||:::           ||||||   ::::::      :::
Db      15892 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 15833

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||            |||   |||          :::|||   ||||||
Db      15832 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 15788

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                  :::|||         :::|||   |||:::|||      ||||||   |||
Db      15787 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 15731

Qy         266 -------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db      15730 CTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 15671

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                           ::::::::::|||::::::      |||      ||||||
Db      15670 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 15614

Qy         293 CysGlnTyrThrTrp 297
                :::||||||:::
Db      15613 ACAGAATATACATTT 15599

RESULT 42
US-09-340-620A-63/c
; Sequence 63, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
;  APPLICANT: Bertin, John
;  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
;  FILE REFERENCE: 07334-124001
;  CURRENT APPLICATION NUMBER: US/09/340,620A
;  CURRENT FILING DATE:  1999-06-28
;  PRIOR APPLICATION NUMBER: US 09/245,281
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: US 09/207,359
;  PRIOR FILING DATE: 1998-12-08
;  PRIOR APPLICATION NUMBER: US 09/099,041
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: US 09/019,942
;  PRIOR FILING DATE: 1998-02-06
;  NUMBER OF SEQ ID NOS: 71
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63
;   LENGTH: 32042
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-340-620A-63

Alignment Scores:
Pred. No.:                1.85          Length:          32042
Score:                    106.00        Matches:         56
Percent Similarity:       44.0%         Conservative:    43
Best Local Similarity:    24.9%         Mismatches:      92
Query Match:              5.6%          Indels:          35
DB:                       3             Gaps:            9

US-10-757-745-2 (1-362) x US-09-340-620A-63 (1-32042)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||     :::   ||||||:::                  :::
Db      16224 GAAGAAACTGCATCAACTAATGAGCAAAA-TCGCCAGCTAACATCATAATGACAGGATCA 16166

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||     ::::::::::|||   ||||||:::|||      ||||||:::
Db      16165 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 16112

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||::::::              |||   |||     :::
Db      16111 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 16055

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  :::::::|||      |||   ::::::
Db      16054 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 16013

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||     :::   |||                  ||||||:::   |||
Db      16012 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 15953

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
```

```
Db      15952 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 15893

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db      15892 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 15833

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db      15832 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 15788

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db      15787 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 15731

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db      15730 CTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 15671

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db      15670 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 15614

Qy        293 CysGlnTyrThrTrp 297
Db      15613 ACAGAATATACATTT 15599

RESULT 43
US-09-728-721-63/c
; Sequence 63, Application US/09728721
; Patent No. 6869775
; GENERAL INFORMATION:
;  APPLICANT: Bertin, John
;  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
;  FILE REFERENCE: 07334-124001
;  CURRENT APPLICATION NUMBER: US/09/728,721
;  CURRENT FILING DATE:  2000-12-01
;  PRIOR APPLICATION NUMBER: 09/340,620
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: US 09/207,359
;  PRIOR FILING DATE: 1998-12-08
;  PRIOR APPLICATION NUMBER: US 09/099,041
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: US 09/019,942
;  PRIOR FILING DATE: 1998-02-06
;  NUMBER OF SEQ ID NOS: 71
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63
;   LENGTH: 32042
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-728-721-63

Alignment Scores:
Pred. No.:              1.85          Length:         32042
Score:                  106.00        Matches:        56
Percent Similarity:     44.0%         Conservative:   43
Best Local Similarity:  24.9%         Mismatches:     92
Query Match:            5.6%          Indels:         35
DB:                     3             Gaps:           9

US-10-757-745-2 (1-362) x US-09-728-721-63 (1-32042)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db      16224 GAAGAAACTGCATCAACTAATGAGCAAAA-TCGCCAGCTAACATCATAATGACAGGATCA 16166

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db      16165 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 16112

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      16111 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 16055

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db      16054 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 16013

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db      16012 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 15953

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db      15952 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 15893

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db      15892 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 15833

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db      15832 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 15788

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db      15787 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 15731

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db      15730 CTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 15671

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db      15670 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 15614

Qy        293 CysGlnTyrThrTrp 297
Db      15613 ACAGAATATACATTT 15599

RESULT 44
US-09-949-002-810/c
; Sequence 810, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 810
;   LENGTH: 39553
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-810

Alignment Scores:
Pred. No.:              2.6           Length:         39553
Score:                  106.00        Matches:        56
Percent Similarity:     44.0%         Conservative:   43
Best Local Similarity:  24.9%         Mismatches:     92
Query Match:            5.6%          Indels:         35
DB:                     3             Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-002-810 (1-39553)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db      20860 GAAGAAACTGCATCAACTAATGAGCAAAA-TCGCCAGCTAACATCATAATGACAGGATCA 20802
```

```
Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db      20801 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 20748

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      20747 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 20691

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db      20690 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 20649

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db      20648 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 20589

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db      20588 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 20529

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
Db      20528 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 20469

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db      20468 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 20424

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db      20423 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 20367

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db      20366 CTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 20307

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db      20306 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 20250

Qy        293 CysGlnTyrThrTrp 297
Db      20249 ACAGAATATACATTT 20235
```

RESULT 45
US-09-949-002-673/c
; Sequence 673, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 673
; LENGTH: 58111
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-673

Alignment Scores:

Pred. No.:	4.84	Length:	58111
Score:	106.00	Matches:	56
Percent Similarity:	44.0%	Conservative:	43
Best Local Similarity:	24.9%	Mismatches:	92
Query Match:	5.6%	Indels:	35

```
DB:                    3              Gaps:          9

US-10-757-745-2 (1-362) x US-09-949-002-673 (1-58111)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db      39463 GAAGAAACTGCATCAACTAATGAGCAAAA-TCGCCAGCTAACATCATAATGACAGGATCA 39405

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db      39404 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 39351

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      39350 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 39294

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db      39293 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 39252

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db      39251 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 39192

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db      39191 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 39132

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
Db      39131 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 39072

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db      39071 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 39027

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db      39026 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 38970

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db      38969 CTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 38910

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db      38909 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 38853

Qy        293 CysGlnTyrThrTrp 297
Db      38852 ACAGAATATACATTT 38838
```

RESULT 46
US-09-949-016-17367/c
; Sequence 17367, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17367
; LENGTH: 102520

```
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17367

Alignment Scores:
Pred. No.:              12.1          Length:         102520
Score:                  106.00        Matches:        56
Percent Similarity:     44.0%         Conservative:   43
Best Local Similarity:  24.9%         Mismatches:     92
Query Match:            5.6%          Indels:         35
DB:                     3             Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-17367 (1-102520)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||     :::   |||||||:::                  :::
Db      57481 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCG 57423

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||:::|||      ||||||:::
Db      57422 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 57369

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::              |||   |||      :::
Db      57368 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 57312

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                    :::::::|||      |||   ::::::
Db      57311 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 57270

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                  ||||||:::   |||
Db      57269 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 57210

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||       ||||||:::   |||:::
Db      57209 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 57150

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                :::::   |||   |||:::            ||||||   ::::::      :::
Db      57149 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 57090

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                :::|||             |||   |||         :::|||   ||||||
Db      57089 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 57045

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                :::|||          :::|||   |||:::|||      ||||||   |||
Db      57044 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 56988

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db      56987 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 56928

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db      56927 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 56871

Qy        293 CysGlnTyrThrTrp 297
                :::||||||:::
Db      56870 ACAGAATATACATTT 56856

RESULT 47
US-09-949-016-12448/c
; Sequence 12448, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12448
;   LENGTH: 102526
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12448

Alignment Scores:
Pred. No.:              12.1          Length:         102526
Score:                  106.00        Matches:        56
Percent Similarity:     44.0%         Conservative:   43
Best Local Similarity:  24.9%         Mismatches:     92
Query Match:            5.6%          Indels:         35
DB:                     3             Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-12448 (1-102526)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||     :::   |||||||:::                  :::
Db      57488 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCG 57430

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||:::|||      ||||||:::
Db      57429 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 57376

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::              |||   |||      :::
Db      57375 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 57319

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                    :::::::|||      |||   ::::::
Db      57318 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 57277

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                  ||||||:::   |||
Db      57276 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 57217

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||       ||||||:::   |||:::
Db      57216 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 57157

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                :::::   |||   |||:::            ||||||   ::::::      :::
Db      57156 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 57097

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                :::|||             |||   |||         :::|||   ||||||
Db      57096 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 57052

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                :::|||          :::|||   |||:::|||      ||||||   |||
Db      57051 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 56995

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db      56994 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 56935

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db      56934 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 56878

Qy        293 CysGlnTyrThrTrp 297
                :::||||||:::
Db      56877 ACAGAATATACATTT 56863

RESULT 48
```

```
US-09-949-016-14353/c
; Sequence 14353, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14353
;   LENGTH: 175236
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14353

Alignment Scores:
Pred. No.:                28.8          Length:          175236
Score:                    106.00        Matches:         56
Percent Similarity:       44.0%         Conservative:    43
Best Local Similarity:    24.9%         Mismatches:      92
Query Match:              5.6%          Indels:          35
DB:                       3             Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-14353 (1-175236)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||  ||||||      :::   ||||||:::                  :::
Db        9404 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 9346

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      :::::::::|||  ||||||:::|||      ||||||:::
Db        9345 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 9292

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||::::::            |||   |||      :::
Db        9291 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 9235

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 ::::::|||      |||   ::::::
Db        9234 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 9193

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||  |||      :::   |||                |||||:::   |||
Db        9192 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 9133

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::      :::         |||      ||||||:::   |||:::
Db        9132 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 9073

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::            ||||||   ::::::      :::
Db        9072 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 9013

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||         |||   |||          :::|||   ||||||
Db        9012 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 8968

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||         :::|||   |||:::|||      ||||||   |||
Db        8967 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 8911

Qy         266 ---------------------------------------ArgAspArgGluValThrArg 272
                                                     :::|||  :::
Db        8910 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 8851

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                           ::::::::::|||:::::      |||      ||||||
Db        8850 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 8794

Qy         293 CysGlnTyrThrTrp 297
                  :::||||||:::
Db        8793 ACAGAATATACATTT 8779

RESULT 49
US-09-949-016-15974/c
; Sequence 15974, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15974
;   LENGTH: 294836
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(294836)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15974

Alignment Scores:
Pred. No.:                66.7          Length:          294836
Score:                    106.00        Matches:         56
Percent Similarity:       44.0%         Conservative:    43
Best Local Similarity:    24.9%         Mismatches:      92
Query Match:              5.6%          Indels:          35
DB:                       3             Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-15974 (1-294836)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||  ||||||      :::   ||||||:::                  :::
Db      269774 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 269716

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      :::::::::|||  ||||||:::|||      ||||||:::
Db      269715 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 269662

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||::::::            |||   |||      :::
Db      269661 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 269605

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 ::::::|||      |||   ::::::
Db      269604 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 269563

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||  |||      :::   |||                |||||:::   |||
Db      269562 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 269503

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::      :::         |||      ||||||:::   |||:::
Db      269502 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 269443

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
```

```
Db     269442 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 269383

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db     269382 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 269338

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db     269337 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 269281

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db     269280 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 269221

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db     269220 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 269164

Qy        293 CysGlnTyrThrTrp 297

Db     269163 ACAGAATATACATTT 269149

RESULT 50
US-09-949-016-12008
; Sequence 12008, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12008
;   LENGTH: 331814
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(331814)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12008

Alignment Scores:
Pred. No.:              80.7          Length:         331814
Score:                  106.00        Matches:        56
Percent Similarity:     44.0%         Conservative:   43
Best Local Similarity:  24.9%         Mismatches:     92
Query Match:            5.6%          Indels:         35
DB:                     3             Gaps:           9

US-10-757-745-2 (1-362) x US-09-949-016-12008 (1-331814)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110

Db     134994 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 135052

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db     135053 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 135106

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db     135107 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 135163

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db     135164 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 135205

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db     135206 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 135265

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db     135266 ATCCTAGTCTCTGACAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 135325

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226

Db     135326 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 135385

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db     135386 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 135430

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db     135431 CACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 135487

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db     135488 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 135547

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db     135548 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 135604

Qy        293 CysGlnTyrThrTrp 297

Db     135605 ACAGAATATACATTT 135619

RESULT 51
US-09-949-016-17056
; Sequence 17056, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17056
;   LENGTH: 331814
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(331814)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17056

Alignment Scores:
Pred. No.:              80.7          Length:         331814
Score:                  106.00        Matches:        56
Percent Similarity:     44.0%         Conservative:   43
Best Local Similarity:  24.9%         Mismatches:     92
Query Match:            5.6%          Indels:         35
DB:                     3             Gaps:           9
```

```
US-10-757-745-2 (1-362) x US-09-949-016-17056 (1-331814)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||  ||||||     :::   ||||||:::                  :::
Db      134994 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 135052

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||      ::::::::::|||   ||||||:::|||      ||||||:::
Db      135053 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 135106

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||  :::   |||::::::              |||   |||      :::
Db      135107 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 135163

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||      |||   ::::::
Db      135164 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 135205

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                    ||||||:::   |||
Db      135206 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 135265

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::          :::        |||       ||||||:::   |||:::
Db      135266 ATCCTAGTCTCTGACAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 135325

Qy         211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||  |||:::          ||||||   ::::::      :::
Db      135326 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 135385

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                 :::|||             |||   |||         :::|||   ||||||
Db      135386 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 135430

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||         :::|||   |||:::|||      ||||||   |||
Db      135431 CACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 135487

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db      135488 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 135547

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                        :::::::::|||:::::      |||      ||||||
Db      135548 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 135604

Qy         293 CysGlnTyrThrTrp 297
               :::|||||:::
Db      135605 ACAGAATATACATTT 135619

RESULT 52
US-09-949-016-13527/c
; Sequence 13527, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13527
;   LENGTH: 206433
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13527

Alignment Scores:
Pred. No.:                49.7          Length:          206433
Score:                    105.00        Matches:         55
Percent Similarity:       44.0%         Conservative:    44
Best Local Similarity:    24.4%         Mismatches:      92
Query Match:              5.5%          Indels:          35
DB:                       3             Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-13527 (1-206433)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||  ||||||     :::   ||||||:::                  :::
Db      190570 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 190512

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||      ::::::::::|||   |||::::::|||      ||||||:::
Db      190511 AATTCACACATAACAATATTAACTTTAAATGTAAATGGA------CTAAATTCTGCAATT 190458

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||  :::   |||::::::              |||   |||      :::
Db      190457 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 190401

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||      |||   ::::::
Db      190400 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 190359

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                    ||||||:::   |||
Db      190358 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 190299

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::          :::        |||       ||||||:::   |||:::
Db      190298 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 190239

Qy         211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||  |||:::          ||||||   ::::::      :::
Db      190238 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATTTAT 190179

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                 :::|||             |||   |||         :::|||   ||||||
Db      190178 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 190134

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||         :::|||   |||:::|||      ||||||   |||
Db      190133 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 190077

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db      190076 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 190017

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                        :::::::::|||:::::      |||      ||||||
Db      190016 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 189960

Qy         293 CysGlnTyrThrTrp 297
               :::|||||:::
Db      189959 ACAGAGTATACATTT 189945

RESULT 53
US-09-949-016-12417/c
; Sequence 12417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
```

```
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12417
;   LENGTH: 254778
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12417

Alignment Scores:
Pred. No.:                69.9          Length:          254778
Score:                    105.00        Matches:         55
Percent Similarity:       44.0%         Conservative:    44
Best Local Similarity:    24.4%         Mismatches:      92
Query Match:              5.5%          Indels:          35
DB:                       3             Gaps:            9

US-10-757-745-2 (1-362) x US-09-949-016-12417 (1-254778)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||     :::   ||||||:::                  :::
Db     190915 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 190857

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||      :::::::::|||   |||:::::|||      ||||||:::
Db     190856 AATTCACACATAACAATATTAACTTTAAATGTAAATGGA------CTAAATTCTGCAATT 190803

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::          |||   |||      :::
Db     190802 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 190746

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::::|||      |||   ::::::
Db     190745 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 190704

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                 ||||||:::   |||
Db     190703 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 190644

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||      ||||||:::   |||:::
Db     190643 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 190584

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::             ||||||   ::::::      :::
Db     190583 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATTTAT 190524

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||         |||  |||          :::|||   ||||||
Db     190523 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 190479

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||  |||:::|||      ||||||   |||
Db     190478 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 190422

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||  :::
Db     190421 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 190362

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db     190361 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 190305

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db     190304 ACAGAGTATACATTT 190290

RESULT 54
US-09-949-016-12805/c
; Sequence 12805, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12805
;   LENGTH: 421491
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(421491)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12805

Alignment Scores:
Pred. No.:                158           Length:          421491
Score:                    105.00        Matches:         47
Percent Similarity:       45.8%         Conservative:    45
Best Local Similarity:    23.4%         Mismatches:      75
Query Match:              5.5%          Indels:          34
DB:                       3             Gaps:            10

US-10-757-745-2 (1-362) x US-09-949-016-12805 (1-421491)

Qy        115 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134
              :::::::::|||   |||::::::||||||          |||          :::
Db     289557 ACAATATTAACCTTAAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACAC 289507

Qy        135 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 154
                 :::   :::::::::          |||   |||   ::::::||||||
Db     289506 AGACTGGCAAATTGGATAAAGAATCAAGACCCATCAGTGTGCTATATTCAGGAG------ 289453

Qy        155 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 174
                          ::::::|||      |||   ::::::          |||   |||
Db     289452 -----------ACCCATCTCACATGCAGAGACACACATAGGCTCAAAATAAAGGGATGG 289405

Qy        175 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194
                 :::   |||                     ||||||:::   |||   :::
Db     289404 AGGAAAATCTACCAAGCAAATGGAAAACAAAAGAAAGCAGGGGTTGCAATCCTAGTCTCT 289345

Qy        195 GluIleIleProPheProSerThrLysMetMetArg-----------AsnLeuLeuCys 210
                 :::   |||      ||||||:::   |||               :::   |||
Db     289344 GATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACTTAATG 289285

Qy        211 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 230
              |||      ::::::          ||||||   ::::::      :::      :::|||
Db     289284 GTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATATGCACACAATACA 289225

Qy        231 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 250
              |||          |||   |||          :::|||   ||||||:::   :::|||:::
Db     289224 AGA---------GCACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAAA 289180

Qy        251 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn-----------LeuArg 266
                 :::|||   |||:::|||      ||||||   |||              :::
Db     289179 GACTTAGATTCCCACACAATAATAATG---GGAGACTTTAACACCCCACTGTCAATATTA 289123

Qy        267 AspArg----------------------GluValThrArgCysGlyGlyLeuProAsn 278
              ||||||                     ::::::                  |||:::
Db     289122 GACAGATCAACCAGACAGAAAGTTAACAAGGATATCCAGGACTTAAACTCAGCTCCACAC 289063
```

```
Qy        279 ------AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 296
                    :::::::::   ::::::      |||      ||||||      :::||||||
Db     289062 CAAGCAGACCTAATACACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACA 289006

Qy        297 Trp 297
                 :::
Db     289005 TTC 289003
```

RESULT 55
US-09-949-016-14060/c
; Sequence 14060, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14060
; LENGTH: 421494
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(421494)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14060

Alignment Scores:

Pred. No.:	158	Length:	421494
Score:	105.00	Matches:	47
Percent Similarity:	45.8%	Conservative:	45
Best Local Similarity:	23.4%	Mismatches:	75
Query Match:	5.5%	Indels:	34
DB:	3	Gaps:	10

US-10-757-745-2 (1-362) x US-09-949-016-14060 (1-421494)

```
Qy        115 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134
              ::::::::::|||  |||:::::|||||      |||      :::
Db     289557 ACAATATTAACCTTAAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACAC 289507

Qy        135 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 154
                :::   :::::::::            |||  |||   ::::::||||||
Db     289506 AGACTGGCAAATTGGATAAAGAATCAAGACCCATCAGTGTGCTATATTCAGGAG------ 289453

Qy        155 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 174
                    :::::|||      |||   ::::::       |||   |||
Db     289452 ------------ACCCATCTCACATGCAGAGACACACATAGGCTCAAAATAAAGGGATGG 289405

Qy        175 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194
              :::   |||          |||||:::   |||   :::
Db     289404 AGGAAAATCTACCAAGCAAATGGAAAACAAAAGAAAGCAGGGGTTGCAATCCTAGTCTCT 289345

Qy        195 GluIleIleProPheProSerThrLysMetMetArg------------AsnLeuLeuCys 210
              :::       |||      ||||||:::   |||            :::   |||
Db     289344 GATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACTTAATG 289285

Qy        211 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 230
              |||      ::::::       ||||||   ::::::      :::      :::|||
Db     289284 GTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATATGCACACAATACA 289225

Qy        231 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 250
              |||         |||   |||         :::|||   ||||||:::   :::|||:::
```

```
Db     289224 AGA---------GCACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAAA 289180

Qy        251 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn------------LeuArg 266
                  :::|||   |||:::|||      ||||||   |||             :::
Db     289179 GACTTAGATTCCCACACAATAATAATG---GGAGACTTTAACACCCCACTGTCAATATTA 289123

Qy        267 AspArg------------------------GluValThrArgCysGlyGlyLeuProAsn 278
              ||||||                       ::::::                 |||:::
Db     289122 GACAGATCAACCAGACAGAAAGTTAACAAGGATATCCAGGACTTAAACTCAGCTCCACAC 289063

Qy        279 ------AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 296
                    :::::::::   ::::::      |||      ||||||      :::||||||
Db     289062 CAAGCAGACCTAATACACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACA 289006

Qy        297 Trp 297
                 :::
Db     289005 TTC 289003
```

RESULT 56
US-09-949-016-17522/c
; Sequence 17522, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17522
; LENGTH: 61462
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(61462)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17522

Alignment Scores:

Pred. No.:	8.1	Length:	61462
Score:	104.50	Matches:	63
Percent Similarity:	43.3%	Conservative:	50
Best Local Similarity:	24.1%	Mismatches:	99
Query Match:	5.5%	Indels:	49
DB:	3	Gaps:	12

US-10-757-745-2 (1-362) x US-09-949-016-17522 (1-61462)

```
Qy         68 GluSerAlaLeuGluArgArgProGluThrIleSerGlu---------------------- 80
              ::::::   |||   |||   ||||||      ::::::
Db      15084 CAGGCCTGCCTTTCAAGAGCTCCAGAAAGAAGCACTAAACATGGAAAGAAAAAACCGGTA 15025

Qy         81 ---------ProLysThrTyrValAspLeuThrAsnGluGluThrThr---------Asp 94
                       |||||||||      :::      :::|||
Db      15024 CCAGCAACTGCAAAAACATACCAAATTGTAAAGACCATCGACACTGTGAAGAAACTGCAA 14965

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114
                    |||||||||:::            |||      :::   |||
Db      14964 CAACTAATGAGCAAAATAACTAGCATCATAATGACA------GGATCAAATTCACACATA 14911

Qy        115 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134
              :::::::::|||   ||||||:::||||||          |||:::      :::
Db      14910 ACAATATTAACCTTGAATATAAATGGGCTA---------AATGTTCCAATTAAAACACAC 14860
```

```
Qy        135 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 154
Db      14859 AGACCGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG------ 14806

Qy        155 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 174
Db      14805 -----------ACCCATCTCATGTGCAAAGACACACATAGGCTCAAAATAAAGGGATGG 14758

Qy        175 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194
Db      14757 AGGAATATTTACCAAGCAAATGGAAAGAAAAAAAAAGCAGGGGTTGCAATCCTAGTTCCT 14698

Qy        195 GluIleIleProPheProSerThrLysMetMetArgAsn-----------LeuLeuCys 210
Db      14697 GATAAAGCAAAATTTAAACCAACAAAGATCAAAAAAGACAAAGAAGGGTATTACATAATG 14638

Qy        211 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 230
Db      14637 GTCAAGGAATCAATGCAACAAGAAGAGCTAACTATCCTAAATATATATGCACCCAATACA 14578

Qy        231 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 250
Db      14577 ---------GGAGCAACCAGATTC------ATAAAGCAAGTTCTTAGAGACCTACAAAGA 14533

Qy        251 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--------------- 265
Db      14532 GACTTAGACTCCCACACAATAATA---GTGGGAGACTTTAACACCCCACTGTCAACATTA 14476

Qy        266 ------ArgAspArgGluValThrArgCys-GlyGlyLeu--------------ProAs 278
Db      14475 GATCAACGAGACAGAAAATTAACAAGGATATTCAGGACTTGAACTCAGCTCCGGACCAAG 14416

Qy        278 n-AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db      14415 CAGACCGAATAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 14360
```

RESULT 57
US-09-949-016-11757/c
; Sequence 11757, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11757
; LENGTH: 131631
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(131631)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11757

Alignment Scores:

Pred. No.:	31.9	Length:	131631
Score:	104.00	Matches:	55
Percent Similarity:	44.0%	Conservative:	44
Best Local Similarity:	24.4%	Mismatches:	92
Query Match:	5.5%	Indels:	35
DB:	3	Gaps:	9

US-10-757-745-2 (1-362) x US-09-949-016-11757 (1-131631)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db     116335 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 116277

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db     116276 AATTCACACATAACAATATTAACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 116223

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db     116222 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 116166

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db     116165 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 116124

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db     116123 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 116064

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db     116063 ATCCTAGTCTCTGATAAAACAGACTTTAAGCCAACAAAGATCAAAAGAGACAAAGAAGGC 116004

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db     116003 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATATAT 115944

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     115943 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGAAAGTCCTGAGT 115899

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db     115898 GACCTACAAAGAGACCTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 115842

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db     115841 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 115782

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     115781 TCAGCTCTGCACCAAGGGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 115725

Qy        293 CysGlnTyrThrTrp 297
Db     115724 ACAGAATATACATTT 115710
```

RESULT 58
US-09-485-473-4
; Sequence 4, Application US/09485473
; Patent No. 6740512
; GENERAL INFORMATION:
; APPLICANT: STOFFEL, WILHELM
; APPLICANT: HOFMANN, KAY
; APPLICANT: TOMIUK, STEPHAN
; TITLE OF INVENTION: NEUTRAL SPHINGOMYELINASE
; FILE REFERENCE: 10496/P61950US1
; CURRENT APPLICATION NUMBER: US/09/485,473
; CURRENT FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1627
; TYPE: DNA
; ORGANISM: Murine sp.
US-09-485-473-4

Alignment Scores:

Pred. No.:	0.0305	Length:	1627
Score:	103.50	Matches:	72

```
Percent Similarity:     38.9%          Conservative:  58
Best Local Similarity:  21.6%          Mismatches:    137
Query Match:            5.4%           Indels:        67
DB:                     3              Gaps:          17

US-10-757-745-2 (1-362) x US-09-485-473-4 (1-1627)

Qy         74 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 93
              ||||||     :::     :::::::::|||           :::
Db        199 CGCCCCCCACCCCCAACCAGCTCGCGTTCCTGGGTCGGGGCAGCGCAGGACAGGGCAATA 258

Qy         94 AspSerThrThrSerLysIleSerPro------------SerGluAspThrGlnGlnGlu 109
                            ||||||              |||          :::
Db        259 AGCCTGTGCGCGCAATCCGCCTCGCCGCCCTTGCTCCGAAGCACTCCAGCCATGAAGCTC 318

Qy        110 AsnGlySerMet------------PheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
              |||   |||:::           |||:::|||      |||:::|||     |||
Db        319 AACTTTTCTCTACGGCTGAGAGTTTTCAATCTCAACTGCTGGGACATCCCCTACCTGAGC 378

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
                 :::   :::::::|||   :::   :::      :::|||   |||   :::
Db        379 ---AAACATAGGGCGGACCGCATGAAGCGCTTGGGAGACTTTCTGAACTTGGAAAACTTT 435

Qy        146 AspValIlePheLeuGlnGluValIleProProTyr---TyrSerTyrLeuLysLysArg 164
              |||:::      |||:::||||||              :::   ||||||::::::|||
Db        436 GATCTGGCTCTCCTGGAGGAGGTGTGGAGTGAGCAGGACTTCCAGTACCTAAGGCAAAGG 495

Qy        165 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 184
                 |||          :::|||      :::          ||||||   :::   |||:::
Db        496 CTATCG------------CTCACCTATCCAGATGCACACTACTTCAGAAGCGGGATGATA 543

Qy        185 LysLysSerArgVal---------LysLeuLysSerGlnGluIleIlePro--------- 198
                                   |||          |||||||||
Db        544 GGCAGTGGCCTCTGTGTGTTCTCCAAACACCCAATCCAGGAAATCTTCCAGCATGTCTAC 603

Qy        199 ------------PheProSerThrLysMetMetArgAsnLeuLeuCys------------ 210
                         :::|||          :::      |||
Db        604 AGTCTGAATGGTTACCCCTACATGTTCCATCATGGAGACTGGTTCTGTGGGAAGTCTGTG 663

Qy        211 ------ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
                   :::   :::   :::||||||        |||                :::||||||
Db        664 GGGCTGCTGGTGCTCCGTCTAAGTGGACTGGTGCTCAATGCCTACGTGACTCATCTACAT 723

Qy        229 SerThrArgGly---------------HisAlaAlaGluArgMetAsnGlnLeuLysMet 243
              :::                     :::   |||   |||:::   |||
Db        724 GCTGAGTACAGCCGACAGAAGGACATCTACTTTGCACACCGTGTGGCCCAAGCTTGGGAA 783

Qy        244 ValLeuLysLysMetGlnGluAlaProGluSerAla---ThrValIlePheAlaGlyAsp 262
              :::   :::  :::           ::::::|||       |||:::       ||||||
Db        784 CTGGCCCAGTTCATCCACCACACATCCAAGAATGCAGATGTGGTTCTATTGTGTGGAGAC 843

Qy        263 ThrAsnLeuArgAspArgGluValThrArgCys---------------GlyGlyLeuPro 277
                 |||:::      :::::::::      |||                ||||||
Db        844 CTCAATATGCACCCCAAAGACCTGGGCTGCTGCCTGCTGAAAGAGTGGACAGGGCTCCAT 903

Qy        278 AsnAsnIleValAspValTrpGluPheLeuGlyLys---------------------Pro 290
              :::      |||:::      :::|||   |||                       |||
Db        904 GATGCTTTCGTTGAGACTGAGGACTTTAAGGGCTCTGATGATGGCTGTACCATGGTACCC 963

Qy        291 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 310
              |||:::|||      :::                       :::||||||
Db        964 AAGAACTGCTACGTCAGC---------------CAGCAGGACCTGGGA---CCGTTTCCG 1005

Qy        311 CysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 330
                 :::|||   |||   :::   ::::::|||   :::   |||   |||:::
Db       1006 TCTGGTATCCGGATTGATTACGTGCTTTACAAGGCAGTCTCT---GAGTTCCACGTCTGC 1062

Qy        331 ProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro------ 348
                 :::|||                 |||           |||           |||
Db       1063 TGTGAGACTCTG-----------------AAAACCACTACAGGCTGTGACCCTCACAGT 1104

Qy        349 ------------SerAspHisTrpGlyLeuLeuCysAsnLeu 358
                         |||||||||      |||:::      |||

Db       1105 GACAAGCCCTTCTCTGATCACGAGGCCCTCATGGCTACTTTG 1146

RESULT 59
US-09-949-016-16246
; Sequence 16246, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16246
;   LENGTH: 152655
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(152655)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16246

Alignment Scores:
Pred. No.:              53.8           Length:        152655
Score:                  103.00         Matches:       54
Percent Similarity:     44.0%          Conservative:  45
Best Local Similarity:  24.0%          Mismatches:    92
Query Match:            5.4%           Indels:        35
DB:                     3              Gaps:          9

US-10-757-745-2 (1-362) x US-09-949-016-16246 (1-152655)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                  :::
Db      30445 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 30503

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                  |||       :::::::::|||   |||::::::|||       ||||||:::
Db      30504 AATTCACACATAACAATATTAACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 30557

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||  :::   |||::::::                |||   |||       :::
Db      30558 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 30614

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::|||       |||   ::::::
Db      30615 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 30656

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                ||||||:::  :::
Db      30657 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGATTGCA 30716

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::          |||       ||||||:::   |||:::
Db      30717 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 30776

Qy        211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::                ||||||   ::::::      :::
Db      30777 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATATAT 30836

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::|||              |||   |||          :::|||   ||||||
Db      30837 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 30881
```

```
Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||          :::|||   |||:::|||      ||||||   |||
Db      30882 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 30938

Qy        266 -------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db      30939 CTGTCAACATTAGACAGATCACCGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 30998

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              :::::::::|||::::::      |||      ||||||
Db      30999 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 31055

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db      31056 GCAGAATATACATTT 31070
```

RESULT 60
US-09-949-016-14214
; Sequence 14214, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14214
; LENGTH: 26711
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14214

Alignment Scores:

Pred. No.:	4.27	Length:	26711
Score:	102.00	Matches:	55
Percent Similarity:	43.6%	Conservative:	43
Best Local Similarity:	24.4%	Mismatches:	93
Query Match:	5.4%	Indels:	35
DB:	3	Gaps:	9

US-10-757-745-2 (1-362) x US-09-949-016-14214 (1-26711)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||  ||||||      :::   ||||||:::                  :::
Db       9133 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 9191

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||      :::::::::|||   |||:::  |||      ||||||:::
Db       9192 AATTCACACATAACAATATTAACTTTAAATGTAAGTGGA------CTAAATTCTGCAATT 9245

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||  |||  :::  :::::::::             |||  |||      :::
Db       9246 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 9302

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                :::::|||     |||  ::::::
Db       9303 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 9344

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                |||||:::   |||
Db       9345 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 9404

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||      ||||||:::   |||:::
```

```
Db       9405 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 9464

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::             ||||||   ::::::      :::
Db       9465 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATATAT 9524

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||   |||          :::|||   ||||||
Db       9525 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 9569

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||          :::|||   |||:::|||      ||||||   |||
Db       9570 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 9626

Qy        266 -------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db       9627 CTGTCAATATTAGACAGATCAACAAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 9686

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              :::::::::||||||:::      |||      ||||||
Db       9687 TCAGCTCTGCACCAAGCAGACCTAATAGACGTCTACAGAACTCTC---CACCCCAAATCA 9743

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db       9744 ACAGAATATACATTT 9758
```

RESULT 61
US-09-949-016-15232
; Sequence 15232, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15232
; LENGTH: 35837
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15232

Alignment Scores:

Pred. No.:	6.87	Length:	35837
Score:	102.00	Matches:	56
Percent Similarity:	43.6%	Conservative:	42
Best Local Similarity:	24.9%	Mismatches:	93
Query Match:	5.4%	Indels:	35
DB:	3	Gaps:	9

US-10-757-745-2 (1-362) x US-09-949-016-15232 (1-35837)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||  ||||||      :::   ||||||:::                  :::
Db      26809 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATCATGACAGGATCA 26867

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||      :::::::::|||   ||||||:::|||      ||||||:::
Db      26868 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 26921

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||  |||  :::  |||::::::             |||  |||      :::
Db      26922 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 26978
```

```
Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  ::::::|||       |||    ::::::
Db      26979 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 27020

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                 ||||||:::   |||
Db      27021 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 27080

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::       :::        |||      ||||||:::   |||:::
Db      27081 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 27140

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::              ||||||   :::         :::
Db      27141 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCCAAATATTTAT 27200

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||  |||          :::||||  ||||||
Db      27201 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 27245

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db      27246 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 27302

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db      27303 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 27362

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db      27363 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 27419

Qy        293 CysGlnTyrThrTrp 297
              :::|||||||:::
Db      27420 ACAGAATATACATTT 27434

RESULT 62
US-09-949-016-13002
; Sequence 13002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13002
;   LENGTH: 168394
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(168394)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13002

Alignment Scores:
Pred. No.:                83.6          Length:          168394
Score:                    102.00        Matches:         54
Percent Similarity:       44.4%         Conservative:    46
Best Local Similarity:    24.0%         Mismatches:      91
Query Match:              5.4%          Indels:          35
DB:                       3             Gaps:            9
```

```
US-10-757-745-2 (1-362) x US-09-949-016-13002 (1-168394)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                  :::
Db     132051 GAAGAAACTGCATCAACTAACGTGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 132109

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||       :::::::::|||   |||::::::|||       |||||||:::
Db     132110 AATTCACACATAACAATATTAACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 132163

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||  :::  :::::::::                |||   |||      :::
Db     132164 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 132220

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  ::::::|||       |||    ::::::
Db     132221 CAGGAA-----------------GCCCATCTCACGTGCAGAGACACACATAGGCTCAAA 132262

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                 ||||||:::   |||
Db     132263 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 132322

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::       :::        |||      ||||||:::   |||:::
Db     132323 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGGGACAAAGAAGGC 132382

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::              ||||||   :::::::      :::
Db     132383 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATAT 132442

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||  |||          :::|||   ||||||
Db     132443 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 132487

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||          :::|||   |||:::|||      ||||||   |||
Db     132488 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 132544

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db     132545 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 132604

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db     132605 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CAACCCAAATCA 132661

Qy        293 CysGlnTyrThrTrp 297
              :::|||||||:::
Db     132662 ACAGAATATACATTT 132676

RESULT 63
US-09-949-016-12574
; Sequence 12574, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12574
;   LENGTH: 72843
;   TYPE: DNA
;   ORGANISM: Human
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(72843)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12574

Alignment Scores:
Pred. No.:                24.9          Length:          72843
Score:                    101.50        Matches:         52
Percent Similarity:       42.4%         Conservative:    43
Best Local Similarity:    23.2%         Mismatches:      96
Query Match:              5.3%          Indels:          34
DB:                       3             Gaps:            8

US-10-757-745-2 (1-362) x US-09-949-016-12574 (1-72843)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               ||| |||  |||||| :::    |||:::                          :::
Db       18324 GAAGAAACTGCATCAACTAACGAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 18382

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||     ::::::::::|||  |||::::::||||||         |||  |||
Db       18383 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTTCA 18433

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                 :::       :::  :::::::::          |||  |||      :::
Db       18434 ATCAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 18493

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||     |||  ::::::
Db       18494 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 18535

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||  |||      :::  |||                 ||||||:::  |||
Db       18536 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 18595

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::       :::       |||      ||||||:::  ||:::
Db       18596 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 18655

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                 ::::::   |||  |||:::            ||||||  ::::::      :::
Db       18656 CATTACATAATGGTAAAGGGATCAATTCAACAGGAAGAGCTAACTATCCTAAATATATAT 18715

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::     |||      |||  |||          :::|||   ||||||
Db       18716 GCACCCAACAATACAGGA------GCACCCAGATTC------ATAAAGCAAGTCCTGAGT 18763

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                 :::|||       :::|||  |||:::|||      ||||||  |||
Db       18764 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATT---GGAGACTTTAACACCCCA 18820

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||  :::
Db       18821 CTGTCAATATTAGACAGATCAACGAGACAAAAAGTTAACAAAGATAGCCAGGAATTGAAC 18880

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                        :::::::::|||:::::      |||     ||||||
Db       18881 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 18937

Qy         293 CysGlnTyrThr 296
                  :::||||||
Db       18938 ACAGAATATACA 18949

RESULT 64
US-09-949-002-577
; Sequence 577, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577
;   LENGTH: 179904
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-577

Alignment Scores:
Pred. No.:                123           Length:          179904
Score:                    101.00        Matches:         53
Percent Similarity:       43.3%         Conservative:    44
Best Local Similarity:    23.7%         Mismatches:      95
Query Match:              5.3%          Indels:          33
DB:                       3             Gaps:            8

US-10-757-745-2 (1-362) x US-09-949-002-577 (1-179904)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               ||| |||  |||||| :::    ||||||:::                          :::
Db       23770 GAAGAAACTCCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 23828

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||     ::::::::::|||  |||::::::||||||         |||
Db       23829 AATTCACACATAACAATATTAACTTTGAATGTAAATGGACTA---------AATGCTCCA 23879

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                 :::       :::  :::::::::          |||  |||      :::
Db       23880 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 23939

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||     |||  ::::::
Db       23940 CAGGAA-----------------ACCCATCTCACATGCAGAGACACACATAGGCTCAAG 23981

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||  |||      :::  |||                 ||||||:::  |||
Db       23982 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAAGCAGGGGTTGCA 24041

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::       :::       |||      ||||||:::  |||:::
Db       24042 ATCCTAGTCTCTGATAAAACAGACTTCAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 24101

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                 ::::::   |||  |||:::            ||||||  |||:::      :::
Db       24102 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTCTCCTAAATTTATAT 24161

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::|||       |||  |||          :::|||   ||||||
Db       24162 GCACCCAATACA---------GGAGCACCAAGATTC------ATAAAGCAAGTCCTGAGT 24206

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly---AspThrAsnLeu 265
                 :::|||       :::|||  |||:::|||      |||   :::|||   |||
Db       24207 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATGGGAGGCTTTAACACCCCACTG 24266

Qy         266 ----------------------------------ArgAspArgGluValThrArgCys 273
                                                 :::|||  :::
Db       24267 TCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAACTCA 24326

Qy         274 GlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCys 293
                        :::::::::|||:::::      |||     ||||||
Db       24327 GCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCAACA 24383

Qy         294 GlnTyrThrTrp 297
                  :::||||||:::
Db       24384 GAATATACATTT 24395

RESULT 65
US-09-949-002-705
```

```
; Sequence 705, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 705
;   LENGTH: 179905
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-705

Alignment Scores:
Pred. No.:                123          Length:          179905
Score:                    101.00       Matches:         53
Percent Similarity:       43.3%        Conservative:    44
Best Local Similarity:    23.7%        Mismatches:      95
Query Match:              5.3%         Indels:          33
DB:                       3            Gaps:            8

US-10-757-745-2 (1-362) x US-09-949-002-705 (1-179905)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db      23770 GAAGAAACTCCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 23828

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db      23829 AATTCACACATAACAATATTAACTTTGAATGTAAATGGACTA---------AATGCTCCA 23879

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      23880 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 23939

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db      23940 CAGGAA-----------------ACCCATCTCACATGCAGAGACACACATAGGCTCAAG 23981

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db      23982 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAAGCAGGGGTTGCA 24041

Qy        191 LeuLysSerGlnGluIleIleProProSerThrLysMetMetArgAsnLeuLeuCys 210
Db      24042 ATCCTAGTCTCTGATAAAACAGACTTCAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 24101

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db      24102 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTCTCCTAAATTTATAT 24161

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db      24162 GCACCCAATACA---------GGAGCACCAAGATTC------ATAAAGCAAGTCCTGAGT 24206

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly---AspThrAsnLeu 265
Db      24207 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATGGGAGGCTTTAACACCCCACTG 24266

Qy        266 ----------------------------------ArgAspArgGluValThrArgCys 273
Db      24267 TCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAACTCA 24326

Qy        274 GlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCys 293
Db      24327 GCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCAACA 24383

Qy        294 GlnTyrThrTrp 297
Db      24384 GAATATACATTT 24395

RESULT 66
US-09-270-767-11336
; Sequence 11336, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11336
;   LENGTH: 879
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-11336

Alignment Scores:
Pred. No.:                0.0303       Length:          879
Score:                    100.00       Matches:         51
Percent Similarity:       35.6%        Conservative:    32
Best Local Similarity:    21.9%        Mismatches:      66
Query Match:              5.3%         Indels:          84
DB:                       3            Gaps:            12

US-10-757-745-2 (1-362) x US-09-270-767-11336 (1-879)

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
Db         21 CAGGGATACGATACGAGGATC---CTGAAAGTGTGGCGGGTGCAGAGCAACCAG------ 71

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
Db         72 ---GTGGCCATCGCGGCCCGGCTGCGTATGCGATCC---------------------- 104

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis------ 233
Db        105 AGTGGTCGGGAATTCTGTGTGGCCACCACCCACTTAAAGGCCCGCCATGGCGCCCTTTTG 164

Qy        234 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 253
Db        165 GCGAAATTACGGAACGAACAGGGACGTGATCTCATTAGGTTTGTGAAGCAGTTTGCCGGA 224

Qy        254 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu------------ 269
Db        225 GATACGCCGCTGCTGCTCTGCGGCGATTTCAATGCCGAACCGGTGGAGCCAATCTATGCC 284

Qy        270 ------------------ValThrArgCysGlyGlyLeu---------------------- 276
Db        285 ACGATCTTGGGATGTGATCTCCTCCGCTTGGGCAGCGCCTATGCCGACGTGAAGCTAGAC 344

Qy        277 ------------------ProAsnAsnIleValAspValTrpGluPheLeuGlyLys--- 289
Db        345 CGCGAGGAGATACTACATCCCAAC------GCGGATGTCGGTGAGTTCGTCGCGAAATCG 398

Qy        290 ------------ProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeu 305
Db        399 ATGAAACGCGAGCCACCCTACACC------ACCTGGAAA---------------------- 431

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 325
Db        432 ----------------------------------------ATTCGCGAAGAGGGCGAG 449

Qy        326 GluGlyHisIleIle------------------ProArgSerLeuAspLeu--------- 336
Db        450 GAATGTCACACCATCGACTACGTCTTTTACACACCCGATCGCCTCAAGATCAAAAACTGT 509
```

```
Qy        337 ----------------------LeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348
                                     :::|||      :::           :::::|||
Db        510 CTAGACTTTCCGGCCGGAGAGCAAATTGGCAAGAATCGCACACCCAGCTTCCAATATCCA 569

Qy        349 SerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
              |||||||||:::   |||:::|||:::   :::::::::
Db        570 TCGGATCACTTTTCCCTGGTGTGTGACTTTGAGCTACTG 608

RESULT 67
US-09-080-855-1
; Sequence 1, Application US/09080855A
; Patent No. 6083721
; GENERAL INFORMATION:
;  APPLICANT: Saras, Jan
;  APPLICANT: Franzn, Petra
;  APPLICANT: Aspenstrm, Pontus
;  APPLICANT: Hellman, Ulf
;  APPLICANT: Gonez, Leonel Jorge
;  APPLICANT: Heldin, Carl-Henrik
;  TITLE OF INVENTION: PARG, A GTPASE ACTIVATING PROTEIN WHICH INTERACTS WITH PTPL1
;  FILE REFERENCE: L0461/7030
;  CURRENT APPLICATION NUMBER: US/09/080,855A
;  CURRENT FILING DATE: 1998-05-18
;  EARLIER APPLICATION NUMBER: 08/805,583
;  EARLIER FILING DATE: 1997-02-25
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 5238
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 184..3966
US-09-080-855-1

Alignment Scores:
Pred. No.:                0.542          Length:          5238
Score:                    100.00         Matches:         61
Percent Similarity:       36.4%          Conservative:    46
Best Local Similarity:    20.7%          Mismatches:      117
Query Match:              5.3%           Indels:          70
DB:                       3              Gaps:            12

US-10-757-745-2 (1-362) x US-09-080-855-1 (1-5238)

Qy         17 GluGlyGluProGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAla 36
              |||   |||      :::||||||   :::||||||||||::::::
Db       1141 GAAGCAGAGAATGCTCTCAAAAAGGCAAAATTATTATGCATGCAACGTCAAGAT------ 1194

Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp--------------- 51
                 :::   |||   ::: :    |||   ||||||   :::
Db       1195 GAATATGAGAAAGCAAAGTCTTCCATGTTTCGTGCAGAAGAGGAGCATCTGTCTTCAAGT 1254

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProPro------ValGluGluSer 69
                 :::   :::   ||||||         |||            :::||||||
Db       1255 GGCGGATTAGCAAAAAATCTCAACAAGCAACTAGAAAAAAAGCGAAGGTTGGAAGAGGAG 1314

Qy         70 AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsn 89
              ||||||::::::      |||                |||     |||
Db       1315 GCTCTCCAAAAAGTAGAAGAAGCAGATGAACTTTACAAAGTTTGTGTGACAAATGTTGAA 1374

Qy         90 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 109
              |||         |||                                |||:::|||::::::|||
Db       1375 GAAAGAAGAAATGAT----------------------GTAGAAAATACCAAAAGAGAA 1410

Qy        110 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp--------------------- 122
                   :::          :::  :::      |||
Db       1411 ATTTTAGCACAACTCCGGACACTTGTTTTCCAGTGTGATCTTACCCTTAAAGCGGTAACA 1470

Qy        123 --------------------GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGly 135
                                   |||         :::|||:::::::|||   :::
Db       1471 GTTAACCTCTTCCACATGCAGCATCTGCAGGCTGCTTCCCTTGCAGACAGATTACAGTCT 1530

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              :::|||          ||||||   |||            ||||||
Db       1531 CTCTGTGGTAGTGCCAAACTCTATGACCCAGGC-----------CAAGAG--------- 1569

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSer--------------------------- 166
              |||       ::::::|||      :::|||
Db       1570 ---TACAGTGAATTTGTCAAGGCCACAAATTCAACTGAAGAAGAAAAAGTTGATGGAAAT 1626

Qy        167 ---AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 185
                 |||      :::   :::           |||:::      |||      |||
Db       1627 GTAAATAAACATTTAAATAGTTCCCAACCTTCAGGATTTGGACCTGCCAACTCTTTA--- 1683

Qy        186 LysSerArgValLysLeuLysSerGlnGluIleIleProPhePro------SerThrLys 203
                                        ::::::::::      |||     |||  |||
Db       1684 ----------------------GAGGATGTTGTACGCCTTCCTGACAGTTCTAATAAA 1719

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu--- 222
                 :::      :::                 :::::::::|||
Db       1720 ATTGAAGAGGACAGATGCTCTAACAGTGCAGATATAACAGGTCCTTCCTTTATAAGATCA 1779

Qy        223 -----------MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
                      |||   |||      ||||||||   |||   :::     |||
Db       1780 TGGACATTTGGGATGTTTAGTGATTCTGAGAGCACTGGAGGGAGCAGCGAATCTAGATCT 1839

Qy        239 AsnGlnLeuLysMetValLeu--------------LysLysMetGlnGluAlaProGlu 253
                    :::   :::                  :::|||:::         |||
Db       1840 CTGGATTCAGAATCTATAAGTCCAGGAGACTTTCATCGAAAACTTCCACGAACACCATCC 1899

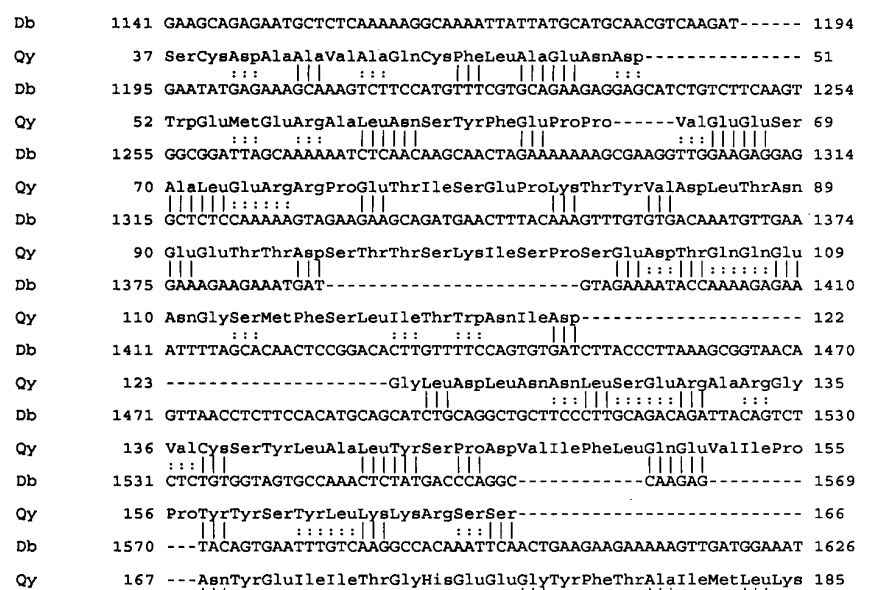
Qy        254 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 267
              |||   |||:::      |||   |||   :::   |||:::
Db       1900 AGTGGAACTATGTCCTCTGCAGATGATCTAGATGAAAGAGAG 1941

RESULT 68
US-09-566-076-1
; Sequence 1, Application US/09566076
; Patent No. 6475775
; GENERAL INFORMATION:
;  APPLICANT: Saras, Jan
;  APPLICANT: Franzn, Petra
;  APPLICANT: Aspenstrm, Pontus
;  APPLICANT: Hellman, Ulf
;  APPLICANT: Gonez, Leonel Jorge
;  APPLICANT: Heldin, Carl-Henrik
;  TITLE OF INVENTION: PARG, A GTPASE ACTIVATING PROTEIN WHICH INTERACTS WITH PTPL1
;  FILE REFERENCE: L0461/7030
;  CURRENT APPLICATION NUMBER: US/09/566,076
;  CURRENT FILING DATE:
;  EARLIER APPLICATION NUMBER: 09/080,855
;  EARLIER FILING DATE: 1998-05-18
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 5238
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 184..3966
US-09-566-076-1

Alignment Scores:
Pred. No.:                0.542          Length:          5238
Score:                    100.00         Matches:         61
Percent Similarity:       36.4%          Conservative:    46
Best Local Similarity:    20.7%          Mismatches:      117
Query Match:              5.3%           Indels:          70
DB:                       3              Gaps:            12

US-10-757-745-2 (1-362) x US-09-566-076-1 (1-5238)

Qy         17 GluGlyGluProGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAla 36
              |||   |||      :::||||||   :::||||||||||::::::
```

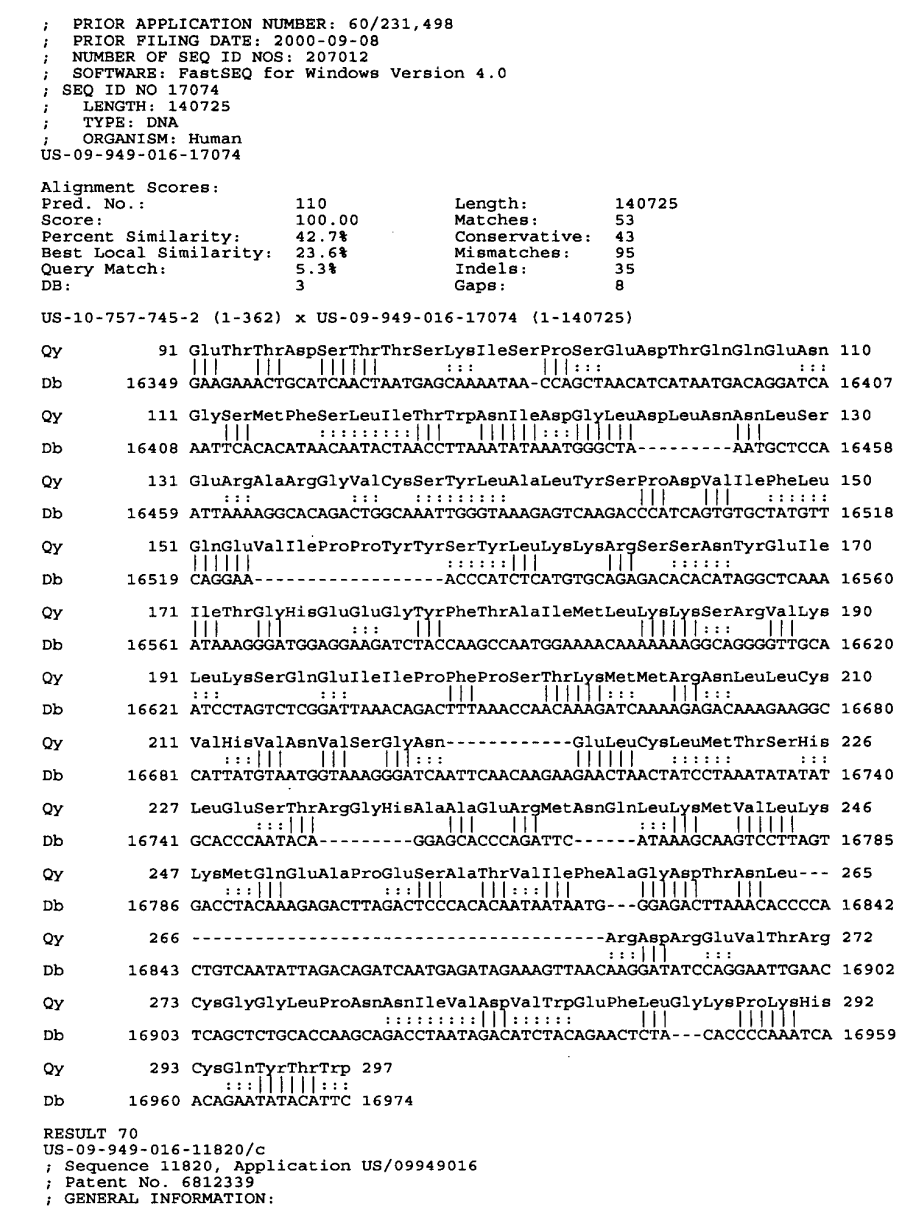


```
RESULT 69
US-09-949-016-17074
; Sequence 17074, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17074
;   LENGTH: 140725
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17074

Alignment Scores:
Pred. No.:                110              Length:           140725
Score:                    100.00           Matches:          53
Percent Similarity:       42.7%            Conservative:     43
Best Local Similarity:    23.6%            Mismatches:       95
Query Match:              5.3%             Indels:           35
DB:                       3                Gaps:             8

US-10-757-745-2 (1-362) x US-09-949-016-17074 (1-140725)
```





```
RESULT 70
US-09-949-016-11820/c
; Sequence 11820, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11820
;   LENGTH: 192700
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-11820

Alignment Scores:
Pred. No.:                183             Length:         192700
Score:                    100.00          Matches:        53
Percent Similarity:       42.7%           Conservative:   43
Best Local Similarity:    23.6%           Mismatches:     95
Query Match:              5.3%            Indels:         35
DB:                       3               Gaps:           8

US-10-757-745-2 (1-362) x US-09-949-016-11820 (1-192700)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::                 :::
Db       48012 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 47954

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||:::||||||         |||
Db       47953 AATTCACACATAACAATATTAACTTTAAATATAAATGGACTA---------AATGCTCCA 47903

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::         :::   :::::::::             |||   |||      :::
Db       47902 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 47843

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||       |||   ::::::
Db       47842 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 47801

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||       :::   |||                |||||:::   |||
Db       47800 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 47741

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::         :::         |||       |||||:::   |||:::
Db       47740 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 47681

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::             ||||||  ::::::      :::
Db       47680 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 47621

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||           |||   |||         :::|||   ||||||
Db       47620 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 47576

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||       ||||||   |||
Db       47575 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 47519

Qy         266 -------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db       47518 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 47459

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                         :::::::::|||::::::       |||      ||||||
Db       47458 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 47402

Qy         293 CysGlnTyrThrTrp 297
                  :::||||||:::
Db       47401 ACAGAATATACATTT 47387

RESULT 71
US-09-949-016-17182/c
; Sequence 17182, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17182
;   LENGTH: 192704
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17182

Alignment Scores:
Pred. No.:                183             Length:         192704
Score:                    100.00          Matches:        53
Percent Similarity:       42.7%           Conservative:   43
Best Local Similarity:    23.6%           Mismatches:     95
Query Match:              5.3%            Indels:         35
DB:                       3               Gaps:           8

US-10-757-745-2 (1-362) x US-09-949-016-17182 (1-192704)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::                 :::
Db       48012 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 47954

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||:::||||||         |||
Db       47953 AATTCACACATAACAATATTAACTTTAAATATAAATGGACTA---------AATGCTCCA 47903

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::         :::   :::::::::             |||   |||      :::
Db       47902 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 47843

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||       |||   ::::::
Db       47842 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 47801

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||       :::   |||                |||||:::   |||
Db       47800 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 47741

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::         :::         |||       |||||:::   |||:::
Db       47740 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 47681

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::             ||||||  ::::::      :::
Db       47680 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 47621

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||           |||   |||         :::|||   ||||||
Db       47620 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 47576

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
```

```
Db      47575 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 47519

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db      47518 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 47459

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db      47458 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 47402

Qy        293 CysGlnTyrThrTrp 297

Db      47401 ACAGAATATACATTT 47387

RESULT 72
US-09-949-016-16065/c
; Sequence 16065, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16065
;   LENGTH: 212139
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16065

Alignment Scores:
Pred. No.:              214          Length:         212139
Score:                  100.00       Matches:        55
Percent Similarity:     45.7%        Conservative:   45
Best Local Similarity:  25.1%        Mismatches:     78
Query Match:            5.3%         Indels:         41
DB:                     3            Gaps:           11

US-10-757-745-2 (1-362) x US-09-949-016-16065 (1-212139)

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db     130240 TCACCTGCTAACATCATAATGACAGGATCAAATTCACACATAACAATATTAACTTTAAAT 130181

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db     130180 GTAAATGGA------CTAAATTCTCCAATTAAAAGACACAGA---CTGGCAAATTGGATA 130130

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db     130129 AAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCAT 130088

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db     130087 CTCACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCA 130028

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db     130027 AATGGAAAATTAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGCCTTTAAA 129968

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------ 218

Db     129967 CCAATAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAA 129908

Qy        219 ------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 236

Db     129907 CAAGAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCC 129857

Qy        237 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 256

Db     129856 AGATTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACA 129803

Qy        257 ValIlePheAlaGlyAsp------------------------ThrAsnLeuArgAspAr 268

Db     129802 TTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAG 129746

Qy        268 gGluValThrArgCysGlyGlyLeuProAsn----------------------------As 279

Db     129745 AAAGTCAACAAGG---------ATACCCAGGGAATTGAACTCAGCTCTGAACCAAGCAGA 129695

Qy        279 nIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297

Db     129694 CCTAATAGACATCTACAGAACTCTC---CACCCCAAATCAAAAGAATATACATTT 129643

RESULT 73
US-09-252-991A-14967
; Sequence 14967, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 14967
;   LENGTH: 840
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14967

Alignment Scores:
Pred. No.:              0.0325       Length:         840
Score:                  99.50        Matches:        56
Percent Similarity:     34.9%        Conservative:   41
Best Local Similarity:  20.1%        Mismatches:     122
Query Match:            5.2%         Indels:         59
DB:                     3            Gaps:           10

US-10-757-745-2 (1-362) x US-09-252-991A-14967 (1-840)

Qy        118 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCys 137

Db         25 ACCATGAAAATCGTCTCTTTCAATATCAACGGTCTGCGCGCCCGGCCGCACCAGTTGGCG 84

Qy        138 SerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu------ValIlePro 155

Db         85 GCGCTGATCGAGCGGCACCAGCCCGACGTGATCGGGCTGCAGGAAACCAAGGTCGCCGAC 144

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175

Db        145 GACCAGTTTCCCCAGGCCGAGATCGAGGCGCTCGGCTACCACGTCCACTACCACGGGCAG 204

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195

Db        205 AAAGGCCACTACGGCGTCGCCCTGCTCTCCCGC---CAGGCGCCCCTGGAACTGCATCGC 261

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys---ValHisValAsn 214

Db        262 GGCTTCCCGAGCGACGGCGAGGAATCCCAGCGGCGCTTCATCTGGGGCACCTTCAGCGAC 321
```

```
Qy        215 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu---------------------- 227
Db        322 GCCCAGGGCAACCCGATCACCGTGATGAACGGTTACTTCCCGCAAGGCGAGAACCGCGAG 381

Qy        228 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 247
Db        382 CATCCGACCAAGTTCCCGGCCAAGCAGCGCTTCTACGCCGATCTCCAGTCGCTGCTGGAG 441

Qy        248 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu------ 265
Db        442 AACAGCTTCAAGCCCGAGCAAGCCCTGGTGGTGATG---GGCGACATCAATATCTCGCCG 498

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db        499 GAGGACTGCGACATCGGCATTGGCGAGGAAAACCGCAAGCGCTGGCTGAAGACCGGCAAG 558

Qy        273 CysGlyGlyLeuProAsn--------------------------------AsnIleVal 281
Db        559 TGCAGCTTCCTGCCGGAAGAGCGCGAATGGCTGGCTACGTTGAAGAACTGGGGCCTGGTC 618

Qy        282 AspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
Db        619 GACAGCTTCCGGGCCCTGAATCCGGAGGTGAACGACCGCTTCAGCTGGTTCGACTACCGC 678

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe------ 319
Db        679 AGCCGTGGCTTCGAGGACAACCCCAAGCGCGGCCTGCGCATCGATGTGATCCTGGCCAGC 738

Qy        320 ------------------PheArgAlaAlaAlaGluGluGlyHisIleIleProArgSer 333
Db        739 CGGGTGCTGCAGTCGCGCTTCAAGGACGCCGGCATCGAC---------------------- 777

Qy        334 LeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
Db        778 TACGACCTGCGCGGCATGGAGAAG-----------------CCCTCCGATCAC 813

RESULT 74
US-09-252-991A-14783/c
; Sequence 14783, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 14783
;   LENGTH: 1149
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14783

Alignment Scores:
Pred. No.:              0.0539      Length:          1149
Score:                  99.50       Matches:         56
Percent Similarity:     34.9%       Conservative:    41
Best Local Similarity:  20.1%       Mismatches:      122
Query Match:            5.2%        Indels:          59
DB:                     3           Gaps:            10

US-10-757-745-2 (1-362) x US-09-252-991A-14783 (1-1149)

Qy        118 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCys 137
Db        819 ACCATGAAAATCGTCTCTTTCAATATCAACGGTCTGCGCGCCCGGCCGCACCAGTTGGCG 760

Qy        138 SerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu------ValIlePro 155
Db        759 GCGCTGATCGAGCGGCACCAGCCCGACGTGATCGGGCTGCAGGAAACCAAGGTCGCCGAC 700

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
Db        699 GACCAGTTTCCCCAGGCCGAGATCGAGGCGCTCGGCTACCACGTCCACTACCACGGGCAG 640

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
Db        639 AAAGGCCACTACGGCGTCGCCCTGCTCTCCCGC---CAGGCGCCCCTGGAACTGCATCGC 583

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys---ValHisValAsn 214
Db        582 GGCTTCCCGAGCGACGGCGAGGAATCCCAGCGGCGCTTCATCTGGGGCACCTTCAGCGAC 523

Qy        215 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu---------------------- 227
Db        522 GCCCAGGGCAACCCGATCACCGTGATGAACGGTTACTTCCCGCAAGGCGAGAACCGCGAG 463

Qy        228 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 247
Db        462 CATCCGACCAAGTTCCCGGCCAAGCAGCGCTTCTACGCCGATCTCCAGTCGCTGCTGGAG 403

Qy        248 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu------ 265
Db        402 AACAGCTTCAAGCCCGAGCAAGCCCTGGTGGTGATG---GGCGACATCAATATCTCGCCG 346

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db        345 GAGGACTGCGACATCGGCATTGGCGAGGAAAACCGCAAGCGCTGGCTGAAGACCGGCAAG 286

Qy        273 CysGlyGlyLeuProAsn--------------------------------AsnIleVal 281
Db        285 TGCAGCTTCCTGCCGGAAGAGCGCGAATGGCTGGCTACGTTGAAGAACTGGGGCCTGGTC 226

Qy        282 AspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
Db        225 GACAGCTTCCGGGCCCTGAATCCGGAGGTGAACGACCGCTTCAGCTGGTTCGACTACCGC 166

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe------ 319
Db        165 AGCCGTGGCTTCGAGGACAACCCCAAGCGCGGCCTGCGCATCGATGTGATCCTGGCCAGC 106

Qy        320 ------------------PheArgAlaAlaAlaGluGluGlyHisIleIleProArgSer 333
Db        105 CGGGTGCTGCAGTCGCGCTTCAAGGACGCCGGCATCGAC---------------------- 67

Qy        334 LeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
Db         66 TACGACCTGCGCGGCATGGAGAAG-----------------CCCTCCGATCAC 31

RESULT 75
US-09-252-991A-14833
; Sequence 14833, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 14833
;   LENGTH: 1461
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14833
```

Alignment Scores:

Pred. No.:	0.0794	Length:	1461
Score:	99.50	Matches:	56
Percent Similarity:	34.9%	Conservative:	41
Best Local Similarity:	20.1%	Mismatches:	122
Query Match:	5.2%	Indels:	59
DB:	3	Gaps:	10

US-10-757-745-2 (1-362) x US-09-252-991A-14833 (1-1461)

```
Qy        118 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCys 137
              |||     |||         ::::::|||   |||       |||          :::
Db         68 ACCATGAAAATCGTCTCTTTCAATATCAACGGTCTGCGCGCCCGGCCGCACCAGTTGGCG 127

Qy        138 SerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu------ValIlePro 155
              :::   :::       :::   ||||||||||||   |||||||||       |||
Db        128 GCGCTGATCGAGCGGCACCAGCCCGACGTGATCGGGCTGCAGGAAACCAAGGTCGCCGAC 187

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              :::         :::       :::       |||   :::             :::
Db        188 GACCAGTTTCCCCAGGCCGAGATCGAGGCGCTCGGCTACCACGTCCACTACCACGGGCAG 247

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              :::|||::::::         :::|||   :::   :::       |||:::
Db        248 AAAGGCCACTACGGCGTCGCCCTGCTCTCCCGC---CAGGCGCCCCTGGAACTGCATCGC 304

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys---ValHisValAsn 214
                  |||         :::       |||       :::             :::
Db        305 GGCTTCCCGAGCGACGGCGAGGAATCCCAGCGGCGCTTCATCTGGGGCACCTTCAGCGAC 364

Qy        215 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu---------------------- 227
                |||||   :::   :::|||       :::
Db        365 GCCCAGGGCAACCCGATCACCGTGATGAACGGTTACTTCCCGCAAGGCGAGAACCGCGAG 424

Qy        228 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 247
              |||:::     |||   :::|||                       |||   :::
Db        425 CATCCGACCAAGTTCCCGGCCAAGCAGCGCTTCTACGCCGATCTCCAGTCGCTGCTGGAG 484

Qy        248 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu------ 265
                  ||||||   |||   |||:::       ||||||   |||:::
Db        485 AACAGCTTCAAGCCCGAGCAAGCCCTGGTGGTGATG---GGCGACATCAATATCTCGCCG 541

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   |||       :::       :::
Db        542 GAGGACTGCGACATCGGCATTGGCGAGGAAAACCGCAAGCGCTGGCTGAAGACCGGCAAG 601

Qy        273 CysGlyGlyLeuProAsn--------------------------------AsnIleVal 281
              |||     ||||||                                    :::|||
Db        602 TGCAGCTTCCTGCCGGAAGAGCGCGAATGGCTGGCTACGTTGAAGAACTGGGGCCTGGTC 661

Qy        282 AspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
              |||   :::     |||       :::       ::::::::::|||
Db        662 GACAGCTTCCGGGCCCTGAATCCGGAGGTGAACGACCGCTTCAGCTGGTTCGACTACCGC 721

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe------ 319
              :::                       ||||||   |||   |||
Db        722 AGCCGTGGCTTCGAGGACAACCCCAAGCGCGGCCTGCGCATCGATGTGATCCTGGCCAGC 781

Qy        320 ------------------PheArgAlaAlaAlaGluGluGlyHisIleIleProArgSer 333
                              |||:::   |||       :::
Db        782 CGGGTGCTGCAGTCGCGCTTCAAGGACGCCGGCATCGAC---------------------- 820

Qy        334 LeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
              ||||||   |||:::||||||                 ||||||||||||
Db        821 TACGACCTGCGCGGCATGGAGAAG-----------------CCCTCCGATCAC 856
```

Search completed: March 10, 2006, 23:21:31
Job time : 3716.68 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:38:30 ; Search time 1400.78 Seconds (without alignments)
1722.336 Million cell updates/sec

Title: US-10-757-745-2
Perfect score: 1902
Sequence: 1 MELGSCLEGGREAAEEEGEP.....DCGRFPSDHWGLLCNLDIIL 362

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170325_3993/app_query.fasta_1
-DB=N_Geneseq -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1589_@runat_09032006_170325_3993 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1902	100.0	1920	3	AAZ47118	Aaz47118 Human CD4
2	1902	100.0	1940	14	ADX06356	Adx06356 Cyclin-de
3	1902	100.0	1948	4	AAI58997	Aai58997 Human pol
4	1902	100.0	1948	5	ADQ99219	Adq99219 DNA encod
5	1902	100.0	1948	9	ADB48979	Adb48979 Novel hum
6	1902	100.0	1958	4	AAI60783	Aai60783 Human pol
7	1902	100.0	2499	2	AAX28153	Aax28153 Topoisome
8	1902	100.0	3152	9	ADA10970	Ada10970 Human cDN
9	1897	99.7	1936	10	ADD19013	Add19013 Human dis
10	1897	99.7	1936	13	ADP25361	Adp25361 PRO polyp
11	1894	99.6	1296	3	AAC98160	Aac98160 Human col
12	1893	99.5	1898	4	AAH15146	Aah15146 Human cDN
13	1258	66.1	1312	3	AAZ47119	Aaz47119 Mouse CD4
14	1199	63.0	1088	5	AAS86254	Aas86254 DNA encod
15	1182	62.1	752	4	AAH08073	Aah08073 Human cDN
c 16	1177	61.9	1079	2	AAX84209	Aax84209 DNA encod
c 17	1177	61.9	1079	3	AAC79438	Aac79438 cDNA sequ
c 18	1177	61.9	1079	6	ABK28982	Abk28982 Human bre
19	1119	58.8	674	11	ADT95548	Adt95548 Colon can
20	1119	58.8	674	11	ADX42030	Adx42030 Human cDN
21	1069	56.2	644	11	ADT95551	Adt95551 Colon can
22	1069	56.2	644	11	ADX42033	Adx42033 Human cDN
23	1068	56.2	1227	5	AAS86255	Aas86255 DNA encod
24	1054	55.4	633	11	ADT95565	Adt95565 Colon can
25	1054	55.4	633	11	ADX42047	Adx42047 Human cDN
26	1031	54.2	625	11	ADT95029	Adt95029 Colon can
27	1031	54.2	625	11	ADX41511	Adx41511 Human cDN
28	1020	53.6	625	11	ADT95512	Adt95512 Colon can
29	1020	53.6	625	11	ADX41994	Adx41994 Human cDN
30	902	47.4	553	11	ADT95464	Adt95464 Colon can
31	902	47.4	553	11	ADX41946	Adx41946 Human cDN
c 32	857	45.1	774	10	ADD34375	Add34375 Mouse mit
33	837	44.0	483	2	AAX40590	Aax40590 Human sec
34	714.5	37.6	602	10	ADD34376	Add34376 Mouse mit
35	651	34.2	401	11	ADT95844	Adt95844 Colon can
36	651	34.2	401	11	ADX42326	Adx42326 Human cDN
c 37	510	26.8	391	11	ADT95307	Adt95307 Colon can
c 38	510	26.8	391	11	ADX41789	Adx41789 Human cDN
39	376	19.8	258	2	AAX41008	Aax41008 Human sec
40	298	15.7	176	6	ABK27684	Abk27684 Human col
c 41	298	15.7	176	6	ABV95959	Abv95959 Human pan
c 42	298	15.7	179	6	ABL36461	Abl36461 Human col
43	267	14.0	725	6	ABQ34462	Abq34462 Oligonucl
c 44	267	14.0	725	6	ABQ34463	Abq34463 Oligonucl
45	231	12.1	1566	3	AAC47239	Aac47239 Arabidops
46	231	12.1	1606	3	AAC41089	Aac41089 Arabidops
c 47	221	11.6	268	6	ABT07150	Abt07150 Human ova
c 48	221	11.6	268	8	ABX73028	Abx73028 Human ova
49	184.5	9.7	725	6	ABQ34465	Abq34465 Oligonucl
c 50	184.5	9.7	725	6	ABQ34464	Abq34464 Oligonucl
51	182.5	9.6	444	4	AAI15311	Aai15311 Probe #52
52	182.5	9.6	691	4	AAI24484	Aai24484 Probe #14
53	169	8.9	267	5	AAS69905	Aas69905 DNA encod
54	158.5	8.3	525	13	ACN59762	Acn59762 Cotton gy
55	147	7.7	9737	14	ADZ62001	Adz62001 Murine Cc
56	140	7.4	2928	12	ADQ97808	Adq97808 Mouse can
57	128	6.7	1767	12	ADE43923	Ade43923 Human DP1
58	128	6.7	1769	12	ADQ97811	Adq97811 Human can
59	128	6.7	1971	14	ADZ49510	Adz49510 Insulin s
60	127.5	6.7	1350	4	ABL08315	Abl08315 Drosophil
61	126	6.6	47040	12	ADQ97807	Adq97807 Mouse can
c 62	126	6.6	90043	11	ACN44608	Acn44608 Mouse gen
63	123	6.5	49355	12	ADQ97810	Adq97810 Human can
64	120	6.3	1644	13	ADS45761	Ads45761 Bacterial
65	117.5	6.2	5299	8	AAD49594	Aad49594 Human cyt
c 66	117	6.2	96960	8	ACF62734	Acf62734 Cancer ba
c 67	117	6.2	96960	8	ADB20849	Adb20849 MRP1 base
c 68	117	6.2	96960	10	ADB87938	Adb87938 Human UGT
c 69	117	6.2	96960	10	ADB96921	Adb96921 Human MDR
c 70	117	6.2	96960	10	ADB92112	Adb92112 Human MDR
71	116	6.1	5496	13	ADR46193	Adr46193 Retrotran
72	114	6.0	5496	13	ADR46200	Adr46200 Retrotran
73	112	5.9	5322	8	AAD49592	Aad49592 Human cyt
74	111	5.8	1602	8	ACA01197	Aca01197 C. glutam
75	111	5.8	1608	5	AAH67526	Aah67526 C glutami

ALIGNMENTS

```
RESULT 1
AAZ47118
ID   AAZ47118 standard; cDNA; 1920 BP.
XX
AC   AAZ47118;
XX
DT   15-MAR-2000  (first entry)
XX
DE   Human CD40 receptor associated protein gene.
XX
KW   Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW   immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW   human; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW   tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW   TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW   diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW   arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW   graft versus host disease; autoimmune disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO9955859-A2.
XX
PD   04-NOV-1999.
XX
PF   28-APR-1999;   99WO-EP003025.
XX
PR   29-APR-1998;   98EP-00201392.
XX
PA   (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI   Pype SMC,  Remacle JEFJG,  Huylebroeck DFE;
XX
DR   WPI; 2000-062029/05.
DR   P-PSDB; AAY56019.
XX
PT   Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT   diseases and for improvement of anti-tumor treatments.
XX
PS   Claim 9; Page 37-39; 48pp; English.
XX
CC   This sequence represents the gene encoding human CD40 receptor associated
CC   protein (CRAP). CRAP is a functional protein capable of interacting with
CC   the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC   necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC   I, where the protein has no homology to TNF receptor associated factor
CC   (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC   the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC   related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC   diseases, and for the improvement of anti-tumour diseases. Diseases which
CC   may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC   systemic lupus erythematosus, graft rejection, graft versus host disease,
CC   allergy, and autoimmune disease. The proteins can be used to sensitize
CC   tumour cells to anti-tumour treatments and to screen for compounds which
CC   interfere with the interaction of the proteins with other protein
CC   components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ   Sequence 1920 BP; 599 A; 327 C; 435 G; 557 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:                2.62e-199       Length:          1920
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2 (1-362) x AAZ47118 (1-1920)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 79

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 139

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 199

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 259

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 319

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 379

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1099

Qy        361 IleLeu 362
              ||||||
Db       1100 ATATTG 1105
```

```
RESULT 2
ADX06356
ID   ADX06356 standard; DNA; 1940 BP.
XX
AC   ADX06356;
XX
DT   21-APR-2005  (first entry)
XX
DE   Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 921.
XX
KW   cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2005012875-A2.
XX
PD   10-FEB-2005.
XX
PF   29-JUL-2004; 2004WO-US024424.
XX
PR   29-JUL-2003; 2003US-0490890P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Li M,  Rupnow BA,  Webster KR,  Jackson DG,  Wong TW;
XX
DR   WPI; 2005-163068/17.
DR   P-PSDB; ADX06357.
XX
PT   Biomarkers useful for predicting or determining the response of a mammal
PT   to a cancer treatment comprising administration of a modulator of cyclin-
PT   dependent kinase activity.
XX
PS   Claim 5; SEQ ID NO 921; 141pp; English.
XX
CC   This invention describes a novel method of predicting or determining
CC   whether a mammal will respond or is responding to an anti-cancer agent
CC   that modulates cyclin-dependent kinase (cdk) activity. The method
CC   comprises measuring the level of one or more biomarkers selected from
CC   2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC   NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC   invention is utilized in a kit for determining or predicting whether
CC   patient would be susceptible or resistant to treatment by an agent
CC   modulating cdk activity. The invention also describes a method for
CC   utilizing individualized genetic profiles for treating diseases and
CC   disorders based on patient's response and molecular level, specialized
CC   microarrays comprising the biomarkers described, antibodies directed
CC   against the biomarkers and a cell culture model to identify biomarkers.
CC   The cdk modulator is preferably N-5-[[5-(1,1-Dimethylethyl)-2-
CC   oxazolyl]methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC   tartaric acid salt. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC   sequence encodes a biomarker used in the method of the invention.
XX
SQ   Sequence 1940 BP; 613 A; 330 C; 439 G; 558 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.65e-199       Length:         1940
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2 (1-362) x ADX06356 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
              ||||||
Db       1106 ATATTG 1111

RESULT 3
AAI58997
```

```
ID   AAI58997 standard; cDNA; 1948 BP.
XX
AC   AAI58997;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 1200.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J,  Zhao QA;
PI   Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39841.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders such
PT   as central nervous system injuries.
XX
PS   Claim 1; SEQ ID NO 1200; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC   encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders. Note: The sequence data for this patent did not form
CC   part of the printed specification
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;
```

Alignment Scores:			
Pred. No.:	2.67e-199	Length:	1948
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2 (1-362) x AAI58997 (1-1948)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuGluGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118
```

```
Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124

RESULT 4
ADQ99219
ID   ADQ99219 standard; cDNA; 1948 BP.
XX
AC   ADQ99219;
XX
DT   23-SEP-2004  (first entry)
XX
DE   DNA encoding human GPCR-like protein seqid 889.
XX
KW   ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
KW   antidiabetic; GPCR-like protein; ophthalmic disorder;
KW   neurological disorder; immunological disorder; nephritic disorder;
KW   hormonal dysfunction; cancer; atherosclerosis; diabetes;
KW   molecular weight marker; food supplement; human; ss.
XX
OS   Homo sapiens.
XX
PN   US6569662-B1.
XX
PD   27-MAY-2003.
XX
PF   19-JUL-2000; 2000US-00620312.
XX
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Zhou P,  Drmanac RT;
XX
DR   WPI; 2001-442255/47.
XX
PT   New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT   useful for treating diseases of ophthalmic, neurological, immunological
PT   and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT   and diabetes.
XX
PS   Example 2; SEQ ID NO 889; 92pp; English.
XX
CC   The invention describes an isolated polynucleotide (I) comprising a fully
CC   defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC   1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC   given in the specification, its translated or protein coding portion, its
CC   extracellular portion or its active domain. The GPCR-like polypeptides
CC   and polynucleotides are useful for the treatment of diseases of
CC   ophthalmic, neurological, immunological and nephritic systems. They may
CC   also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC   diabetes. The antibodies are useful for detecting or quantitating the
CC   polypeptide in tissue. The polypeptides can also be used as molecular
CC   weight markers and as a food supplement. This sequence represents a human
CC   polynucleotide of the invention.
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.67e-199      Length:          1948
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       5              Gaps:            0

US-10-757-745-2 (1-362) x ADQ99219 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124
```

```
RESULT 5
ADB48979
ID   ADB48979 standard; cDNA; 1948 BP.
XX
AC   ADB48979;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Novel human cDNA SEQ ID NO 889.
XX
KW   ss; cancer; neurodegenerative disease; human.
XX
OS   Homo sapiens.
XX
PN   US2003104529-A1.
XX
PD   05-JUN-2003.
XX
PF   04-JAN-2002; 2002US-00037270.
XX
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   19-JUL-2000; 2000US-00620312.
XX
PA   (ZHOU/) ZHOU P.
PA   (TANG/) TANG Y T.
PA   (LIUC/) LIU C.
PA   (ASUN/) ASUNDI V.
PA   (DRMA/) DRMANAC R T.
XX
PI   Zhou P,  Tang YT,  Liu C,  Asundi V,  Drmanac RT;
XX
DR   WPI; 2003-678194/64.
XX
PT   New polynucleotide, useful for treating diseases e.g., cancer or
PT   neurodegenerative diseases.
XX
PS   Claim 1; SEQ ID NO 889; 99pp; English.
XX
CC   The invention relates to a polynucleotide comprising a sequence given in
CC   the specification, or its mature protein-coding portion, or its
CC   complement. The polynucleotide is useful for treating diseases e.g.,
CC   cancer or neurodegenerative diseases and many others listed in the
CC   specification. The present sequence represents a novel human cDNA. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=20030104529.
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.67e-199       Length:          1948
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2 (1-362) x ADB48979 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124

RESULT 6
AAI60783
ID   AAI60783 standard; cDNA; 1958 BP.
XX
AC   AAI60783;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4772.
XX
```

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41627.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 4772; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1958 BP; 613 A; 339 C; 442 G; 564 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.69e-199	Length:	1958
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2 (1-362) x AAI60783 (1-1958)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 110

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 170

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 230

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 290

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 350

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 410

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1130

Qy        361 IleLeu 362
              ||||||
Db       1131 ATATTG 1136
```

RESULT 7
AAX28153
ID AAX28153 standard; DNA; 2499 BP.

```
XX
AC   AAX28153;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Topoisomerase II binding protein 1 coding sequence.
XX
KW   Topoisomerase II binding protein; TopBP; anticancer agent; ds.
XX
OS   Homo sapiens.
XX
PN   JP11075856-A.
XX
PD   23-MAR-1999.
XX
PF   17-SEP-1997;   97JP-00251544.
XX
PR   17-SEP-1997;   97JP-00251544.
XX
PA   (TSUR/) TSURUO T.
PA   (CHUS ) CHUGAI PHARM CO LTD.
XX
DR   WPI; 1999-257704/22.
DR   P-PSDB; AAY03182.
XX
PT   New Topoisomerase II- binding protein - useful as an anticancer agent.
XX
PS   Disclosure; Page 18-19; 28pp; Japanese.
XX
CC   This sequence encodes the topoisomerase II binding protein (TopBP) of the
CC   invention. The TopBP protein is useful as an anticancer agent. TopBP can
CC   be used as the target molecule for anticancer agent
XX
SQ   Sequence 2499 BP; 720 A; 505 C; 587 G; 687 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.84e-199       Length:          2499
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       2               Gaps:            0

US-10-757-745-2 (1-362) x AAX28153 (1-2499)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 541

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 601

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 661

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 721

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 781

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 841

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 901

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 961

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1021

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1081

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1141

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1201

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1261

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1321

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1381

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1441

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1501

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1561

Qy        361 IleLeu 362
              ||||||
Db       1562 ATATTG 1567

RESULT 8
ADA10970
ID   ADA10970 standard; cDNA; 3152 BP.
XX
AC   ADA10970;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human cDNA differentially expressed in colon cancer #60.
XX
KW   ss; differential expression; colon cancer; cancer; human.
XX
OS   Homo sapiens.
XX
PN   US2002160382-A1.
XX
PD   31-OCT-2002.
XX
PF   11-OCT-2001; 2001US-00981353.
XX
PR   11-OCT-2000; 2000US-0239841P.
XX
PA   (LASE/) LASEK A W.
PA   (JONE/) JONES D A.
XX
```

```
PI   Lasek AW,  Jones DA;
XX
DR   WPI; 2003-265756/26.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   colon disorder, useful for diagnosing, treating, staging or monitoring
PT   treatment for colon cancers.
XX
PS   Claim 1; SEQ ID NO 88; 231pp; English.
XX
CC   The invention relates to a combination comprising cDNAs that are
CC   differentially expressed in colon disorder. The methods and compositions
CC   of the present invention are useful for diagnosing, treating, staging or
CC   monitoring treatment for colon cancer. They are also useful in high
CC   throughput methods for using cDNAs to detect differential expression of
CC   nucleic acids in a sample, screening molecules or compounds to identify a
CC   ligand which specifically binds a cDNA and using a protein to screen
CC   molecules or compounds to identify at least one ligand which specifically
CC   binds the protein. The present sequence represents a human cDNA
CC   differentially expressed in colon cancer.
XX
SQ   Sequence 3152 BP; 875 A; 643 C; 707 G; 926 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                5.38e-199       Length:          3152
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2 (1-362) x ADA10970 (1-3152)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 1311

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 1371

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 1431

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 1491

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1492 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 1551

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 1611

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 1671

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1672 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 1731

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1732 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1791

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1792 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1851

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1852 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1911

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1912 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1971

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1972 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 2031

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2032 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 2091

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2092 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 2151

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2152 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 2211

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2212 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 2271

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2272 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 2331

Qy        361 IleLeu 362
              ||||||
Db       2332 ATATTG 2337

RESULT 9
ADD19013
ID   ADD19013 standard; DNA; 1936 BP.
XX
AC   ADD19013;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human disease related protein DNA sequence SeqID502.
XX
KW   human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW   antiarteriosclerotic; vulnerary; gene therapy;
KW   hypoxia-regulated condition; tumourigenesis; angiogenesis; apoptosis;
KW   inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW   glucose transportation; catecholamine synthesis; iron transport;
KW   nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW   retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW   inflammatory condition; wound healing; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003018621-A2.
XX
PD   06-MAR-2003.
XX
PF   23-AUG-2002; 2002WO-GB003892.
XX
PR   23-AUG-2001; 2001GB-00020558.
PR   05-OCT-2001; 2001GB-00024037.
XX
PA   (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI   Kingsman SM,  White J,  Ward NR,  Harris RA,  Naylor S,  Mundy CR;
XX
DR   WPI; 2003-290046/28.
DR   P-PSDB; ADD19012.
XX
```

```
PT   New substantially purified polypeptide, useful for diagnosing or treating
PT   a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT   injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT   wound healing.
XX
PS   Claim 27; SEQ ID NO 502; 424pp; English.
XX
CC   This invention relates to novel human genes and gene product which are
CC   implicated in certain disease states. Compounds which modulate the
CC   proteins of the invention may have cytostatic, antiinflammatory,
CC   ophthalmological, antiarteriosclerotic or vulnerary activities. The
CC   sequences of the invention may be useful for gene therapy. The invention
CC   may be useful for diagnosing or treating a hypoxia-regulated condition,
CC   such as tumourigenesis, angiogenesis, apoptosis, inflammation,
CC   erythropoiesis, or the biological response to hypoxia conditions
CC   including processes such as glycolysis, gluconeogenesis, glucose
CC   transportation, catecholamine synthesis, iron transport or nitric oxide
CC   synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC   injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC   inflammatory conditions or wound healing. The present sequence is that of
CC   a disease related protein encoding DNA sequence of the invention.
XX
SQ   Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                  9.43e-199       Length:          1936
Score:                      1897.00         Matches:         361
Percent Similarity:         99.7%           Conservative:    0
Best Local Similarity:      99.7%           Mismatches:      1
Query Match:                99.7%           Indels:          0
DB:                         10              Gaps:            0

US-10-757-745-2 (1-362) x ADD19013 (1-1936)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 76

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db         77 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGAGGTTTGCCTCGGTCGCAAGCTGCGATGCC 136

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 196

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 256

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 316

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 376

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 436

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 496

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 556

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 616

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 676

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 736

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 796

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 856

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 916

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 976

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1036

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1096

Qy        361 IleLeu 362
              ||||||
Db       1097 ATATTG 1102

RESULT 10
ADP25361
ID   ADP25361 standard; cDNA; 1936 BP.
XX
AC   ADP25361;
XX
DT   18-NOV-2004  (first entry)
XX
DE   PRO polypeptide encoding cDNA SEQ ID NO:475.
XX
KW   ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
KW   immunosuppressive; osteopathic; antidiabetic; dermatological;
KW   antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW   gene therapy; immune system.
XX
OS   Unidentified.
XX
PN   WO2004041170-A2.
XX
PD   21-MAY-2004.
XX
PF   30-OCT-2003; 2003WO-US034312.
XX
PR   01-NOV-2002; 2002US-0423394P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Clark H,  Schoenfeld J,  Van Lookeren M,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-419628/39.
DR   P-PSDB; ADP25362.
XX
PT   New PRO polypeptides and polynucleotides, useful for treating e.g.
PT   erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT   renal disease, or demyelinating diseases of the central or peripheral
```

```
PT   nervous system.
XX
PS   Claim 1; SEQ ID NO 475; 2940pp; English.
XX
CC   The invention relates to a novel isolated nucleic acid and the PRO
CC   polypeptide encoded by it. A protein of the invention has
CC   antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC   osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC   antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC   of the invention may have a use in gene therapy. The PRO polypeptide, its
CC   agonist, antagonist, or antibody that specifically binds to the
CC   polypeptide is useful for treating an immune related disorder such as
CC   systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC   juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC   idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC   vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC   thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC   disease, a demyelinating disease of the central or peripheral nervous
CC   system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC   a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC   disease, infectious or autoimmune chronic active hepatitis, primary
CC   biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC   inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC   disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC   disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC   disease, asthma, allergic rhinitis, atopic dermatitis, food
CC   hypersensitivity, urticaria, an immunologic disease of the lung,
CC   eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC   pneumonitis, a transplantation associated disease, graft rejection or
CC   graft-versus-host disease. The present sequence encodes a PRO protein of
CC   the invention.
XX
SQ   Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.43e-199      Length:          1936
Score:                    1897.00        Matches:         361
Percent Similarity:       99.7%          Conservative:    0
Best Local Similarity:    99.7%          Mismatches:      1
Query Match:              99.7%          Indels:          0
DB:                       13             Gaps:            0

US-10-757-745-2 (1-362) x ADP25361 (1-1936)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 76

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db         77 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGAGGTTTGCCTCGGTCGCAAGCTGCGATGCC 136

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 196

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 256

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 316

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 376

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 436

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 496

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 556

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 616

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 676

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 736

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 796

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 856

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 916

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 976

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1036

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1096

Qy        361 IleLeu 362
              ||||||
Db       1097 ATATTG 1102

RESULT 11
AAC98160
ID   AAC98160 standard; cDNA; 1296 BP.
XX
AC   AAC98160;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human colon cancer antigen nucleotide sequence SEQ ID NO:170.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW   immunomodulatory; muscular; gynaecological; gastrointestinal;
KW   nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW   neural disorder; immune system disorder; muscular disorder;
KW   reproductive disorder; gastrointestinal disorder; renal disorder;
KW   infectious disease; cardiovascular disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055351-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005883.
XX
```

```
PR   12-MAR-1999;   99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587534/55.
DR   P-PSDB; AAB53403.
XX
PT   Colon cancer associated gene sequences, referred to as colon cancer
PT   antigens, useful for the treatment, prevention, and diagnosis of colon
PT   disorders such as colon cancer.
XX
PS   Claim 1; Page 597; 2104pp; English.
XX
CC   AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC   called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC   human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC   neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC   vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC   can be used in gene therapy. The colon cancer antigen polynucleotides,
CC   proteins and antibodies to the proteins are useful for the prevention,
CC   treatment and diagnosis of colon disorders, such as colon cancer. The
CC   polynucleotides may be used in diagnostics and research, such as for
CC   chromosome identification, and as hybridisation probes. The proteins may
CC   also be used to prevent diseases such as neural disorders, immune system
CC   disorders, muscular disorders, reproductive disorders, gastrointestinal
CC   disorders, wounds, renal disorders, infectious diseases, and
CC   cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC   sequences used in the exemplification of the present invention
XX
SQ   Sequence 1296 BP; 376 A; 250 C; 326 G; 333 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.:              1.13e-198       Length:         1296
Score:                  1894.00         Matches:        361
Percent Similarity:     99.7%           Conservative:   0
Best Local Similarity:  99.7%           Mismatches:     1
Query Match:            99.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2 (1-362) x AAC98160 (1-1296)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 83

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 143

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 203

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 263

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 323

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 383

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 443

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 503

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACA 563

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 623

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTT 683

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 743

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 803

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 863

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 GTGGATGTCTGGGAGTTTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAA 923

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 983

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        984 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1043

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1103

Qy        361 IleLeu 362
              ||||||
Db       1104 ATATTG 1109

RESULT 12
AAH15146
ID   AAH15146 standard; cDNA; 1898 BP.
XX
AC   AAH15146;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:13209.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
```

```
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 8; SEQ ID NO 13209; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 1898 BP; 590 A; 327 C; 426 G; 555 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.53e-198      Length:          1898
Score:                    1893.00        Matches:         361
Percent Similarity:       99.7%          Conservative:    0
Best Local Similarity:    99.7%          Mismatches:      1
Query Match:              99.5%          Indels:          0
DB:                       4              Gaps:            0

US-10-757-745-2 (1-362) x AAH15146 (1-1898)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1080

Qy        361 IleLeu 362
              ||||||
Db       1081 ATATTG 1086

RESULT 13
AAZ47119
ID   AAZ47119 standard; cDNA; 1312 BP.
XX
AC   AAZ47119;
XX
DT   15-MAR-2000  (first entry)
XX
DE   Mouse CD40 receptor associated protein gene.
XX
KW   Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW   immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW   mouse; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW   tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW   TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW   diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW   arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW   graft versus host disease; autoimmune disease; ds.
XX
OS   Mus musculus.
XX
PN   WO9955859-A2.
```

```
XX
PD   04-NOV-1999.
XX
PF   28-APR-1999;   99WO-EP003025.
XX
PR   29-APR-1998;   98EP-00201392.
XX
PA   (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI   Pype SMC,  Remacle JEFJG,  Huylebroeck DFE;
XX
DR   WPI; 2000-062029/05.
DR   P-PSDB; AAY56020.
XX
PT   Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT   diseases and for improvement of anti-tumor treatments.
XX
PS   Claim 10; Page 41-43; 48pp; English.
XX
CC   This sequence represents the gene encoding mouse CD40 receptor associated
CC   protein (CRAP). CRAP is a functional protein capable of interacting with
CC   the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC   necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC   I, where the protein has no homology to TNF receptor associated factor
CC   (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC   the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC   related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC   diseases, and for the improvement of anti-tumour diseases. Diseases which
CC   may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC   systemic lupus erythematosus, graft rejection, graft versus host disease,
CC   allergy, and autoimmune disease. The proteins can be used to sensitize
CC   tumour cells to anti-tumour treatments and to screen for compounds which
CC   interfere with the interaction of the proteins with other protein
CC   components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ   Sequence 1312 BP; 359 A; 279 C; 352 G; 322 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.87e-128      Length:          1312
Score:                    1258.00        Matches:         240
Percent Similarity:       80.6%          Conservative:    43
Best Local Similarity:    68.4%          Mismatches:      64
Query Match:              66.1%          Indels:          4
DB:                       3              Gaps:            3

US-10-757-745-2 (1-362) x AAZ47119 (1-1312)

Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31
              |||||||||:::   |||   |||      |||||||:::|||||||||   |||:::
Db        191 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 247

Qy         32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51
              ||||||   |||      ||||||      :::              |||   |||||||||
Db        248 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 307

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
              |||:::   ::::::||||||::::::|||||||||   |||       :::
Db        308 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 367

Qy         72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91
              |||:::|||   |||       :::   :::   |||||||||||||||||||:::
Db        368 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 427

Qy         92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111
              |||:::|||      :::   |||||||||      |||       |||:::
Db        428 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 484

Qy        112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131
              |||       |||   ||||||||||||||||||||||||||||       ||||||   |||
Db        485 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 544

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
              ||||||||||||||||||||   ||||||||||||||||||||||:::||||||||||

Db        545 AGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 604

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
              ||||||||||||||||||   :::|||||||||||||:::::::::|||   ||||||
Db        605 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 664

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191
              ||||||:::|||||||||||||||||||||||||:::|||||||||   |||||||||
Db        665 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTT 724

Qy        192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        725 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 784

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
              :::|||::::::   |||||||||   ||||||||||||||||||||||||||||||||
Db        785 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 844

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
              |||:::|||||||||:::   |||||||||   ||||||   ||||||||||||||||
Db        845 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 904

Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271
              |||:::|||   ||||||||||||||||||||||||||||||||||||:::||||||
Db        905 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 964

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
              :::|||||||||||||||:::|||:::   |||   ||||||||||||||||||||||
Db        965 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 1024

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
              ||||||||||||||||||||||||||:::   |||:::||||||   |||   ||||||
Db       1025 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1084

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              |||   ||||||||||||||||||||||||      ||||||||||||||:::||||||
Db       1085 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1138

Qy        332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
              :::|||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db       1139 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1198

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||||   |||:::::::::|||
Db       1199 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1231

RESULT 14
AAS86254
ID   AAS86254 standard; cDNA; 1088 BP.
XX
AC   AAS86254;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #22058.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
```

```
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG22067.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 22058; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1088 BP; 345 A; 182 C; 236 G; 325 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.64e-122      Length:          1088
Score:                    1199.00        Matches:         226
Percent Similarity:       99.1%          Conservative:    1
Best Local Similarity:    98.7%          Mismatches:      2
Query Match:              63.0%          Indels:          0
DB:                       5              Gaps:            0

US-10-757-745-2 (1-362) x AAS86254 (1-1088)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy        333 SerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrp 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 660

Qy        353 GlyLeuLeuCysAsnLeuAspIleIle 361
              |||||||||||||||||||||||||||
Db        661 GGTCTTCTGTGCAACTTAGATATAATA 687

RESULT 15
AAH08073
ID   AAH08073 standard; cDNA; 752 BP.
XX
AC   AAH08073;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:4908.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 1; SEQ ID NO 4908; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
```

```
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 752 BP; 225 A; 148 C; 200 G; 176 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:                2.04e-120       Length:          752
Score:                    1182.00         Matches:         238
Percent Similarity:       94.8%           Conservative:    1
Best Local Similarity:    94.4%           Mismatches:      10
Query Match:              62.1%           Indels:          3
DB:                       4               Gaps:            0

US-10-757-745-2 (1-362) x AAH08073 (1-752)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArg-AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 660

Qy        220 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 240
              |||||||  ||||||  |||    :::|||  ||||||      ||||||||||||||
Db        661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy        240 lnLeuLysMetValLeu-LysLysMetGln 249
              ||   ||||||   |||    |||||||||
Db        721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750
```

```
RESULT 16
AAX84209/c
ID   AAX84209 standard; cDNA; 1079 BP.
XX
AC   AAX84209;
XX
DT   08-SEP-1999  (first entry)
XX
DE   DNA encoding human breast tumour protein immunogenic fragment.
XX
KW   Breast tumour protein; immunogenic fragment; vaccine; detection;
KW   breast cancer development; therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9933869-A2.
XX
PD   08-JUL-1999.
XX
PF   22-DEC-1998;   98WO-US027416.
XX
PR   24-DEC-1997;   97US-00998253.
PR   24-DEC-1997;   97US-00998255.
PR   17-JUL-1998;   98US-00118554.
PR   17-JUL-1998;   98US-00118627.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J;
XX
DR   WPI; 1999-405486/34.
XX
PT   New breast tumor protein genes used, in vaccines for immunotherapy, or
PT   for diagnosis of breast cancer.
XX
PS   Claim 3; Page 60; 70pp; English.
XX
CC   This sequence encodes a human breast tumour protein immunogenic fragment
CC   of the invention. The polypeptides or nucleic acids encoding them are
CC   useful in vaccines and pharmaceutical compositions for manufacture of
CC   medicaments for inhibiting the development of breast cancer in a patient.
CC   They can also be used to treat breast cancer. Antibodies against these
CC   polypeptides can be used to detect and monitor progression of breast
CC   cancer in patients. Primers and probes derived from the polynucleotides
CC   encoding the breast proteins are useful for detection of breast cancer.
CC   Peripheral blood cells from a patient incubated in the presence of at
CC   least one polypeptide, such that T cells proliferate, are useful in
CC   manufacture of a medicament for treating breast cancer in a patient.
CC   Antigen presenting cells incubated in the presence of at least one
CC   polypeptide are also useful for treating breast cancer
XX
SQ   Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.23e-119       Length:          1079
Score:                    1177.00         Matches:         226
Percent Similarity:       99.6%           Conservative:    0
Best Local Similarity:    99.6%           Mismatches:      1
Query Match:              61.9%           Indels:          1
DB:                       2               Gaps:            0

US-10-757-745-2 (1-362) x AAX84209 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
                |||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398

RESULT 17
AAC79438/c
ID   AAC79438 standard; cDNA; 1079 BP.
XX
AC   AAC79438;
XX
DT   07-FEB-2001  (first entry)
XX
DE   cDNA sequence of human breast tumour clone 1015D11.
XX
KW   Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;
KW   vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061756-A2.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009688.
XX
PR   09-APR-1999;   99US-00288950.
PR   02-JUL-1999;   99US-00346327.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J,  Dillon DC;
XX
DR   WPI; 2000-638568/61.
XX
PT   A novel isolated polypeptide comprising an immunogenic portion of a
PT   breast cancer protein useful in the detection and treatment of breast
PT   cancer.
XX
PS   Claim 4; Page 77; 95pp; English.
XX
CC   The present sequence was isolated from a breast tumour cDNA library. It
CC   is provided in a specification relating to compounds for immunotherapy
CC   and diagnosis of breast cancer. Breast tumour antigens and the
CC   polynucleotides that encode them may be used in the production of a
CC   pharmaceutical composition to be used in the treatment of breast cancer.
CC   Proliferated T cells and incubated antigen presenting cells are also
CC   required. The polypeptides and polynucleotides may also be used to
CC   produce a vaccine
XX
SQ   Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.23e-119      Length:          1079
Score:                    1177.00        Matches:         226
Percent Similarity:       99.6%          Conservative:    0
Best Local Similarity:    99.6%          Mismatches:      1
Query Match:              61.9%          Indels:          1
DB:                       3              Gaps:            0

US-10-757-745-2 (1-362) x AAC79438 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
                |||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398

RESULT 18
ABK28982/c
```

```
ID   ABK28982 standard; cDNA; 1079 BP.
XX
AC   ABK28982;
XX
DT   23-APR-2002  (first entry)
XX
DE   Human breast tumour polypeptide cDNA clone #11.
XX
KW   Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
KW   immunostimulant.
XX
OS   Homo sapiens.
XX
PN   WO200198339-A2.
XX
PD   27-DEC-2001.
XX
PF   12-JUN-2001; 2001WO-US019032.
XX
PR   22-JUN-2000; 2000US-00602877.
PR   12-OCT-2000; 2000US-00687507.
PR   06-FEB-2001; 2001US-00778381.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J,  Dillon DC,  Retter MW,  Harlocker SL;
XX
DR   WPI; 2002-147792/19.
XX
PT   Polynucleotides encoding breast tumor polypeptides, useful for treating
PT   breast cancer or stimulating an immune response.
XX
PS   Claim 1; Page 131; 150pp; English.
XX
CC   The invention relates to polynucleotides encoding breast tumour
CC   polypeptides. The sequences are useful for treating cancer, preferably
CC   breast cancer, in a patient or for stimulating an immune response. The
CC   polynucleotides and polypeptides are also useful in the diagnosis and
CC   monitoring of breast cancer. A method for detecting the presence of a
CC   cancer in a patient, comprises obtaining a biological sample from the
CC   patient, contacting the biological sample with a binding agent that binds
CC   to a breast tumour polypeptide, detecting in the sample an amount of
CC   polypeptide that binds to the binding agent, and comparing the amount of
CC   polypeptide to a predetermined cut-off value, therefore determining the
CC   presence of a cancer in the patient. Sequences ABK28920-ABK29025
CC   represent cDNA clones encoding human breast tumour polypeptides of the
CC   invention
XX
SQ   Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.23e-119       Length:          1079
Score:                    1177.00         Matches:         226
Percent Similarity:       99.6%           Conservative:    0
Best Local Similarity:    99.6%           Mismatches:      1
Query Match:              61.9%           Indels:          1
DB:                       6               Gaps:            0

US-10-757-745-2 (1-362) x ABK28982 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398

RESULT 19
ADT95548
ID   ADT95548 standard; cDNA; 674 BP.
XX
AC   ADT95548;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1067.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
```

```
XX
PS   Claim 1; SEQ ID NO 1067; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.:              1.57e-113      Length:          674
Score:                  1119.00        Matches:         216
Percent Similarity:     98.6%          Conservative:    0
Best Local Similarity:  98.6%          Mismatches:      2
Query Match:            58.8%          Indels:          1
DB:                     11             Gaps:            0

US-10-757-745-2 (1-362) x ADT95548 (1-674)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 77

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        253 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        272 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy        292 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

Qy        312 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProAr 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy        332 gSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAsp 350
              |||||||||||||||||||||    ||||||||   ||||||||||||||||||
Db        618 AAGTTTGGACCTTCTTGGATTANAAAAACTGGACTGNGGTAGATTTCCTAGTGAT 672

RESULT 20
ADX42030
ID   ADX42030 standard; cDNA; 674 BP.
XX
AC   ADX42030;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1067.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1067; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.:              1.57e-113      Length:          674
Score:                  1119.00        Matches:         216
Percent Similarity:     98.6%          Conservative:    0
Best Local Similarity:  98.6%          Mismatches:      2
Query Match:            58.8%          Indels:          1
DB:                     11             Gaps:            0

US-10-757-745-2 (1-362) x ADX42030 (1-674)
```

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 77

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        253 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        272 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy        292 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

Qy        312 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProAr 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy        332 gSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAsp 350
              ||||||||||||||||||||||    ||||||||  ||||||||||||||||||
Db        618 AAGTTTGGACCTTCTTGGATTANAAAAAACTGGACTGNGGTAGATTTCCTAGTGAT 672
```

RESULT 21
ADT95551
ID ADT95551 standard; cDNA; 644 BP.
XX
AC ADT95551;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #1070.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1070; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.86e-108	Length:	644
Score:	1069.00	Matches:	203
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	99.5%	Mismatches:	0
Query Match:	56.2%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2 (1-362) x ADT95551 (1-644)

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391
```

```
Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy        333 SerLeuAspLeu 336
              ||||||||||||
Db        632 AGTTTGGACCTT 643
```

RESULT 22
ADX42033
ID ADX42033 standard; cDNA; 644 BP.
XX
AC ADX42033;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 1070.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1070; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.86e-108	Length:	644
Score:	1069.00	Matches:	203
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	99.5%	Mismatches:	0
Query Match:	56.2%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2 (1-362) x ADX42033 (1-644)

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy        333 SerLeuAspLeu 336
              ||||||||||||
Db        632 AGTTTGGACCTT 643
```

RESULT 23
AAS86255
ID AAS86255 standard; cDNA; 1227 BP.
XX
AC AAS86255;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #22059.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.

```
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG22068.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 22059; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1227 BP; 311 A; 239 C; 360 G; 317 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.6e-107        Length:          1227
Score:                    1068.00         Matches:         248
Percent Similarity:       69.4%           Conservative:    11
Best Local Similarity:    66.5%           Mismatches:      25
Query Match:              56.2%           Indels:          89
DB:                       5               Gaps:            5

US-10-757-745-2 (1-362) x AAS86255 (1-1227)

Qy          1 MetGluLeuGlySerCysLeu-GluGlyGlyArgGluAlaAlaGluGluGluGlyGluPr 20
              |||||||||||||||||||||     |||  |||||||||||||||||||||||||||||
Db         74 ATGGAGTTGGGGAGTTGCCTGTTATGGCTTTAGGGAGGCGGCGGAGGAAGAGGGCGAGCC 133

Qy         20 oGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAl 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGC 193

Qy         40 aAlaVal-AlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnS 60
              |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 CGCAGTGAGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACT 253

Qy         60 erTyrPheGluProProValGluGluSer-AlaLeuGluArgArgProGluThrIleSer 79
              ||||||||||||||||  |||||||| ||||||||||||||||||||||||||||||||
Db        254 CCTACTTCGAGCCTCCGGCGGAGGAGAGCAGCCTTGGAACGCCGACCTGAAACCATCTCT 313

Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAA 373

Qy        100 IleSerProSerGluAspThrGlnGlnGluAsn-GlySerMetPheSerLeu-IleThrT 119
              |||||||||||||||||||||||||||||||||| |||||||||||||||| ||||||| |
Db        374 ATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGGCAGCATGTTCTCTCTTCATTACCT 433

Qy        119 rp-AsnIle-AspGlyLeuAsp---LeuAsnAsnLeu-SerGlu-ArgAlaArgGlyVal 136
              || |||||| ||||||:::      |||||||||||| |||||| ||||||||||||||||
Db        434 GGGAATATTTGATGGGATTAGGATCTTAAACAATCTGTTCAGAGGAGGGCTCGAGGGGTG 493

Qy        137 CysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProPro 156
              ||||||||||||||||||           ::::::|||          |||
Db        494 TGTTCCTACTTAGCTTTGTAAGTATTATCACTTCTATTTAATAGTAATGTG--------- 544

Qy        157 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 176
                     ||||||:::
Db        545 ------TCTTATGTA---------------------------------------------- 553

Qy        177 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 196

Db        553 ------------------------------------------------------------ 553

Qy        197 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 216

Db        553 ------------------------------------------------------------ 553

Qy        217 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 236
                                                                    |||
Db        554 ----------------------------------------------------TATACGGAA 562

Qy        237 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 CGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACA 622

Qy        257 ValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr---ArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||:::   :::
Db        623 GTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTGAGTGACAAAGTCAGTTTG 682

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              :::      :::  :::  :::||      :::      |||       |||||||||
Db        683 ATGTTTGAGGACAGATTATTGCTATGGAGACTAATTTAGGTTCCT---TTTTGCCAGTAT 739

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 799

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGAC 859

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 919

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        920 TGCAACTTAGATATAATATTG 940

RESULT 24
ADT95565
ID   ADT95565 standard; cDNA; 633 BP.
XX
AC   ADT95565;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1084.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
```

PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1084; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	2.14e-106	Length:	633
Score:	1054.00	Matches:	200
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	55.4%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2 (1-362) x ADT95565 (1-633)

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 633
```

RESULT 25
ADX42047
ID ADX42047 standard; cDNA; 633 BP.
XX
AC ADX42047;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 1084.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1084; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense

```
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                2.14e-106     Length:          633
Score:                    1054.00       Matches:         200
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              55.4%         Indels:          0
DB:                       11            Gaps:            0

US-10-757-745-2 (1-362) x ADX42047 (1-633)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 633

RESULT 26
ADT95029
ID   ADT95029 standard; cDNA; 625 BP.
XX
AC   ADT95029;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #548.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 548; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:                7.29e-104     Length:          625
Score:                    1031.00       Matches:         198
Percent Similarity:       99.0%         Conservative:    0
Best Local Similarity:    99.0%         Mismatches:      1
Query Match:              54.2%         Indels:          1
DB:                       11            Gaps:            0

US-10-757-745-2 (1-362) x ADT95029 (1-625)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy        173 -GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 192
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy        192 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy        212 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        232 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy        252 oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        272 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy        292 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy        312 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 623
```

```
RESULT 27
ADX41511
ID   ADX41511 standard; cDNA; 625 BP.
XX
AC   ADX41511;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 548.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 548; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              7.29e-104     Length:        625
Score:                  1031.00       Matches:       198
Percent Similarity:     99.0%         Conservative:  0
Best Local Similarity:  99.0%         Mismatches:    1
Query Match:            54.2%         Indels:        1
DB:                     11            Gaps:          0

US-10-757-745-2 (1-362) x ADX41511 (1-625)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||    |||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy        173 -GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 192
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy        192 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy        212 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        232 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy        252 oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        272 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy        292 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy        312 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 623

RESULT 28
ADT95512
ID   ADT95512 standard; cDNA; 625 BP.
XX
AC   ADT95512;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1031.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
```

KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1031; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	1.19e-102	Length:	625
Score:	1020.00	Matches:	194
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	53.6%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2 (1-362) x ADT95512 (1-625)

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 581

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 326
              ||||||||||||||||||||||||||||||||||||||||||
Db        582 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAG 623
```

RESULT 29
ADX41994
ID ADX41994 standard; cDNA; 625 BP.
XX
AC ADX41994;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 1031.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1031; 244pp; English.
XX

```
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                1.19e-102       Length:          625
Score:                    1020.00         Matches:         194
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              53.6%           Indels:          0
DB:                       11              Gaps:            0

US-10-757-745-2 (1-362) x ADX41994 (1-625)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 581

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 326
              ||||||||||||||||||||||||||||||||||||||||||
Db        582 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAG 623

RESULT 30
ADT95464
ID   ADT95464 standard; cDNA; 553 BP.
XX
AC   ADT95464;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #983.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 983; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:                1.06e-89        Length:          553
Score:                    902.00          Matches:         171
Percent Similarity:       98.3%           Conservative:    0
Best Local Similarity:    98.3%           Mismatches:      3
Query Match:              47.4%           Indels:          0
DB:                       11              Gaps:            0

US-10-757-745-2 (1-362) x ADT95464 (1-553)

Qy        129 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 148
              |||          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 CTCTANATGCATGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATA 89

Qy        149 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 168
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          90 TTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTAT 149

Qy         169 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 188
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         150 GAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 209

Qy         189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         210 GTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTT 269

Qy         209 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         270 TTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAG 329

Qy         229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         330 AGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATG 389

Qy         249 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 268
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         390 CAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGA 449

Qy         269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         450 GAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 509

Qy         289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 302
                |||||||||||||||||||||||||||||||||||||||
Db         510 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 551

RESULT 31
ADX41946
ID   ADX41946 standard; cDNA; 553 BP.
XX
AC   ADX41946;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 983.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 983; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:                1.06e-89        Length:          553
Score:                    902.00          Matches:         171
Percent Similarity:       98.3%           Conservative:    0
Best Local Similarity:    98.3%           Mismatches:      3
Query Match:              47.4%           Indels:          0
DB:                       11              Gaps:            0

US-10-757-745-2 (1-362) x ADX41946 (1-553)

Qy         129 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 148
                |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          30 CTCTANATGCATGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATA 89

Qy         149 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 168
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          90 TTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTAT 149

Qy         169 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 188
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         150 GAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 209

Qy         189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         210 GTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTT 269

Qy         209 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         270 TTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAG 329

Qy         229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         330 AGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATG 389

Qy         249 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 268
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         390 CAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGA 449

Qy         269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         450 GAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 509

Qy         289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 302
                |||||||||||||||||||||||||||||||||||||||
Db         510 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 551

RESULT 32
ADD34375/c
ID   ADD34375 standard; DNA; 774 BP.
XX
AC   ADD34375;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Mouse mitochondrial DNA sequence SEQ ID NO:2153.
XX
KW   ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW   mitochondrial disease; oxidative phosphorylation dysfunction;
KW   oxidative stress; apoptosis; aging.
XX
OS   Mus musculus.
XX
```

PN WO2003020220-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027886.
XX
PR 30-AUG-2001; 2001US-0316323P.
PR 31-AUG-2001; 2001CA-02356540.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Wallace DC, Levy S, Kerstann K, Procaccio V;
XX
DR WPI; 2003-300821/29.
XX
PT Array containing probes for genes involved in mitochondrial biology, useful for determining mitochondrial biology gene expression profiles for use in diagnosing pathologies and identifying biochemical pathways.
XX
PS Claim 2; SEQ ID NO 2153; 201pp; English.
XX
CC The invention relates to a novel array comprising at least two isolated nucleotide molecules, each molecule having a sequence capable of uniquely hybridising to a nucleic acid molecule which is an expression product of a gene involved in mitochondrial biology. The array comprises two or more isolated nucleic acid molecules or spots, each molecule having a sequence chosen from sequence of 994 human probes and 2046 mouse probes. An array of the invention is useful for determining an expression profile of a mouse or human sample containing nucleic acid, by contacting the array with the sample under conditions allowing selective hybridisation, and measuring hybridisation of nucleic acid in the sample to the array to produce an expression profile. The array is also useful for determining an expression profile of a first labelled sample containing nucleic acid relative to a second, differently labelled sample containing nucleic acid. The second sample is a reference or a standard. An array is useful for determining an expression profile diagnostic of an energy-metabolism-related physiological condition. An array of the invention is useful for determining mitochondrial biology gene expression profiles of organisms, such as human, mice and closely related species, tissue and organs of such organisms, which are useful for determining expression profiles diagnostic of energy metabolism-related physiological conditions, diagnosing such physiological conditions, identifying biochemical pathways, genes, and mutations involved in such physiological conditions, identifying therapeutic agents useful for preventing and/or treating such physiological conditions, evaluating and/or monitoring the efficacy of such therapies, and creating and identifying animal models of human energy metabolism-related physiological conditions. An array is also useful for defining expression signatures or profiles for mitochondrial diseases, as well as distinguishing clinical disorders that result from oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress, apoptosis and aging. An array of the invention contains probes of genes not previously recognised to participate in mitochondrial biology. The sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA clones used to make the probes of the invention. Some sequences are not present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905, 1906, 2408 and 2643.
XX
SQ Sequence 774 BP; 218 A; 169 C; 135 G; 252 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.61e-84	Length:	774
Score:	857.00	Matches:	166
Percent Similarity:	88.4%	Conservative:	24
Best Local Similarity:	77.2%	Mismatches:	23
Query Match:	45.1%	Indels:	4
DB:	10	Gaps:	1

US-10-757-745-2 (1-362) x ADD34375 (1-774)

```
Qy        148 IlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsn 167
              :::|||||||||||||   |||||||||||||   :::|||||||||||||||||   |||
Db        768 GTATTTCTACAGGAAGT-ATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGC-AGCCAG 711

Qy        168 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 187
              |||   |||||||||||:::|||||||||||||||||||||||:::|||||||||
Db        710 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 651

Qy        188 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 207
              |||||||||   |||||||||||||||||||||||||||:::|||||||||||||||:::
Db        650 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAGC 591

Qy        208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
              ||||||||||||:::|||::::::   |||||||||   ||||||||||||||||||||
Db        590 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 531

Qy        228 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 247
              ||||||||||||   |||:::|||||||||:::   |||||||||   ||||||   |||
Db        530 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 471

Qy        248 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 267
              |||||||||||||||:::|||   ||||||||||||||||||||||||||||||||||||
Db        470 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 411

Qy        268 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
              :::|||||||   :::|||||||||||||||:::|||:::   |||   ||||||||||||
Db        410 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 351

Qy        288 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 307
              |||||||||||||||||||||||||||||||||||:::   |||:::||||||   |||
Db        350 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 291

Qy        308 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 327
              ||||||   |||   ||||||||||||||||||||||||       ||||||||||||
Db        290 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGG 237

Qy        328 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 347
              |||:::||||||:::|||||||||||:::|||||||||||||||||||||||||||||||
Db        236 CACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTT 177

Qy        348 ProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||||||   |||:::::::::|||
Db        176 CCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 132
```

RESULT 33
AAX40590
ID AAX40590 standard; cDNA; 483 BP.
XX
AC AAX40590;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human secreted protein 5' EST SEQ ID No: 190.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis; forensic; gene therapy; chromosome mapping; signal peptide; prostate; upstream regulatory sequence; cytokine activity; cell proliferation; differentiation; haematopoiesis regulation; tissue growth regulation; reproductive hormone regulation; chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS Homo sapiens.
XX
PN WO9906550-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB001232.
XX
PR 01-AUG-1997; 97US-00905144.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
DR WPI; 1999-153780/13.
DR P-PSDB; AAY11868.

XX
PT New isolated prostate-derived nucleic acids - used to develop products
PT which may have cytokine, immune regulatory, haematopoiesis regulating,
PT anti-inflammatory or tumour inhibition activity.
XX
PS Claim 1; Page 298; 675pp; English.
XX
CC AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins expressed in prostate, and encode the proteins
CC given in AAY11716 to AAY11993 respectively. The proteins given represent
CC the signal peptide and an N-terminal fragment of a secreted protein. The
CC nucleic acid sequences can be used for producing secreted human gene
CC products. They can also be used to develop products for diagnosis and
CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation and differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptides can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 483 BP; 123 A; 111 C; 139 G; 110 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.31e-82	Length:	483
Score:	837.00	Matches:	161
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	44.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2 (1-362) x AAX40590 (1-483)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 Leu 161
              |||
Db        481 CTA 483
```

RESULT 34
ADD34376
ID ADD34376 standard; DNA; 602 BP.
XX
AC ADD34376;
XX
DT 15-JAN-2004 (first entry)
XX
DE Mouse mitochondrial DNA sequence SEQ ID NO:2154.
XX
KW ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW mitochondrial disease; oxidative phosphorylation dysfunction;
KW oxidative stress; apoptosis; aging.
XX
OS Mus musculus.
XX
PN WO2003020220-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027886.
XX
PR 30-AUG-2001; 2001US-0316323P.
PR 31-AUG-2001; 2001CA-02356540.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Wallace DC, Levy S, Kerstann K, Procaccio V;
XX
DR WPI; 2003-300821/29.
XX
PT Array containing probes for genes involved in mitochondrial biology,
PT useful for determining mitochondrial biology gene expression profiles for
PT use in diagnosing pathologies and identifying biochemical pathways.
XX
PS Claim 2; SEQ ID NO 2154; 201pp; English.
XX
CC The invention relates to a novel array comprising at least two isolated
CC nucleotide molecules, each molecule having a sequence capable of uniquely
CC hybridising to a nucleic acid molecule which is an expression product of
CC a gene involved in mitochondrial biology. The array comprises two or more
CC isolated nucleic acid molecules or spots, each molecule having a sequence
CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC of the invention is useful for determining an expression profile of a
CC mouse or human sample containing nucleic acid, by contacting the array
CC with the sample under conditions allowing selective hybridisation, and
CC measuring hybridisation of nucleic acid in the sample to the array to
CC produce an expression profile. The array is also useful for determining
CC an expression profile of a first labelled sample containing nucleic acid
CC relative to a second, differently labelled sample containing nucleic
CC acid. The second sample is a reference or a standard. An array is useful
CC for determining an expression profile diagnostic of an energy-metabolism-
CC related physiological condition. An array of the invention is useful for
CC determining mitochondrial biology gene expression profiles of organisms,
CC such as human, mice and closely related species, tissue and organs of
CC such organisms, which are useful for determining expression profiles
CC diagnostic of energy metabolism-related physiological conditions,
CC diagnosing such physiological conditions, identifying biochemical
CC pathways, genes, and mutations involved in such physiological conditions,
CC identifying therapeutic agents useful for preventing and/or treating such
CC physiological conditions, evaluating and/or monitoring the efficacy of
CC such therapies, and creating and identifying animal models of human
CC energy metabolism-related physiological conditions. An array is also
CC useful for defining expression signatures or profiles for mitochondrial
CC diseases, as well as distinguishing clinical disorders that result from
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC apoptosis and aging. An array of the invention contains probes of genes
CC not previously recognised to participate in mitochondrial biology. The
CC sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC clones used to make the probes of the invention. Some sequences are not
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC 1906, 2408 and 2643.
XX

```
SQ   Sequence 602 BP; 194 A; 119 C; 134 G; 155 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.01e-69        Length:         602
Score:                  714.50          Matches:        141
Percent Similarity:     83.2%           Conservative:   23
Best Local Similarity:  71.6%           Mismatches:     32
Query Match:            37.6%           Indels:         1
DB:                     10              Gaps:           1

US-10-757-745-2 (1-362) x ADD34376 (1-602)

Qy         73 ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 92
              |||:::|||   |||      :::   :::   |||||||||||||||||||:::
Db         15 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 74

Qy         93 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 112
              |||:::|||      :::   ||||||||      |||     |||:::   |||
Db         75 AATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGCAGC 131

Qy        113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132
                 |||   ||||||||||||||||||||||||||      ||||||   ||||||
Db        132 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 191

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||   |||||||||||||||||||||||:::|||||||||||||
Db        192 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 251

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||   :::|||||||||||||||:::::::::|||   |||||||||
Db        252 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 311

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              |||:::|||||||||||||||||||||||:::|||||||||   |||||||||   |||
Db        312 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 371

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||:::|||||||||||||:::|||||||||||||:::
Db        372 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAGCCTGCTATGCGTAAAT 431

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              |||::::::   ||||||||   ||||||||||||||||||||||||||||||
Db        432 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 491

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              |||:::|||||||:::   |||||||||   ||||||   ||||||||||||||||||
Db        492 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 551

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
              :::|||   |||||||||||||||||||||||||||||||||||:::|||
Db        552 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAA 602

RESULT 35
ADT95844
ID   ADT95844 standard; cDNA; 401 BP.
XX
AC   ADT95844;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1363.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1363; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 401 BP; 132 A; 67 C; 89 G; 113 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.41e-62        Length:         401
Score:                  651.00          Matches:        127
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            34.2%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2 (1-362) x ADT95844 (1-401)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 78

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 138

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 198

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258
```

```
Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 318

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 378

Qy        253 GluSerAlaThrValIlePhe 259
              |||||||||||||||||||||
Db        379 GAGTCAGCTACAGTTATATTT 399
```

RESULT 36
ADX42326
ID ADX42326 standard; cDNA; 401 BP.
XX
AC ADX42326;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 1363.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1363; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 401 BP; 132 A; 67 C; 89 G; 113 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.41e-62	Length:	401
Score:	651.00	Matches:	127
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	34.2%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2 (1-362) x ADX42326 (1-401)

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 78

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 138

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 198

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 318

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 378

Qy        253 GluSerAlaThrValIlePhe 259
              |||||||||||||||||||||
Db        379 GAGTCAGCTACAGTTATATTT 399
```

RESULT 37
ADT95307/c
ID ADT95307 standard; cDNA; 391 BP.
XX
AC ADT95307;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #826.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 826; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)

CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 391 BP; 83 A; 100 C; 103 G; 104 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	1.21e-46	Length:	391
Score:	510.00	Matches:	98
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	26.8%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2 (1-362) x ADT95307 (1-391)

```
Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 AGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGG 278

Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 GCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAA 218

Qy         77 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 96
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACC 158

Qy         97 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 ACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTC 98

Qy        117 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 ATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 44
```

RESULT 38
ADX41789/c
ID ADX41789 standard; cDNA; 391 BP.
XX
AC ADX41789;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 826.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 826; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 391 BP; 83 A; 100 C; 103 G; 104 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	1.21e-46	Length:	391
Score:	510.00	Matches:	98
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	26.8%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2 (1-362) x ADX41789 (1-391)

```
Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 AGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGG 278

Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 GCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAA 218

Qy         77 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 96
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACC 158

Qy         97 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 ACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTC 98

Qy        117 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 ATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 44
```

RESULT 39
AAX41008
ID AAX41008 standard; cDNA; 258 BP.
XX
AC AAX41008;
XX
DT 18-JUN-1999 (first entry)
XX

```
DE   Human secreted protein 5' EST SEQ ID NO: 220.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS   Homo sapiens.
XX
PN   WO9906554-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB001238.
XX
PR   01-AUG-1997;   97US-00905134.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Lacroix B;
XX
DR   WPI; 1999-153784/13.
DR   P-PSDB; AAY12175.
XX
PT   New nucleic acids encoding human secreted proteins - obtained from cDNA
PT   libraries prepared from kidney, fetal kidney, dystrophic muscle, muscle
PT   and heart tissue.
XX
PS   Claim 1; Page 319; 622pp; English.
XX
CC   AAX40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for
CC   human secreted proteins, and encode the proteins given in AAY01602 and
CC   AAY11994 to AAY12260, respectively. The proteins given represent the
CC   signal peptide and an N-terminal fragment of a secreted protein. The
CC   nucleic acid sequences can be used for producing secreted human gene
CC   products. They can also be used to develop products for diagnosis and
CC   therapy. The proteins obtained may have cytokine activity, cell
CC   proliferation/differentiation activity, haematopoiesis regulating
CC   activity, tissue growth regulating activity, reproductive hormone
CC   regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC   thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, tumour inhibition activity or other activities. The products
CC   can be used in forensic, gene therapy and chromosome mapping procedures.
CC   The sequences can also be used for obtaining corresponding promoter
CC   sequences. The nucleic acids encoding the signal peptide can be used for
CC   directing extracellular secretion of a polypeptide or the insertion of a
CC   polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ   Sequence 258 BP; 54 A; 54 C; 105 G; 40 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.:                4.09e-32        Length:          258
Score:                    376.00          Matches:         72
Percent Similarity:       96.0%           Conservative:    0
Best Local Similarity:    96.0%           Mismatches:      3
Query Match:              19.8%           Indels:          0
DB:                       2               Gaps:            0

US-10-757-745-2 (1-362) x AAX41008 (1-258)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 90

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              |||||||||||||||||||||||||    ||||||   ||||||||||||||||||||||
Db         91 GAGGTGAAAAAGCGGCGACTTCTGTGTSTRGAGTTTRCCTCGGTCGCAAGCTGCGATGCC 150

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        151 GCAGTGGCTCAGTGCTTCCTGGCCGAKAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 210

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 75
              |||||||||||||||||||||||||||||||||||||||||||||
Db        211 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCA 255

RESULT 40
ABK27684
ID   ABK27684 standard; cDNA; 176 BP.
XX
AC   ABK27684;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human colon cancer expressed sequence tag, Seq ID no 121.
XX
KW   Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
KW   expressed sequence tag.
XX
OS   Homo sapiens.
XX
PN   WO200196390-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US018577.
XX
PR   09-JUN-2000; 2000US-0210821P.
PR   18-DEC-2000; 2000US-0256571P.
PR   10-MAY-2001; 2001US-0290240P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Hepler WT,  Clapper JD,  Wang A,  Secrist H;
XX
DR   WPI; 2002-139708/18.
XX
PT   Novel isolated polynucleotide encoding a polypeptide comprising a portion
PT   of colon tumor protein, useful for detection, diagnosis and therapy of
PT   human colon cancer.
XX
PS   Claim 1; Page 182; 220pp; English.
XX
CC   The invention relates to an isolated polynucleotide (I) encoding a
CC   polypeptide (II) comprising at least a portion of a colon tumour protein.
CC   (I), (II) and antibody (III) to (II) are useful for determining the
CC   presence of a cancer in a patient. (I), (II) or antigen presenting cells
CC   expressing (I) is useful for stimulating and/or expanding T cells
CC   specific for a tumour protein, by contacting T cells with (I), (II) or
CC   antigen-presenting cells that express (I), under conditions and for a
CC   time sufficient to permit the stimulation and/or expansion of T cells.
CC   (I), (II), or antigen presenting cells that express (II) are useful for
CC   treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
CC   isolated from a patient with (I), (II) or antigen presenting cells that
CC   express (II), such that T cells proliferate, and administering to the
CC   patient an effective amount of the proliferated T cells, thus inhibiting
CC   the development of a cancer in the patient. (I) or (II) is useful in
CC   vaccines and pharmaceutical compositions for prevention and treatment of
CC   colon malignancies and for the diagnosis and monitoring of such cancers.
CC   (I), (II) or (III) is useful for detection, diagnosis and/or therapy of
CC   human colon cancer. (I) is useful as a probe or primer for nucleic acid
CC   hybridisation, and in the design and preparation of ribozyme molecules
CC   for inhibiting expression of (II) in tumour cells. ABK27564-ABK27807
CC   represent novel human colon cancer coding sequences and primers of the
CC   invention
XX
SQ   Sequence 176 BP; 68 A; 27 C; 31 G; 50 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.58e-24        Length:          176
Score:                    298.00          Matches:         58
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              15.7%           Indels:          0
```

```
DB:                    6              Gaps:          0

US-10-757-745-2 (1-362) x ABK27684 (1-176)

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 62

Qy        164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 122

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176

RESULT 41
ABV95959/c
ID   ABV95959 standard; cDNA; 176 BP.
XX
AC   ABV95959;
XX
DT   14-JAN-2003  (first entry)
XX
DE   Human pancreatic cancer expressed cDNA SEQ ID NO 1367.
XX
KW   Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW   cytostatic; tumour; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200260317-A2.
XX
PD   08-AUG-2002.
XX
PF   30-JAN-2002; 2002WO-US002781.
XX
PR   30-JAN-2001; 2001US-0265305P.
PR   31-JAN-2001; 2001US-0265682P.
PR   09-FEB-2001; 2001US-0267568P.
PR   21-MAR-2001; 2001US-0278651P.
PR   28-APR-2001; 2001US-0287112P.
PR   16-MAY-2001; 2001US-0291631P.
PR   12-JUL-2001; 2001US-0305484P.
PR   20-AUG-2001; 2001US-0313999P.
PR   27-NOV-2001; 2001US-0333626P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Benson DR,  Kalos MD,  Lodes MJ,  Persing DH,  Hepler WT,  Jiang Y;
XX
DR   WPI; 2002-627435/67.
XX
PT   New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT   diagnosing, preventing and/or treating cancer, particularly pancreatic
PT   cancer.
XX
PS   Claim 1; SEQ ID NO 1367; 300pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (I) comprising: (a)
CC   any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
CC   complements of (a); (c) sequences consisting of at least 20 contiguous
CC   residues of (a); (d) sequences that hybridize to (a), under moderately
CC   stringent conditions; (e) sequences having at least 75% or 90% identity
CC   to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
CC   ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer
CC   in a patient and compositions comprising polypeptides, polynucleotides,
CC   antibodies, fusion proteins, T cell populations and antigen presenting
CC   cells expressing the polypeptide are useful in treating pancreatic cancer
CC   and stimulating an immune response. The polynucleotides can be used as
CC   probes or primers for nucleic acid hybridisation, in the design and
CC   preparation of ribozyme molecules for inhibiting expression of the tumour
CC   polypeptides and proteins in the tumour cells, in vaccines and for gene
CC   therapy. Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 176 BP; 50 A; 31 C; 27 G; 68 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              9.58e-24      Length:          176
Score:                  298.00        Matches:         58
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            15.7%         Indels:          0
DB:                     6             Gaps:            0

US-10-757-745-2 (1-362) x ABV95959 (1-176)

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy        164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 42
ABL36461/c
ID   ABL36461 standard; cDNA; 179 BP.
XX
AC   ABL36461;
XX
DT   08-APR-2002  (first entry)
XX
DE   Human colon tumour antigen polynucleotide SEQ ID NO:50.
XX
KW   Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW   colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200196388-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US018557.
XX
PR   09-JUN-2000; 2000US-0210899P.
PR   20-FEB-2001; 2001US-0270216P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Harlocker SL,  Secrist H;
XX
DR   WPI; 2002-114514/15.
XX
PT   Novel isolated colon tumor polynucleotide differentially expressed in
PT   colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT   useful for inhibiting development of cancer in patient.
XX
PS   Claim 1; SEQ ID NO 50; 105pp; English.
XX
CC   ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC   which were isolated from human colon tumour and colon metastatic tumour
CC   cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC   production. (I) can be used for stimulating and/or expanding T cells
CC   specific for a tumour protein on contact with the T cells. They are also
CC   useful for inhibiting the development of cancer in a patient. (I) can be
CC   used as probes or primers for nucleic acid hybridisation, for preparing
CC   mutant species primers, or primers for use in genetic constructions. (I)
CC   can be used in the diagnosis of a colon tumour
```

```
XX
SQ   Sequence 179 BP; 51 A; 31 C; 27 G; 68 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              9.81e-24        Length:         179
Score:                  298.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            15.7%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2 (1-362) x ABL36461 (1-179)

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy        164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 43
ABQ34462
ID   ABQ34462 standard; DNA; 725 BP.
XX
AC   ABQ34462;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 21053.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K,  Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 725 BP; 131 A; 64 C; 251 G; 279 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.99e-19        Length:         725
Score:                  267.00          Matches:        63
Percent Similarity:     56.8%           Conservative:   0
Best Local Similarity:  56.8%           Mismatches:     19
Query Match:            14.0%           Indels:         30
DB:                     6               Gaps:           1

US-10-757-745-2 (1-362) x ABQ34462 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 ATGGAGTTGGGGAGTTGTTTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGTTT 382

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||  ||||||||||||||||  ||||||  |||||||||
Db        383 GAGGTGAAAAAGCGGCGATTTTTGTGTGTGGAGTTTGTTTCGGTCGTAAGTTGCGATGTC 442

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu--------------- 55
                 |||      |||||||||   ||||||||||||||||||
Db        443 GTAGTGGTTTAGTGTTTTTTGGTCGAGAACGATTGGGAGATGGA-AGTACGTTTTCGTTG 501

Qy         55 ------------------------------------------------------------ 55

Db        502 TTTTCGTTTTCGGTTTTATTATTTAGGTCGGTGGAGCGTATAAATTTTATTTTTTGTGTT 561

Qy         56 ------------ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                        |||   ||||||   |||||||||       ||||||||||||   |||
Db        562 TTTGTTTTTTAGAGGGTTTTGAATTTTTATTTCGAGTTTTCGGTGGAGGAGAGCGTTTTG 621

Qy         72 GluArgArgProGluThrIleSerGluProLys 82
              |||||||||   |||   |||   |||   |||
Db        622 GAACGTCGATTTGAAATTATTTTTGAGTTTAAG 654

RESULT 44
ABQ34463/c
ID   ABQ34463 standard; DNA; 725 BP.
XX
AC   ABQ34463;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 21054.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
```

```
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K,  Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 725 BP; 279 A; 251 C; 64 G; 131 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.99e-19        Length:         725
Score:                    267.00          Matches:        63
Percent Similarity:       56.8%           Conservative:   0
Best Local Similarity:    56.8%           Mismatches:     19
Query Match:              14.0%           Indels:         30
DB:                       6               Gaps:           1

US-10-757-745-2 (1-362) x ABQ34463 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 ATGGAGTTGGGGAGTTGTTTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGTTT 344

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||   ||||||||||||||   ||||||   |||||||||
Db        343 GAGGTGAAAAAGCGGCGATTTTTGTGTGTGGAGTTTGTTTCGGTCGTAAGTTGCGATGTC 284

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu--------------- 55
                |||       |||||||||   ||||||||||||||||||
Db        283 GTAGTGGTTTAGTGTTTTTTGGTCGAGAACGATTGGGAGATGGA-AGTACGTTTTCGTTG 225

Qy         55 ------------------------------------------------------------ 55

Db        224 TTTTCGTTTTCGGTTTTATTATTTAGGTCGGTGGAGCGTATAAATTTTATTTTTTGTGTT 165

Qy         56 ------------ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                        |||   ||||||   |||||||||       ||||||||||||   |||
Db        164 TTTGTTTTTTAGAGGGTTTTGAATTTTTATTTCGAGTTTTCGGTGGAGGAGAGCGTTTTG 105

Qy         72 GluArgArgProGluThrIleSerGluProLys 82
              |||||||||   |||   |||   |||   |||
Db        104 GAACGTCGATTTGAAATTATTTTTGAGTTTAAG 72

RESULT 45
AAC47239
ID   AAC47239 standard; DNA; 1566 BP.
XX
AC   AAC47239;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 53077.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
```

PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	5.76e-15	Length:	1566
Score:	231.00	Matches:	84
Percent Similarity:	44.2%	Conservative:	39
Best Local Similarity:	30.2%	Mismatches:	109
Query Match:	12.1%	Indels:	46
DB:	3	Gaps:	11

US-10-757-745-2 (1-362) x AAC47239 (1-1566)

```
Qy        113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132
              :::  |||         |||     :::  |||:::||||||          |||
Db        523 ATCTTGAGCTACAATGTTTGGTTTCGTGAAGACCTTGAGTTGAAT---------CTAAGG 573

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
                |||   :::          :::   |||:::||||||   :::|||        ||||||
Db        574 ATGAGAGCGATTGGTCATCTTATTCAGCTTCATTCTCCTCATCTCATATGTTTTCAGGAA 633

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArg----------------SerSer 166
              |||  |||      |||        :::|||                    |||
```

```
Db        634 GTCACTCCTGAGATATATGATATTTTCCGCAAGTCGAATTGGTGGAAAGCGTATTCTTGT 693

Qy        167 AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186

Db        694 TCTGTCTCAGTTGATGTGGCAGTTTCGAGAGGCTACTACTGCATGCTGTTG--------- 744

Qy        187 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 206

Db        745 AGCAAGTTGGGAGTGAAATCTTTCAGCAGCAAAAGCTTTGGGAACTCGATAATGGGAAGA 804

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGlu---LeuCysLeuMetThrSer 225

Db        805 GAACTATCTATCGCAGAAGTTGAAGTTCCAGGGAGAAAGCCACTTGTTTTTGCAACTAGC 864

Qy        226 HisLeuGluSer----------ThrArgGlyHisAlaAlaGluArg 237

Db        865 CATCTTGAGAGTCCCTGTCCAGGTCCTCCAAAATGGGACCAAATGTTCAGTCGAGAACGC 924

Qy        238 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257

Db        925 GTTGAACAAGCCAAGGAGGCAATTGAGATTCTCAGA---------CCTAATGCCAATGTG 975

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly------ 275

Db        976 ATATTCGGGGGAGACATGAACTGGTGCGAC-----------AAGCTAGATGGGAAGTTT 1023

Qy        276 ---LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294

Db       1024 CCTTTGCCAGACAAATGGGTGGATGTTTGGGAGGTTTTG---AAACCAGGTGATTTGGGT 1080

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314

Db       1081 TTTACATATGACACAAAAGCAAACCCAATGTTGTCCGGTAACCGGGCATTGCAGAAACGG 1140

Qy        315 PheAspArgIlePhePheArg---------------------------------------- 321

Db       1141 TTAGACCGGATTTTGTGTCGTCTGGATGATTACAAGCTTGGTGGAATCGAGATGGTTGGG 1200

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGly---LeuGlu 340

Db       1201 AAAGAGGCCATTCCTGGTTTATCATATGTAAAAGAGAAGAAAGTGAGGGGTGACATTAAG 1260

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358

Db       1261 AAGTTAGAGCTTCCGGTGTTGCCTAGTGATCATTTCGGTCTGCTTGTAACCCTT 1314

RESULT 46
AAC41089
ID   AAC41089 standard; DNA; 1606 BP.
XX
AC   AAC41089;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 30607.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
```

PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	5.97e-15	Length:	1606
Score:	231.00	Matches:	84
Percent Similarity:	44.2%	Conservative:	39
Best Local Similarity:	30.2%	Mismatches:	109
Query Match:	12.1%	Indels:	46
DB:	3	Gaps:	11

US-10-757-745-2 (1-362) x AAC41089 (1-1606)

```
Qy        113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132
              :::  |||      |||     :::  |||:::||||||          |||
Db        512 ATCTTGAGCTACAATGTTTGGTTTCGTGAAGACCTTGAGTTGAAT---------CTAAGG 562

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
                |||  :::         :::  |||:::||||||  :::|||        ||||||
Db        563 ATGAGAGCGATTGGTCATCTTATTCAGCTTCATTCTCCTCATCTCATATGTTTTCAGGAA 622

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArg----------------SerSer 166
              |||  |||      |||      :::|||                     |||
Db        623 GTCACTCCTGAGATATATGATATTTTCCGCAAGTCGAATTGGTGGAAAGCGTATTCTTGT 682

Qy        167 AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186
              :::      :::                  ||||||:::      ::::::
Db        683 TCTGTCTCAGTTGATGTGGCAGTTTCGAGAGGCTACTACTGCATGCTGTTG--------- 733

Qy        187 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 206
              |||::::::  :::||||||          |||  ::::::  |||     |||
Db        734 AGCAAGTTGGGAGTGAAATCTTTCAGCAGCAAAAGCTTTGGGAACTCGATAATGGGAAGA 793

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGlu---LeuCysLeuMetThrSer 225
                |||          |||  |||  |||  :::  |||          ||||||
Db        794 GAACTATCTATCGCAGAAGTTGAAGTTCCAGGGAGAAAGCCACTTGTTTTTGCAACTAGC 853

Qy        226 HisLeuGluSer----------------------ThrArgGlyHisAlaAlaGluArg 237
              ||||||||||||                      :::      :::   ||||||
Db        854 CATCTTGAGAGTCCCTGTCCAGGTCCTCCAAAATGGGACCAAATGTTCAGTCGAGAACGC 913

Qy        238 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
              :::  |||  |||      ::::::  ::::::          :::|||   |||
Db        914 GTTGAACAAGCCAAGGAGGCAATTGAGATTCTCAGA---------CCTAATGCCAATGTG 964

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly------ 275
```

```
                ||||||   ||||||   |||      |||           :::      |||
Db        965 ATATTCGGGGGAGACATGAACTGGTGCGAC-----------AAGCTAGATGGGAAGTTT 1012

Qy        276 ---LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294
                  ||||||:::      ||||||||||||||||   |||   ||||||
Db       1013 CCTTTGCCAGACAAATGGGTGGATGTTTGGGAGGTTTTG---AAACCAGGTGATTTGGGT 1069

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314
                :::|||:::|||||:::   |||      |||           |||   :::   |||
Db       1070 TTTACATATGACACAAAAGCAAACCCAATGTTGTCCGGTAACCGGGCATTGCAGAAACGG 1129

Qy        315 PheAspArgIlePhePheArg------------------------------------- 321
                  |||||||||   |||
Db       1130 TTAGACCGGATTTTGTGTCGTCTGGATGATTACAAGCTTGGTGGAATCGAGATGGTTGGG 1189

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGly---LeuGlu 340
                  |||      |||              :::          :::   |||   ::::::
Db       1190 AAAGAGGCCATTCCTGGTTTATCATATGTAAAAGAGAAGAAAGTGAGGGGTGACATTAAG 1249

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358
                ||||||:::          ||||||||||||:::|||||||||      |||
Db       1250 AAGTTAGAGCTTCCGGTGTTGCCTAGTGATCATTTCGGTCTGCTTGTAACCCTT 1303

RESULT 47
ABT07150/c
ID   ABT07150 standard; cDNA; 268 BP.
XX
AC   ABT07150;
XX
DT   07-NOV-2002  (first entry)
XX
DE   Human ovarian cancer associated coding sequence SEQ ID NO: 312.
XX
KW   Human; ovarian cancer; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2002076715-A1.
XX
PD   20-JUN-2002.
XX
PF   06-JUN-2001; 2001US-00876889.
XX
PR   23-SEP-1998;   98US-00159320.
PR   08-FEB-1999;   99US-00246429.
PR   16-SEP-1999;   99US-00397787.
XX
PA   (BENS/) BENSON D R.
PA   (LODE/) LODES M J.
PA   (MITC/) MITCHAM J L.
PA   (KING/) KING G E.
XX
PI   Benson DR,  Lodes MJ,  Mitcham JL,  King GE;
XX
DR   WPI; 2002-598720/64.
XX
PT   Composition for detecting and treating ovarian cancer, comprises a
PT   specific polypeptide, polynucleotide, T cell population, or antigen
PT   presenting cell.
XX
PS   Example 1; Page 116; 188pp; English.
XX
CC   The present invention relates to a method of detecting the presence of
CC   ovarian cancer in a patient, involving detecting ovarian cancer
CC   associated polynucleotides. The method is not only used to detect the
CC   presence of cancer, preferably ovarian cancer in a patient, but also is
CC   used to stimulate and/or expand T cells specific for an ovarian tumour
CC   protein. The sequences can be used in vaccines used to treat cancer. The
CC   present sequence is an ovarian cancer associated coding sequence
XX
SQ   Sequence 268 BP; 87 A; 62 C; 43 G; 76 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:              5.59e-15        Length:         268
Score:                  221.00          Matches:        41
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            11.6%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2 (1-362) x ABT07150 (1-268)

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAA 208

Qy        342 LeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 CTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATA 148

Qy        362 Leu 362
              |||
Db        147 TTG 145

RESULT 48
ABX73028/c
ID   ABX73028 standard; cDNA; 268 BP.
XX
AC   ABX73028;
XX
DT   14-MAR-2003  (first entry)
XX
DE   Human ovarian carcinoma antigen partial cDNA sequence #294.
XX
KW   Human; cancer detection; ovarian carcinoma antigen; ovarian cancer;
KW   tumour antigen; tumour; OV2; OV3; OV6; OV9; OV10; OV12; OV14; OV17; OV18;
KW   OV23; OV24; OV27; OV41; OV54; OV57; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US6468758-B1.
XX
PD   22-OCT-2002.
XX
PF   16-SEP-1999;   99US-00397787.
XX
PR   23-SEP-1998;   98US-00159320.
PR   08-FEB-1999;   99US-00246429.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Benson DR,  Lodes MJ,  Mitcham JL,  King GE;
XX
DR   WPI; 2003-147101/14.
XX
PT   Determining presence or absence of cancer in patient by contacting
PT   patient sample with oligonucleotide that hybridizes to polynucleotide
PT   encoding ovarian carcinoma antigen, and detecting amount of hybridization
PT   in sample.
XX
PS   Example 1; Col 223-224; 152pp; English.
XX
CC   The present invention relates to compositions and methods for determining
CC   presence or absence of cancer in patient. The method comprising
CC   contacting a biological sample with an oligonucleotide that hybridises to
CC   partial polynucleotide sequence encoding for human ovarian carcinoma
CC   antigen, or its complement. The method and compositions are useful for
CC   the therapy and diagnosis of ovarian cancer. The method can be used to
CC   identify tumour antigens that are secreted from ovarian carcinoma and/or
CC   other tumours. Effective cancer detection is achieved using the method of
CC   the invention. ABX72720-ABX73050 represent partial cDNA sequences
CC   encoding human ovarian carcinoma antigens
XX
SQ   Sequence 268 BP; 87 A; 62 C; 43 G; 76 T; 0 U; 0 Other;

Alignment Scores:
```

```
Pred. No.:              5.59e-15        Length:         268
Score:                  221.00          Matches:        41
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            11.6%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2 (1-362) x ABX73028 (1-268)

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAA 208

Qy        342 LeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 CTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATA 148

Qy        362 Leu 362
              |||
Db        147 TTG 145

RESULT 49
ABQ34465
ID   ABQ34465 standard; DNA; 725 BP.
XX
AC   ABQ34465;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 21056.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K,  Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 725 BP; 318 A; 213 C; 64 G; 130 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.55e-10        Length:         725
Score:                  184.50          Matches:        40
Percent Similarity:     58.9%           Conservative:   26
Best Local Similarity:  35.7%           Mismatches:     17
Query Match:            9.7%            Indels:         30
DB:                     6               Gaps:           1

US-10-757-745-2 (1-362) x ABQ34465 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              :::::::|||   :::   |||:::      ::::::      :::::::::   ||||||
Db        323 ATAAAATTAAAAAATTACCTAAAAAACGAAAAAAAAACGACGAAAAAAAAAAACGAACCT 382

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ::::::|||||||||||||||||||   ::::::|||   |||:::|||:::   |||
Db        383 AAAATAAAAAAACGACGACTTCTATATATAAAATTTACCTCGATCGCAAACTACGATACC 442

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGlu---------------------- 53
              |||:::   |||   ||||||   |||||||||   :::
Db        443 GCAATAACTCAATACTTCCTAACCGAAAACGACTA-AAAAATAAAAATACGCTTTCGCTA 501

Qy         53 ------------------------------------------------------------ 53

Db        502 TCCTCGCTTCCGACTCCATTACCTAAATCGATAAAACGCACAAACCTCATCCCTTATACC 561

Qy         54 ------MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                    ::::::   ||||||||||||||||||||||:::::::::::||||||
Db        562 TTTATTTTTCAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAACGCCTTA 621

Qy         72 GluArgArgProGluThrIleSerGluProLysThr 83
              :::|||||||||:::||||||||||:::|||||||||
Db        622 AAACGCCGACCTAAAACCATCTCTAAACCCAAAACC 657

RESULT 50
ABQ34464/c
ID   ABQ34464 standard; DNA; 725 BP.
XX
AC   ABQ34464;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 21055.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K,  Guetig D;
XX
```

```
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 725 BP; 130 A; 64 C; 213 G; 318 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.55e-10      Length:          725
Score:                    184.50        Matches:         40
Percent Similarity:       58.9%         Conservative:    26
Best Local Similarity:    35.7%         Mismatches:      17
Query Match:              9.7%          Indels:          30
DB:                       6             Gaps:            1

US-10-757-745-2 (1-362) x ABQ34464 (1-725)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ::::::|||   :::   |||:::      ::::::      :::::::::   ||||||
Db        403 ATAAAATTAAAAAATTACCTAAAAAACGAAAAAAAAAACGACGAAAAAAAAAAAACGAACCT 344

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ::::::||||||||||||||||||   ::::::|||   |||:::|||:::   |||
Db        343 AAAATAAAAAAACGACGACTTCTATATATAAAATTTACCTCGATCGCAAACTACGATACC 284

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGlu---------------------- 53
              |||:::   |||   ||||||   |||||||||   :::
Db        283 GCAATAACTCAATACTTCCTAACCGAAAACGACTA-AAAAATAAAAATACGCTTTCGCTA 225

Qy         53 ------------------------------------------------------------ 53

Db        224 TCCTCGCTTCCGACTCCATTACCTAAATCGATAAAACGCACAAACCTCATCCCTTATACC 165

Qy         54 ------MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                    ::::::   ||||||||||||||||||||||||:::::::::::||||||
Db        164 TTTATTTTTCAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAACGCCTTA 105

Qy         72 GluArgArgProGluThrIleSerGluProLysThr 83
              :::|||||||||:::||||||||||:::|||||||||
Db        104 AAACGCCGACCTAAAACCATCTCTAAACCCAAAACC 69

RESULT 51
AAI15311
ID   AAI15311 standard; DNA; 444 BP.
XX
AC   AAI15311;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #5244 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 5244; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 444 BP; 138 A; 60 C; 80 G; 166 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.08e-10      Length:          444
Score:                    182.50        Matches:         37
Percent Similarity:       74.5%         Conservative:    4
Best Local Similarity:    67.3%         Mismatches:      11
Query Match:              9.6%          Indels:          3
DB:                       4             Gaps:            1

US-10-757-745-2 (1-362) x AAI15311 (1-444)

Qy        135 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 151
              |||:::|||       |||                |||||||||||||||||||||||||
Db        221 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 280

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 340

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186
              ||||||:::          |||   |||   :::       :::|||
Db        341 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 385

RESULT 52
AAI24484
ID   AAI24484 standard; DNA; 691 BP.
XX
AC   AAI24484;
XX
```

```
DT   12-OCT-2001  (first entry)
XX
DE   Probe #14417 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 14417; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 691 BP; 230 A; 81 C; 112 G; 268 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.95e-10        Length:          691
Score:                    182.50          Matches:         37
Percent Similarity:       74.5%           Conservative:    4
Best Local Similarity:    67.3%           Mismatches:      11
Query Match:              9.6%            Indels:          3
DB:                       4               Gaps:            1

US-10-757-745-2 (1-362) x AAI24484 (1-691)

Qy         135 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 151
               |||:::|||     |||                 ||||||||||||||||||||||||
Db          83 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 142

Qy         152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         143 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 202

Qy         172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186
               ||||||:::       |||   |||   :::     :::|||
Db         203 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 247

RESULT 53
AAS69905
ID   AAS69905 standard; cDNA; 267 BP.
XX
AC   AAS69905;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #5709.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG05718.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 5709; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 267 BP; 49 A; 73 C; 81 G; 64 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.06e-09        Length:          267
Score:                    169.00          Matches:         32
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              8.9%            Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2 (1-362) x AAS69905 (1-267)

Qy          24 LysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaValAla 43
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           3 AAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCT 62

Qy          44 GlnCysPheLeuAlaGluAsnAspTrpGluMetGlu 55
```

```
                     ||||||||||||||||||||||||||||||||||
Db         63 CAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAA 98

RESULT 54
ACN59762
ID   ACN59762 standard; cDNA; 525 BP.
XX
AC   ACN59762;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Cotton gynoecium tissue EST Clone ID: LIB3829-034-Q1-K6-G12, SEQ:14543.
XX
KW   Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoecium;
KW   variety Nucotton33B; library LIB3829; molecular tag; molecular marker;
KW   genetic mapping; molecular mapping; seed germination; plant growth;
KW   plant quality; plant yield; plant breeding; tissue printing; ss.
XX
OS   Gossypium hirsutum.
XX
PN   US2004123340-A1.
XX
PD   24-JUN-2004.
XX
PF   12-DEC-2001; 2001US-00021323.
XX
PR   14-DEC-2000; 2000US-0255619P.
XX
PA   (DEIK/) DEIKMAN J.
PA   (FENG/) FENG P C C.
PA   (FINC/) FINCHER K L.
PA   (ZIEG/) ZIEGLER T E.
XX
PI   Deikman J,  Feng PCC,  Fincher KL,  Ziegler TE;
XX
DR   WPI; 2004-479808/45.
XX
PT   New isolated nucleic acid molecule that encodes a plant protein or its
PT   fragment, useful for isolating a variety of agronomically significant
PT   genes associated with plant growth, quality or yield, and as molecular
PT   tags to map genes.
XX
PS   Claim 1; SEQ ID NO 14543; 34pp; English.
XX
CC   The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC   ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC   from primed or non-primed seeds from variety DP50B, mature seeds from
CC   variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC   tissue, developing fibres, carpel walls and septa from variety
CC   Nucotton33B. The invention also relates to substantially purified
CC   proteins or their fragments encoded by nucleic acid molecules of the
CC   invention, and to transformed plants having a nucleic acid construct
CC   comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC   molecular tags to isolate genetic regions, to isolate genes, to map
CC   genes, to determine gene function and to determining whether genes are
CC   members of a particular gene family. The nucleic acid molecules may be
CC   used for isolating a variety of agronomically significant genes
CC   associated with plant growth, quality, yield, and could also serve as
CC   links in metabolic and catabolic pathways. The nucleic acid molecules are
CC   also useful for identifying genes important in initiating and maintaining
CC   seed germination or that may be used to mitigate stresses encountered
CC   during seed germination. The ESTs additionally enable the acquisition of
CC   promoters and cis-regulatory elements which will be useful to express
CC   agronomically significant genes in these tissues and/or other tissues,
CC   and also permits the acquisition of molecular markers useful in breeding
CC   schemes, genetic and molecular mapping, and in cloning of agronomically
CC   significant genes. The nucleic acid molecules are further useful for
CC   detecting the expression level or pattern of a protein or mRNA and for
CC   detecting the presence or quantity of a protein by tissue printing. The
CC   present sequence represents a specifically claimed EST isolated from a
CC   cotton variety Nucotton33B gynoecium tissue cDNA library (LIB3829). The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the US
CC   patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX
SQ   Sequence 525 BP; 141 A; 88 C; 134 G; 162 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.18e-07      Length:          525
Score:                  158.50        Matches:         53
Percent Similarity:     46.1%         Conservative:    29
Best Local Similarity:  29.8%         Mismatches:      63
Query Match:            8.3%          Indels:          33
DB:                     13            Gaps:            7

US-10-757-745-2 (1-362) x ACN59762 (1-525)

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   :::   :::         :::   |||:::|||||||||:::|||
Db         19 AAGAGGATGAAAGCTATTGGTGACCTTATACAACTTCATTCCCCTGATATCATTTGTTTT 78

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              |||||||||   |||   |||             |||   ||||||:::
Db         79 CAAGAGGTTACTCCGATGATTTATGACATTTTT------CGTGGGTCTAATTGG------ 126

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIle---------------------- 182
                          :::|||||||      :::|||
Db        127 -----------TGGAAAGGGTACCGGTGCTCTATCTCGGACAACACGGCGAGTTTGAGA 174

Qy        183 ---------MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 199
                      ::::::  |||::::::  :::||||||              ||||||
Db        175 GCTTACTTTTGTATACAGTTAAGCAAATTGCCTGTGAAATCCTTTCGTCGTGAACCTTTT 234

Qy        200 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 219
              ::::::   |||   |||   |||             |||   |||||||||   |||
Db        235 GATAATTCGATAATGGGGCGAGAACTTTGCATGACTGAGGTTGAAGTTTCTGGGGGTGAA 294

Qy        220 ------LeuCysLeuMetThrSerHisLeuGluSer----------------------- 229
                  |||   :::   |||:::||||||||||||
Db        295 ACAACGCTTGTTGTTGCTACTACCCATCTCGAGAGTCCGTGTCCTGCTCCGCCGAAATGG 354

Qy        230 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 249
              :::   ::::::   ||||||:::   |||   |||:::   :::   ::::::
Db        355 GATCAAATGTATAGTAAGGAACGTGTAGAACAAGCGAAAGTAGCGATCAATATTCTCGAG 414

Qy        250 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
              |||             |||||||||   ||||||   |||       |||:::
Db        415 AATAACCCG---------AATGTTATTTTTGGGGGTGATATGAATTGGGATGACAAATTA 465

Qy        270 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
              |||         |||   :::       :::|||   |||       |||
Db        466 GATGGTCGATTT---CCTTTAACTGATGGATGGATTGATGCTTGGAACGAGTTA 516

RESULT 55
ADZ62001
ID   ADZ62001 standard; DNA; 9737 BP.
XX
AC   ADZ62001;
XX
DT   30-JUN-2005  (first entry)
XX
DE   Murine Ccr4 gene, SEQ ID 1725.
XX
KW   Drug screening; Peroxisome Proliferator-Activated Receptor gamma;
KW   PPAR gamma; gene; ds.
XX
OS   Mus musculus.
XX
PN   US2005084872-A1.
XX
PD   21-APR-2005.
XX
PF   23-JAN-2004; 2004US-00764420.
XX
PR   24-JAN-2003; 2003US-0442797P.
PR   30-MAY-2003; 2003US-0474413P.
```

```
XX
PA   (LUMP/) LUM P Y.
PA   (TANY/) TAN Y.
PA   (DAIH/) DAI H.
PA   (MUIS/) MUISE E S.
PA   (BERG/) BERGER J P.
PA   (THOM/) THOMPSON J R.
XX
PI   Lum PY,  Tan Y,  Dai H,  Muise ES,  Berger JP,  Thompson JR;
XX
DR   WPI; 2005-313963/32.
DR   GENBANK; U70139.
XX
PT   Determining whether agent has defined biological activity, by making
PT   comparison e.g. comparing efficacy value, toxicity value and classifier
PT   value of agent to reference values, and using results to determine
PT   whether agent has activity.
XX
PS   Example 3; SEQ ID NO 1725; 51pp; English.
XX
CC   The present invention relates to a methods for screening biologically
CC   active agents, such as candidate drugs, to determine whether the agent
CC   possesses a defined biological activity. The methods involve making one
CC   or more comparisons chosen from comparing efficacy value of agent to
CC   reference efficacy value, comparing a toxicity value of the agent to
CC   reference toxicity value, comparing a classifier value of the agent to
CC   reference classifier value, and using the comparison result(s) to
CC   determine whether the agent possesses the defined biological activity.
CC   Also claimed is a population of oligonucleotide probes (I) specific for
CC   measuring the expression levels of members of a classifier population of
CC   genes or a toxicity-related population of genes. (I) are useful for
CC   measuring the expression levels of genes that are useful for identifying
CC   agonists or partial agonists of Peroxisome Proliferator-Activated
CC   Receptor (PPAR) gamma. ADZ62007-ADZ62272 are oligonucleotide probes which
CC   are useful in the method of the invention to measure the expression
CC   pattern of the classifier-related population of mouse genes ADZ61748-
CC   ADZ62006. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from USPTO at seqdata.uspto.gov/sequence.html?DocID=20050084872.
XX
SQ   Sequence 9737 BP; 2569 A; 2228 C; 2243 G; 2686 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.:                0.000155        Length:          9737
Score:                    147.00          Matches:         86
Percent Similarity:       33.7%           Conservative:    53
Best Local Similarity:    20.8%           Mismatches:      161
Query Match:              7.7%            Indels:          113
DB:                       14              Gaps:            16

US-10-757-745-2 (1-362) x ADZ62001 (1-9737)

Qy         35 ValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp--------- 51
Db       7429 ATGGCGTGATGCCCACAGGCAGTGGGGCCCTGGCTCATACCGGAGAGGGACAGTGAAGGC 7488

Qy         52 ----------------------------------------------------TrpGluMetGlu 55
Db       7489 CCTGACTGACAGCTGGCTATGGGACTGTTTCTCCTGCAGTGAGTTCCATGGGAA------ 7542

Qy         56 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 75
Db       7543 --------------------ACGGCACCAGTCGACTCTACAGTGCTCTCGCCAAGACCG 7581

Qy         76 GluThr------------IleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91
Db       7582 TCAACAGCAGTGCAGCTGCTCAGCACCCGGAGTACTTGGTGTCAACTGACCCCGAACATC 7641

Qy         92 ThrThrAspSer---------------------------------------------------- 95
Db       7642 TGGAGCCCATCGATCCTAAGGAACTTCTGGAGGAATGCAGGGCTGTTCTGCACACTCGGC 7701

Qy         96 ------ThrThrSerLysIleSerProSerGluAspThrGln-GlnGluAsnGlySerMe 113

Db       7702 CACCCCGCTACCAGCGGATTTTGTGGACCTGAGGACAGATTGCTCCAGCAGCCACTCTCC 7761

Qy        113 tPheSerLeuIleThrTrpAsnIle-----------------AspGlyLeuAsp----- 125
Db       7762 CATTCGCGTCATGCAGTGGAACATCCTCGCCCAAGCTCTCGGAGAAGGCAAAGACAACTT 7821

Qy        126 -------------------LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
Db       7822 TGTGCAGTGCCCCGTGGAAGCGCTCAAATGGGAAGAGAGGAAGTGCCTGATCCTGGAGGA 7881

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
Db       7882 GATCCTGGCTTACCAGCCAGACATACTGTGCCTCCAGGAAGTG---GACCACTACTTTGA 7938

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
Db       7939 CACCTTCCAGCCACTCCTCAGTAGA------------CTGGGCTACCAAGGCACGTTTTT 7986

Qy        179 e-------------------------------------------------------ThrAl 181
Db       7987 CCCCAAGCCCTGGTCACCATGTCTAGATGTGGAACACAACAACGGTCCAGATGGCTGTGC 8046

Qy        181 aIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSe 201
Db       8047 CTTATTTTTTCTCCAAAACCGATTCAAGCTTATCAGCAGCACCAATATTAGGCTGACAGC 8106

Qy        201 rThrLysMetMetArgAsnLeuLeuCysVal------------HisValAsnValSerGl 217
Db       8107 CATGACCCTAAAAACCAATCAGGTGGCCATCGCACAGACCCTGGAGTGCAAGGAGTCTGG 8166

Qy        217 yAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluAr 237
Db       8167 GCGACAGTTCTGCATTGCTGTCACCCACTTAAAAGCCCGCACTGGCTGGGAGCGGTTCCG 8226

Qy        237 gMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVa 257
Db       8227 GTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTGCAGAACATCACCCAGGGAGCAAAGAT 8286

Qy        257 l------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu------ValThrArgCy 273
Db       8287 CCCCCTGATCGTCTGCGGGGACTTCAACGCAGAGCCAACCGAAGAGGTCTACAAACACTT 8346

Qy        273 sGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------GlyLysPr 290
Db       8347 TGCGTCCTCCAGCCTCAACCTCAACAGCGCCTACAAGCTGCTGAGTCCCGATGGACAGTC 8406

Qy        290 oLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAl 310
Db       8407 GGAGCCTCCCTACACAACCTGG-----------------AAGATCCGGACCTCAGGCGA 8448

Qy        310 aCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIl 330
Db       8449 GTGTCGCCACACGCTGGACTATATCTGGTACTCCAGACATGCTCTGAGT---------GT 8499

Qy        330 eProArgSerLeuAspLeuLeu---------------GlyLeuGluLysLeuAspCysGl 345
Db       8500 CACGTCCGCCCTGGATCTACTCACTGAAGAACAGATTGGGCCCAACCGGCTCCCATCCTT 8559

Qy        345 yArgPheProSerAspHisTrpGlyLeuLeuCysAsn 357
Db       8560 CCATTACCCCTCGGACCACCTGTCCCTGGTGTGTGAC 8596

RESULT 56
ADQ97808
ID   ADQ97808 standard; DNA; 2928 BP.
XX
AC   ADQ97808;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Mouse cancer associated sequence MR11-001, SEQ ID 785.
XX
```

```
KW   Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX
OS   Mus musculus.
XX
PN   WO2004060304-A2.
XX
PD   22-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US041389.
XX
PR   27-DEC-2002; 2002US-00330773.
XX
PA   (SAGR-) SAGRES DISCOVERY INC.
XX
PI   Morris DW,  Malandro MS;
XX
DR   WPI; 2004-543781/52.
XX
PT   New isolated cancer associated nucleic acids comprising at least 10
PT   contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT   cancers such as leukemia and lymphoma.
XX
PS   Claim 1; SEQ ID NO 785; 199pp; English.
XX
CC   The present invention relates to cancer associated sequences (ADQ97025-
CC   ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC   treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic formate directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2928 BP; 679 A; 840 C; 704 G; 685 T; 0 U; 20 Other;

Alignment Scores:
Pred. No.:                0.000159        Length:          2928
Score:                    140.00          Matches:         89
Percent Similarity:       35.0%           Conservative:    55
Best Local Similarity:    21.7%           Mismatches:      163
Query Match:              7.4%            Indels:          104
DB:                       12              Gaps:            19

US-10-757-745-2 (1-362) x ADQ97808 (1-2928)

Qy          12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArg------------------ 25
               :::||||||          |||          :::   |||
Db         666 CAGGCGGCGGCGGCCTCGGGCGCCGCGAGGTCGCGTCCCCGAACAGTGAGTTCCATGGGA 725

Qy          26 ------------ArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
                        ||||||                 :::|||   |||      ||||||
Db         726 AACGGCACCAGTCGACTCTACAGTGCTCTCGCCAAGACCGTCAACAGCAGTGCAGCTGCT 785

Qy          42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
                     :::|||          |||   |||                      :::
Db         786 CAGCACCCGGAGTACTTGGTGTCAACTGACCCCGAA--------------------CAT 824

Qy          62 PheGluPro------ProValGlu--------------GluSerAlaLeuGluArgArg 74
                  ||||||       |||   |||                  :::   |||      |||
Db         825 CTGGAGCCCATCGATCCTAAGGAACTTCTGGAGGAATGCAGGGCTGTTCTGCACACTCGG 884

Qy          75 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 94
               |||                  :::   :::||||||||||                ||||||
Db         885 CCACCCCGCTACCAG------CGGGATTTTGTGGACCTG------------AGGACAGAT 926

Qy          95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114
                  ::::::|||                                          |||
Db         927 TGCTCCAGCAGCCAC----------------------------------TCTCCCATT 950

Qy         115 SerLeuIleThrTrpAsnIle-----------------AspGlyLeuAsp--------- 125
                  ::::::   |||||||||                ::::|||   |||
Db         951 CGCGTCATGCAGTGGAACATCCTCGCCCAAGCTCTCGGAGAAGGCAAAGACAACTTTGTG 1010

Qy         126 --------------LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
                                          ||||||         :::        :::

Db        1011 CAGTGCCCCGTGGAAGCGCTCAAATGGGAAGAGAGGAAGTGCCTGATCCTGGAGGAGATC 1070

Qy         141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
                  |||   ||||||::::::   ||||||||||||          |||:::
Db        1071 CTGGCTTACCAGCCAGACATACTGTGCCTCCAGGAAGTG---GACCACTACTTTGACACC 1127

Qy         161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe--- 179
                  :::          |||                |||::::::      :::|||
Db        1128 TTCCAGCCACTCCTCAGTAGA------------CTGGGCTACCAAGGCACGTTTTTCCCC 1175

Qy         180 ------------------------------------------------ThrAlaIle 182
                                                              |||:::
Db        1176 AAGCCCTGGTCACCATGTCTAGATGTGGAACACAACAACGGTCCAGATGGCTGTGCCTTA 1235

Qy         183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 202
                     ::::::|||   ||||||   |||          |||          :::
Db        1236 TTTTTTCTCCAAAACCGATTCAAGCTTATCAGCAGCACCAATATTAGGCTGACAGCCATG 1295

Qy         203 LysMetMetArgAsnLeuLeuCysVal------------HisValAsnValSerGlyAsn 218
                  :::          |||   :::   :::                      ||||||
Db        1296 ACCCTAAAAACCAATCAGGTGGCCATCGCACAGACCCTGGAGTGCAAGGAGTCTGGGCGA 1355

Qy         219 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
                  :::   |||:::       :::||||||:::::       |||              |||
Db        1356 CAGTTCTGCATTGCTGTCACCCACTTAAAAGCCCGCACTGGCTGGGAGCGGTTCCGGTCA 1415

Qy         239 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal--- 257
                  |||          :::|||:::   :::|||          :::   |||   :::
Db        1416 GCTCAGGGCTGTGACCTCCTCCAGAACCTGCAGAACATCACCCAGGGAGCAAAGATCCCC 1475

Qy         258 ---IlePheAlaGlyAspThrAsnLeuArgAspArgGlu------ValThrArgCysGly 274
                     |||        ||||||   |||                |||
Db        1476 CTGATCGTCTGCGGGGACTTCAACGCAGAGCCAACCGAAGAGGTCTACAAACACTTTGCG 1535

Qy         275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------GlyLysProLys 291
                        |||:::          ::::::   |||          |||:::   :::
Db        1536 TCCTCCAGCCTCAACCTCAACAGCGCCTACAAGCTGCTGAGTCCCGATGGACAGTCGGAG 1595

Qy         292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
                        ||||||                    :::      :::      |||
Db        1596 CCTCCCTACACAACCTGG-----------------AAGATCCGGACCTCAGGCGAGTGT 1637

Qy         312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
                  :::          |||   |||::::::          |||              :::
Db        1638 CGCCACACGCTGGACTATATCTGGTACTCCAGACATGCTCTGAGT---------GTCACG 1688

Qy         332 ArgSerLeuAspLeuLeu--------------GlyLeuGluLysLeuAspCysGlyArg 346
                  :::||||||||||||                ||||      ::::|||
Db        1689 TCCGCCCTGGATCTACTCACTGAAGAACAGATTGGGCCCAACCGGCTCCCATCCTTCCAT 1748

Qy         347 PheProSerAspHisTrpGlyLeuLeuCysAsn 357
                  :::||||||||||||      |||:::|||:::
Db        1749 TACCCCTCGGACCACCTGTCCCTGGTGTGTGAC 1781

RESULT 57
ADE43923
ID   ADE43923 standard; cDNA; 1767 BP.
XX
AC   ADE43923;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Human DP160 encoding cDNA SEQ ID NO:33.
XX
KW   DP119; DP444; DP810; DP685; WE474; DP160; RA977; RA770; antidiabetic;
KW   anorectic; neuroprotective; gene therapy; pancreatic disorder; diabetes;
KW   neurodegenerative disorder; hyperglycaemia; impaired glucose tolerance;
KW   obesity; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             194. .1489
```

```
FT                   /*tag=  a
FT                   /product= "DP160"
XX
PN   WO2003099318-A2.
XX
PD   04-DEC-2003.
XX
PF   30-MAY-2003; 2003WO-EP005700.
XX
PR   29-MAY-2002; 2002EP-00011963.
PR   17-SEP-2002; 2002EP-00020829.
XX
PA   (DEVE-) DEVELOGEN ENTWICKLUNGSBIOLOGISCHE FORSCH.
XX
PI   Dohrmann C,  Austen M;
XX
DR   WPI; 2004-035045/03.
DR   P-PSDB; ADE43924.
XX
PT   Use of DP119, DP444, DP810, DP685, WE474, DP160, RA977 or RA770 nucleic
PT   acid molecules or polypeptides, or their effectors or modulators, for
PT   manufacturing an agent for the diagnosis, prevention or treatment of
PT   pancreatic disorders.
XX
PS   Claim 3; SEQ ID NO 33; 125pp; English.
XX
CC   The present invention describes a nucleic acid molecule (I) selected from
CC   DP119, DP444, DP810, DP685, WE474, DP160, RA977 and RA770, or its encoded
CC   polypeptide, or a fragment or variant of the nucleic acid or polypeptide,
CC   or an effector/modulator of the nucleic acid or polypeptide. (I) is used
CC   in manufacturing a pharmaceutical agent, in monitoring or controlling the
CC   function of a gene or gene product influenced or modified by a DP119,
CC   DP444, DP810, DP685, WE474, DP160, RA977 or RA770 polypeptide, or in
CC   identifying substances that interact with the polypeptide. (I) has
CC   antidiabetic, anorectic and neuroprotective activities, and can be used
CC   in gene therapy. (I) is useful for diagnostic or therapeutic
CC   applications, or for manufacturing an agent for diagnosis, monitoring,
CC   prevention or treatment of pancreatic disorders, such as diabetes and
CC   related disorders, as well as neurodegenerative disorders, and other
CC   diseases. It may also be used for detecting and/or verifying, for the
CC   treatment, alleviation and/or prevention, of a pancreatic dysfunction
CC   (e.g. diabetes, hyperglycaemia or impaired glucose tolerance), and
CC   related disorders including obesity, and neurodegenerative disorders, and
CC   others, in cells, cell masses, organs and/or subjects. It is also used
CC   for regenerating or for promoting the differentiation and/or function of
CC   beta-cells in vitro and/or in vivo. (I) may also be used in preparing a
CC   non-human animal which over- or under-expresses the DP119, DP444, DP810,
CC   DP685, WE474, DP160, RA977 or RA770 gene product. In addition, the
CC   nucleic acid molecule, polypeptide or an effector/modulator of the
CC   nucleic acid or polypeptide, is used in monitoring or controlling the
CC   function of a gene or gene product influenced or modified by a DP119,
CC   DP444, DP810, DP685, WE474, DP160, RA977 or RA770 polypeptide, or in
CC   identifying substances that interact with the polypeptide. The present
CC   sequence is used in the exemplification of the present invention.
XX
SQ   Sequence 1767 BP; 375 A; 538 C; 444 G; 410 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.00161        Length:              1767
Score:                    128.00         Matches:             83
Percent Similarity:       33.0%          Conservative:        55
Best Local Similarity:    19.9%          Mismatches:          164
Query Match:              6.7%           Indels:              116
DB:                       12             Gaps:                18

US-10-757-745-2 (1-362) x ADE43923 (1-1767)

Qy          20 ProGluValLysLysArgArgLeuLeuCysValGluPheAla---------SerValAla 36
               |||                 |||||||||           |||          :::
Db         308 CCCAGGCCCGCATCCCCCCGGCTGCTGGCGGCGGCCTCGGCGGCCTCGGGCGCCGCGAGG 367

Qy          37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
               ||||||          |||      |||   :::                  :::
```

```
Db         368 TCGTGTTCCCGAACAGTG------TGTTCCATGGGAACCGGTACAAGCAGACTCTATAGT 421

Qy          57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
               ||||||                    :::      ||||||      :::   ||||||
Db         422 GCTCTC-----------GCCAAGACACTGAACAGCAGCGCTGCCTCCCAGCACCCAGAG 469

Qy          77 ThrIleSerGluPro---------------------------------------------- 81
                  :::     |||
Db         470 TATTTGGTGTCACCTGACCCAGAGCATCTGGAGCCCATTGATCCTAAAGAGCTTCTTGAG 529

Qy          82 ----------------------------------------LysThrTyrValAspLeu 87
                                                      :::   :::||||||||||
Db         530 GAATGCAGGGCCGTCCTGCACACCCGACCTCCCCGGTTCCAGAGGGATTTTGTGGATCTG 589

Qy          88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 107
                  :::           ||||||       |||
Db         590 AGGACAGAT------TGCCCTAGTACCCACCCACCTATC---------------------- 622

Qy         108 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle------------------- 121
                                 ::::::   |||||||||
Db         623 --------------------AGGGTTATGCAATGGAACATCCTCGCCCAAGCTCTTGGA 661

Qy         122 AspGlyLeuAsp-----------------------LeuAsnAsnLeuSerGluArgAla 133
               :::|||   |||                                 ||||||
Db         662 GAAGGCAAAGACAACTTTGTACAGTGCCCTGTTGAAGCACTCAAATGGGAAGAAAGGAAA 721

Qy         134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
                  :::         :::      |||   ||||||::::::   ||||||||||||
Db         722 TGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTCCAAGAGGTG 781

Qy         154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
                     |||:::          :::              |||                  |||
Db         782 ---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA------------CTAGGC 826

Qy         174 HisGluGluGlyTyrPhe------------------------------------------ 179
               ::::::      :::|||
Db         827 TATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAACACAACAAT 886

Qy         180 ------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
                           |||:::        ::::::|||   ||||||   :::
Db         887 GGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTCAACAGTGCC 946

Qy         196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal------------ 211
                  |||    :::          :::      |||   :::   :::
Db         947 AATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCACAGACCCTG 1006

Qy         212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
                        ||||||   :::   |||:::      :::|||||:::::
Db        1007 GAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAAGCACGCACT 1066

Qy         232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
               |||          |||      |||        :::|||:::    :::|||
Db        1067 GGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTGCAAAACATC 1126

Qy         252 ProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
                  :::   |||   :::       |||      ||||||   |||          |||
Db        1127 ACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAGCCAACAGAA 1186

Qy         270 ------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
                                                  |||:::          ::::::   |||
Db        1187 GAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTACAAGCTGCTG 1246

Qy         288 ---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304
                        |||:::   :::          ||||||
Db        1247 AGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG-----------------AAG 1288

Qy         305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 324
               :::      :::      |||:::          |||   |||::::::          |||
Db        1289 ATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCTAAACATGCT 1348

Qy         325 GluGluGlyHisIleIleProArgSerLeuAspLeuLeu--------------GlyLeu 339
                           :::      :::||||||||||||||                |||
Db        1349 CTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAGATTGGACCC 1399
```

```
Qy        340 GluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsn 357
              :::|||           :::||||||||||||     |||:::|||:::
Db       1400 AACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGTGAC 1453

RESULT 58
ADQ97811
ID   ADQ97811 standard; DNA; 1769 BP.
XX
AC   ADQ97811;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Human cancer associated sequence HR11-001, SEQ ID 788.
XX
KW   Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004060304-A2.
XX
PD   22-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US041389.
XX
PR   27-DEC-2002; 2002US-00330773.
XX
PA   (SAGR-) SAGRES DISCOVERY INC.
XX
PI   Morris DW,  Malandro MS;
XX
DR   WPI; 2004-543781/52.
XX
PT   New isolated cancer associated nucleic acids comprising at least 10
PT   contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT   cancers such as leukemia and lymphoma.
XX
PS   Claim 1; SEQ ID NO 788; 199pp; English.
XX
CC   The present invention relates to cancer associated sequences (ADQ97025-
CC   ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC   treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic formate directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1769 BP; 376 A; 536 C; 444 G; 413 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.00161        Length:          1769
Score:                    128.00         Matches:         83
Percent Similarity:       33.0%          Conservative:    55
Best Local Similarity:    19.9%          Mismatches:      164
Query Match:              6.7%           Indels:          116
DB:                       12             Gaps:            18

US-10-757-745-2 (1-362) x ADQ97811 (1-1769)

Qy         20 ProGluValLysLysArgArgLeuLeuCysValGluPheAla---------SerValAla 36
              |||              |||||||||           |||          :::
Db        308 CCCAGGCCCGCATCCCCCCGGCTGCTGGCGGCGGCCTCGGCGGCCTCGGGCGCCGCGAGG 367

Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
              ||||||         |||      |||  :::                 :::
Db        368 TCGTGTTCCCGAACAGTG------TGTTCCATGGGAACCGGTACAAGCAGACTCTATAGT 421

Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
              ||||||                  :::      ||||||     :::   ||||||
Db        422 GCTCTC------------GCCAAGACACTGAACAGCAGCGCTGCCTCCCAGCACCCAGAG 469

Qy         77 ThrIleSerGluPro---------------------------------------------- 81
                 :::      |||
Db        470 TATTTGGTGTCACCTGACCCAGAGCATCTGGAGCCCATTGATCCTAAAGAGCTTCTTGAG 529

Qy         82 ------------------------------------------LysThrTyrValAspLeu 87
                                                        :::   ::|||||||||
Db        530 GAATGCAGGGCCGTCCTGCACACCCGACCTCCCCGGTTCCAGAGGGATTTTGTGGATCTG 589

Qy         88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 107
              :::          ||||||         |||
Db        590 AGGACAGAT------TGCCCTAGTACCCACCCACCTATC---------------------- 622

Qy        108 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle------------------ 121
                                      ::::::   |||||||||
Db        623 --------------------AGGGTTATGCAATGGAACATCCTCGCCCAAGCTCTTGGA 661

Qy        122 AspGlyLeuAsp-----------------------LeuAsnAsnLeuSerGluArgAla 133
              :::|||   |||                                      ||||||
Db        662 GAAGGCAAAGACAACTTTGTACAGTGCCCTGTTGAAGCACTCAAATGGGAAGAAAGGAAA 721

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
                 :::          :::      |||   ||||||::::::   ||||||||||||
Db        722 TGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTCCAAGAGGTG 781

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
                    |||:::          :::          |||                  |||
Db        782 ---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA------------CTAGGC 826

Qy        174 HisGluGluGlyTyrPhe------------------------------------------ 179
              ::::::      :::|||
Db        827 TATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAACACAACAAT 886

Qy        180 ------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
                           |||:::          ::::::|||    ||||||   :::
Db        887 GGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTCAACAGTGCC 946

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal------------ 211
                 |||          :::      :::       |||   :::   :::
Db        947 AATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCACAGACCCTG 1006

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
                       ||||||   :::   |||:::       :::||||||::::::
Db       1007 GAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAAGCACGCACT 1066

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
              |||              |||       |||          :::|||:::   :::|||
Db       1067 GGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTGCAAAACATC 1126

Qy        252 ProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
                 :::   |||   :::      |||      ||||||   |||              |||
Db       1127 ACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAGCCAACAGAA 1186

Qy        270 ------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
                                                 |||:::          ::::::   |||
Db       1187 GAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTACAAGCTGCTG 1246

Qy        288 ---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304
                       |||:::   :::          ||||||
Db       1247 AGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG-----------------AAG 1288

Qy        305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 324
              :::      :::      |||:::          |||   |||::::::          |||
Db       1289 ATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCTAAACATGCT 1348

Qy        325 GluGluGlyHisIleIleProArgSerLeuAspLeuLeu--------------GlyLeu 339
                          :::      :::|||||||||||||                  |||
Db       1349 CTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAGATTGGACCC 1399

Qy        340 GluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsn 357
              :::|||              :::||||||||||||      |||:::|||:::
Db       1400 AACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGTGAC 1453

RESULT 59
ADZ49510
ID   ADZ49510 standard; DNA; 1971 BP.
XX
```

```
AC   ADZ49510;
XX
DT   30-JUN-2005  (first entry)
XX
DE   Insulin signaling pathway related gene, SEQ ID 839.
XX
KW   Antidiabetic; Gene Therapy; Non-insulin dependent diabetes;
KW   insulin resistance; gene; ds.
XX
OS   Homo sapiens.
XX
PN   US2005085436-A1.
XX
PD   21-APR-2005.
XX
PF   08-JUL-2004; 2004US-00887553.
XX
PR   08-JUL-2003; 2003US-0485883P.
XX
PA   (LIHH/) LI H.
PA   (MAJJ/) MA J.
XX
PI   Li H,  Ma J;
XX
DR   WPI; 2005-305194/31.
XX
PT   Treating, preventing or ameliorating pathological conditions associated
PT   with dysregulation of the insulin signaling pathway (ISP) comprises
PT   administering to a subject an amount of a modulator of any of the
PT   proteins regulated by ISP.
XX
PS   Disclosure; SEQ ID NO 839; 70pp; English.
XX
CC   The present invention relates to a method for treating, preventing or
CC   ameliorating pathological conditions associated with dysregulation of the
CC   insulin signaling pathway (ISP). The method comprises administering to a
CC   subject a modulator for ISP-regulated proteins or a pharmaceutical
CC   composition comprising the described modulator. The method is useful for
CC   treating, preventing or ameliorating pathological conditions associated
CC   with dysregulation of the ISP such as Type II diabetes or Type A syndrome
CC   of insulin resistance. The present sequence is a human homolog of a
CC   Drosophila gene regulated by ISP. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20050085436.
XX
SQ   Sequence 1971 BP; 455 A; 574 C; 481 G; 461 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.00189        Length:          1971
Score:                    128.00         Matches:         83
Percent Similarity:       33.0%          Conservative:    55
Best Local Similarity:    19.9%          Mismatches:      164
Query Match:              6.7%           Indels:          116
DB:                       14             Gaps:            18

US-10-757-745-2 (1-362) x ADZ49510 (1-1971)

Qy          20 ProGluValLysLysArgArgLeuLeuCysValGluPheAla---------SerValAla 36
               |||                |||||||||              |||          :::
Db         308 CCCAGGCCCGCATCCCCCCGGCTGCTGGCGGCGGCCTCGGCGGCCTCGGGCGCCGCGAGG 367

Qy          37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
               ||||||      |||       |||  :::                 :::
Db         368 TCGTGTTCCCGAACAGTG------TGTTCCATGGGAACCGGTACAAGCAGACTCTATAGT 421

Qy          57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
               ||||||                  :::      ||||||      :::   ||||||
Db         422 GCTCTC------------GCCAAGACACTGAACAGCAGCGCTGCCTCCCAGCACCCAGAG 469

Qy          77 ThrIleSerGluPro---------------------------------------------- 81
                  :::      |||
Db         470 TATTTGGTGTCACCTGACCCAGAGCATCTGGAGCCCATTGATCCTAAAGAGCTTCTTGAG 529

Qy          82 ----------------------------------------LysThrTyrValAspLeu 87
                                                       :::   :::||||||||
Db         530 GAATGCAGGGCCGTCCTGCACACCCGACCTCCCCGGTTCCAGAGGGATTTTGTGGATCTG 589

Qy          88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 107
                  :::          ||||||          |||
Db         590 AGGACAGAT------TGCCCTAGTACCCACCCACCTATC---------------------- 622

Qy         108 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle------------------ 121
                                        ::::::   |||||||||
Db         623 --------------------AGGGTTATGCAATGGAACATCCTCGCCCAAGCTCTTGGA 661

Qy         122 AspGlyLeuAsp-----------------------LeuAsnAsnLeuSerGluArgAla 133
               :::|||   |||                                 ||||||
Db         662 GAAGGCAAAGACAACTTTGTACAGTGCCCTGTTGAAGCACTCAAATGGGAAGAAAGGAAA 721

Qy         134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
                  :::         :::          |||  |||||::::::  ||||||||||
Db         722 TGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTCCAAGAGGTG 781

Qy         154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
                  |||:::          :::          |||                  |||
Db         782 ---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA-----------CTAGGC 826

Qy         174 HisGluGluGlyTyrPhe----------------------------------------- 179
               ::::::      :::|||
Db         827 TATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAACACAACAAT 886

Qy         180 -----------ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
                        |||:::          :::::::|||   |||||   :::
Db         887 GGACCAGATGGTTGTGCCTTATTTTTCTTCAAAACCGATTCAAGCTAGTCAACAGTGCC 946

Qy         196 IleIleProProProSerThrLysMetMetArgAsnLeuLeuCysVal------------ 211
                  |||          :::      :::       |||   :::   :::
Db         947 AATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCACAGACCCTG 1006

Qy         212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
                        ||||||   :::   |||:::      :::||||||::::::
Db        1007 GAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAAGCACGCACT 1066

Qy         232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
               |||          |||          |||          :::|||:::   :::|||
Db        1067 GGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTGCAAAACATC 1126

Qy         252 ProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
                  :::   |||   :::      |||       ||||||   |||          |||
Db        1127 ACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAGCCAACAGAA 1186

Qy         270 ------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
                           ||:::          ::::::   |||
Db        1187 GAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTACAAGCTGCTG 1246

Qy         288 ---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304
                        |||:::   :::      ||||||
Db        1247 AGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------------AAG 1288

Qy         305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 324
                  :::      :::      |||:::          |||   |||::::::        |||
Db        1289 ATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCTAAACATGCT 1348

Qy         325 GluGluGlyHisIleIleProArgSerLeuAspLeuLeu--------------GlyLeu 339
                           :::      :::||||||||||||                |||
Db        1349 CTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAGATTGGACCC 1399

Qy         340 GluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsn 357
                  :::|||          :::|||||||||||      |||:::|||:::
Db        1400 AACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGTGAC 1453

RESULT 60
ABL08315
ID   ABL08315 standard; cDNA; 1350 BP.
XX
```

```
AC   ABL08315;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 19427.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC, Adams M, Li PWD, Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB64212.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Claim 1; SEQ ID NO 19427; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1350 BP; 328 A; 376 C; 380 G; 266 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.00124        Length:          1350
Score:                    127.50         Matches:         74
Percent Similarity:       34.3%          Conservative:    53
Best Local Similarity:    20.0%          Mismatches:      108
Query Match:              6.7%           Indels:          135
DB:                       4              Gaps:            19

US-10-757-745-2 (1-362) x ABL08315 (1-1350)

Qy          87 LeuThrAsnGluGluThrThrAsp----------------------SerThrThrSerLys 99
               |||   :::|||:::      |||                     ::::::   |||::: 
Db         262 CTGGCGGACGAGGATGTAGCCGACGGACTGCGATGCCTCAAGCTGAATTCCGTGTCGAGG 321

Qy         100 Ile------SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 117
               :::      :::|||   |||                  |||             |||::: 
Db         322 GTTTGCAGTGCGCCAGTCGAG-----------------GGTGATGACATTCGATTGCTG 363

Qy         118 ThrTrpAsnIle---------------------------AspGlyLeu------------ 124
                  |||||||||                          ||||||
Db         364 CAGTGGAACATCCTATCACAAACTCTCGGGCAGCACAACGATGGATTTGTGCGCTGTCCC 423

Qy         125 ---AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyr 143
                  :::                  |||          :::         :::
Db         424 GAGGAGGCTCTCACCTGGGAGCACCGCAAGTACCTGATTGTCCAGGAGATTCTGCAGAAC 483

Qy         144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
                  |||||||||||   |||||||||||         ::::::   :::|||:::
Db         484 CAGCCCGATGTGATTTGTCTGCAAGAGGTA------GACCACTTCAAGTTCCTGCAGACC 537

Qy         164 ------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe------ 179
                     |||   ||||||                                |||   |||
Db         538 GTGTTGGGTAGCCAGAACTAC--------------------GCAGGCATCTTCTTCCCG 576

Qy         180 ----------------------------------------------------ThrAlaIle 182
                                                                    ||||||
Db         577 AAACCGGATTCCCCGTGCCTGTACATCGAGCAGAACAATGGACCCGATGGCTGTGCCATC 636

Qy         183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 202
                  |||:::   :::::::::|||:::      :::
Db         637 TTCTACAAGAGGGACAAACTGCAGCTGCAGGGATACGACACGAGGATCCTGGAAGTGTGG 696

Qy         203 LysMetMetArgAsnLeuLeuCysVal-----------HisValAsnValSerGlyAsn 218
               ::::::      |||   :::   :::                :::      ||||||
Db         697 CGGGTGCAGAGCAACCAGGTGGCCATCGCGGCCCGGCTGCGTATGCGATCCAGTGGTCGG 756

Qy         219 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis------AlaAlaGlu 236
                  |||   |||:::   |||:::||||||::::::      |||          |||
Db         757 GAATTCTGTGTGGCCACCACCCACTTAAAGGCCCGCCATGGCGCCCTTTTGGCGAAATTA 816

Qy         237 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 256
                  |||       |||   :::   :::::::::   :::::::::
Db         817 CGGAACGAACAGGGACGTGATCTCATTAGGTTTGTGAAGCAGTTTGCCGGAGATACGCCG 876

Qy         257 ValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu---------------------- 269
               ::::::      ||||||   |||          |||
Db         877 CTGCTGCTCTGCGGCGATTTCAATGCCGAACCGGTGGAGCCCATCTATGCCACGATCTTG 936

Qy         270 ---------ValThrArgCysGlyGlyLeu------------------------------ 276
                        :::   |||   |||
Db         937 GGATGTGATCTCCTCCGCTTGGGCAGCGCCTATGCCGACGTGAAGCTAGACCGCGAGGAG 996

Qy         277 ---------ProAsnAsnIleValAspValTrpGluPheLeuGlyLys------------ 289
                        ||||||          ||||||  ||||||:::   |||
Db         997 ATACTACATCCCAAC------GCGGATGTCGGTGAGTTCGTCGCGAAATCGATGAAACGC 1050

Qy         290 ---ProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
                  |||  :::          ||||||
Db        1051 GAGCCACCCTACACC------ACCTGGAAA---------------------------- 1074

Qy         309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328
                                                   |||         ||||||   |||
Db        1075 --------------------------------ATTCGCGAAGAGGGCGAGGAATGTCAC 1101

Qy         329 IleIle------------------ProArgSerLeuAspLeu----------------- 336
                  |||                 |||     |||   :::
Db        1102 ACCATCGACTACGTCTTTTACACACCCGATCGCCTCAAGATCAAAAACTGTCTAGACTTT 1161

Qy         337 ---------------LeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
                            :::|||      :::           ::::::||||||||||||
Db        1162 CCGGCCGGAGAGCAAATTGGCAAGAATCGCACACCCAGCTTCCAATATCCATCGGATCAC 1221

Qy         352 TrpGlyLeuLeuCysAsnLeuAspIleIle 361
               :::   |||:::|||:::   :::::::::
Db        1222 TTTTCCCTGGTGTGTGACTTTGAGCTACTG 1251

RESULT 61
ADQ97807
ID   ADQ97807 standard; DNA; 47040 BP.
XX
AC   ADQ97807;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Mouse cancer associated sequence MD11-001, SEQ ID 784.
XX
KW   Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX
OS   Mus musculus.
XX
```

```
PN   WO2004060304-A2.
XX
PD   22-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US041389.
XX
PR   27-DEC-2002; 2002US-00330773.
XX
PA   (SAGR-) SAGRES DISCOVERY INC.
XX
PI   Morris DW,  Malandro MS;
XX
DR   WPI; 2004-543781/52.
XX
PT   New isolated cancer associated nucleic acids comprising at least 10
PT   contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT   cancers such as leukemia and lymphoma.
XX
PS   Claim 1; SEQ ID NO 784; 199pp; English.
XX
CC   The present invention relates to cancer associated sequences (ADQ97025-
CC   ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC   treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic formate directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 47040 BP; 12069 A; 10729 C; 10594 G; 12790 T; 0 U; 858 Other;

Alignment Scores:
Pred. No.:                0.319          Length:          47040
Score:                    126.00         Matches:         57
Percent Similarity:       37.9%          Conservative:    42
Best Local Similarity:    21.8%          Mismatches:      114
Query Match:              6.6%           Indels:          48
DB:                       12             Gaps:            10

US-10-757-745-2 (1-362) x ADQ97807 (1-47040)

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ||||||         :::         :::       |||   |||||::::::   |||
Db      35153 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 35212

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              |||||||||          |||:::          :::          |||
Db      35213 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 35260

Qy        171 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 179
                    |||::::::      :::|||
Db      35261 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 35317

Qy        180 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
                                  |||:::          ::::::|||   ||||||
Db      35318 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 35377

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 211
              |||          |||         :::       :::       |||   :::   :::
Db      35378 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 35437

Qy        212 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
                          ||||||   :::   |||:::      :::|||||:::
Db      35438 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 35497

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
              :::       |||              |||       |||          :::|||:::   :::
Db      35498 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 35557

Qy        249 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 266
              |||         :::   |||   :::       |||       ||||||   |||
Db      35558 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 35617

Qy        267 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
                  |||                                  |||:::         :::
```

```
Db      35618 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 35677

Qy        285 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
              :::   |||         |||:::   :::            |||||
Db      35678 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 35725

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321
                    :::       :::       |||:::          |||    |||::::::
Db      35726 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 35779

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu------------ 337
                  |||                 :::      :::||||||||||||
Db      35780 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 35830

Qy        338 ---GlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356
                 |||       :::|||             :::||||||||||||      |||:::|||
Db      35831 ATTGGGCCCAACCGGCTCCCATCCTTCCATTACCCCTCGGACCACCTGTCCCTGGTGTGT 35890

Qy        357 Asn 357
              :::
Db      35891 GAC 35893

RESULT 62
ACN44608/c
ID   ACN44608 standard; DNA; 90043 BP.
XX
AC   ACN44608;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Mouse genomic sequence mCG22175.
XX
KW   Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
OS   Mus musculus.
XX
PN   WO2003073826-A2.
XX
PD   12-SEP-2003.
XX
PF   28-FEB-2003; 2003WO-US006235.
XX
PR   01-MAR-2002; 2002US-00087192.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-328604/31.
XX
PT   Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT   comprises a nucleotide sequence.
XX
PS   Claim 1; SEQ ID NO 1141; 0pp; English.
XX
CC   The present invention relates to novel DNA and protein sequences which
CC   are associated with carcinomas. The sequences are useful for: (i) for
CC   screening drug candidates; (ii) for screening of bioactive agent capable
CC   of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC   a bioactive agent capable of modulating the activity of CAP; (iv) for
CC   evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC   carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC   carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC   (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC   determining Carcinoma Associated (CA) gene copy number. In addition, the
CC   CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC   carcinoma including lymphoma. The present sequence is one such CA coding
CC   sequence. Note: This patent is an equivalent to basic patent
CC   US2002182586A1, for which no sequence data was published
XX
SQ   Sequence 90043 BP; 23656 A; 18143 C; 20242 G; 26180 T; 0 U; 1822 Other;

Alignment Scores:
```

```
Pred. No.:                0.821          Length:         90043
Score:                    126.00         Matches:        57
Percent Similarity:       37.9%          Conservative:   42
Best Local Similarity:    21.8%          Mismatches:     114
Query Match:              6.6%           Indels:         48
DB:                       11             Gaps:           10

US-10-757-745-2 (1-362) x ACN44608 (1-90043)

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ||||||         :::         :::      |||   ||||||::::::   |||
Db      85949 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 85890

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              |||||||||          |||:::         :::          |||
Db      85889 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 85842

Qy        171 IleThrGlyHisGluGluGlyTyrPhe-------------------------------- 179
                  |||::::::      :::|||
Db      85841 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 85785

Qy        180 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
                                  |||:::         :::::::|||   ||||||
Db      85784 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 85725

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 211
              |||         |||          :::      :::       |||   :::   :::
Db      85724 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 85665

Qy        212 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
                      ||||||   :::   |||:::      :::|||||:::
Db      85664 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 85605

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
              :::      |||             |||       |||         :::|||:::   :::
Db      85604 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 85545

Qy        249 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 266
              |||     :::   |||   :::       |||      ||||||   |||
Db      85544 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 85485

Qy        267 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
                  |||                                   |||:::         :::
Db      85484 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 85425

Qy        285 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
              :::   |||         |||:::  :::              ||||||
Db      85424 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 85377

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321
                  :::      :::      |||:::        |||   |||::::::
Db      85376 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 85323

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu------------ 337
                  |||               :::      :::||||||||||||
Db      85322 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 85272

Qy        338 ---GlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356
                 |||      :::|||          :::||||||||||||      |||:::|||
Db      85271 ATTGGGCCCAACCGGCTCCCATCCTTCCATTACCCCTCGGACCACCTGTCCCTGGTGTGT 85212

Qy        357 Asn 357
              :::
Db      85211 GAC 85209
```

```
RESULT 63
ADQ97810
ID   ADQ97810 standard; DNA; 49355 BP.
XX
AC   ADQ97810;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Human cancer associated sequence HD11-001, SEQ ID 787.
XX
KW   Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004060304-A2.
XX
PD   22-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US041389.
XX
PR   27-DEC-2002; 2002US-00330773.
XX
PA   (SAGR-) SAGRES DISCOVERY INC.
XX
PI   Morris DW,  Malandro MS;
XX
DR   WPI; 2004-543781/52.
XX
PT   New isolated cancer associated nucleic acids comprising at least 10
PT   contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT   cancers such as leukemia and lymphoma.
XX
PS   Claim 1; SEQ ID NO 787; 199pp; English.
XX
CC   The present invention relates to cancer associated sequences (ADQ97025-
CC   ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC   treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic formate directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 49355 BP; 12474 A; 10173 C; 10092 G; 13597 T; 0 U; 3019 Other;

Alignment Scores:
Pred. No.:                0.733          Length:         49355
Score:                    123.00         Matches:        56
Percent Similarity:       37.9%          Conservative:   43
Best Local Similarity:    21.5%          Mismatches:     114
Query Match:              6.5%           Indels:         48
DB:                       12             Gaps:           10

US-10-757-745-2 (1-362) x ADQ97810 (1-49355)

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ||||||         :::         :::      |||   ||||||::::::   |||
Db      38299 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 38358

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              |||||||||          |||:::         :::          |||
Db      38359 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 38406

Qy        171 IleThrGlyHisGluGluGlyTyrPhe-------------------------------- 179
                  |||::::::      :::|||
Db      38407 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 38463

Qy        180 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
                                  |||:::         :::::::|||   ||||||
Db      38464 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 38523

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 211
              :::         |||          :::      :::       |||   :::   :::
Db      38524 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 38583

Qy        212 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
                      ||||||   :::   |||:::      :::|||||:::
Db      38584 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 38643

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
              :::      |||             |||       |||         :::|||:::   :::
Db      38644 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 38703
```

```
Qy        249 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 266
Db      38704 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 38763

Qy        267 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
Db      38764 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 38823

Qy        285 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
Db      38824 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 38871

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321
Db      38872 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 38925

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu------------ 337
Db      38926 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 38976

Qy        338 ---GlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356
Db      38977 ATTGGACCCAACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGT 39036

Qy        357 Asn 357
Db      39037 GAC 39039
```

RESULT 64
ADS45761
ID ADS45761 standard; cDNA; 1644 BP.
XX
AC ADS45761;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #504.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 24191; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1644 BP; 463 A; 339 C; 388 G; 454 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.0111	Length:	1644
Score:	120.00	Matches:	73
Percent Similarity:	40.7%	Conservative:	65
Best Local Similarity:	21.5%	Mismatches:	123
Query Match:	6.3%	Indels:	78
DB:	13	Gaps:	18

US-10-757-745-2 (1-362) x ADS45761 (1-1644)

```
Qy         53 GluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu--- 71
Db         28 GAAGCCTGGAAAGCATTGCAGCAACATTTTGAG---CAGATAAAAGATACTCACCTGCGA 84

Qy         72 ------------------GluArgArgProGluThrIleSer---------GluProLys 82
Db         85 GATCTGTTTGAACAGGACAAGGCGCGTTTTGCGACATTCTCTGCCACATTTGACGATCAG 144

Qy         83 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 102
Db        145 ATGCTGGTGGATTATTCCAAAAACCGTATTACCGCAGAAACACTGGAGAAATTACAGGCT 204

Qy        103 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 122
Db        205 CTGGCGAAAGAAACTGATTTATCCGGGGCTATTCATAGCATGTTTTCC------------ 252

Qy        123 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 142
Db        253 GGTGAAAAGATTAACCGTACTGAAGATCGCGCT---------GTATTGCATATTGCTTTG 303

Qy        143 ------------TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 158
Db        304 CGTAACCGCAGTAACTCTCCGATTAAGGTTGATGAGGAAGATGTTATGCCGCAGGTCAAT 363

Qy        159 SerTyrLeu---LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGly 177
Db        364 GCGGTACTGGCAAAAATGAAATCATTCAGTGAGCGCATTATTAGTGGGGAGTGGAAAGGT 423

Qy        178 TyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle 197
Db        424 TATACCGGTAAAGCTATTACTGATGTGGTAAACATCGGCATTGGTGGTTCAGATCTTGGC 483

Qy        198 ProPheProSerThr---------LysMetMetArgAsnLeuLeuCysValHisVal--- 213
Db        484 CCATATATGGTGACAGAGGCACTGAAAGCCTATAAGAACCACCTTAATATGCACTTTGTA 543

Qy        214 ---AsnValSerGlyAsnGluLeuCys-------------------------------- 221
```

```
Db        544 TCCAATGTGGATGGCACTCATATCGCGGAAACACTGAAGTCTCTGGATCCGGAAACCACG 603

Qy        222 ----------------------LeuMetThrSerHisLeuGluSerThrArgGlyHisAla 234

Db        604 CTGTTCCTGATTGCATCGAAGACTTTTACCACGCAGGAAACGATGACTAACGCACATTCT 663

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254

Db        664 GCC---CGTGATTGGTTCTTGCAAAGTGCAGTTGATGAAGTACATGTTGCCAAACATTTT 720

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274

Db        721 GCGGCACTT------------TCAACGAAT------GAACAGGAAGTCAGCAAGTTCGGT 762

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294

Db        763 ATTGATCCGCAAAACATGTTTGAGTTCTGGGACTGGGTCGGCGGC------------CGT 810

Qy        295 TyrThr---TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeu 313

Db        811 TATTCTCTTTGG--------------TCTGCGATTGGTTTATCCATTGTCCTGTCCATT 855

Qy        314 ArgPheAspArgIle-------------------------------------PhePhe 320

Db        856 GGATATGATAATTTCGAGCAATTGCTGAATGGTGCCCATGCTATGGATAAGCACTTTGCA 915

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeu 339

Db        916 CAAACTGCTCCTGAACAGAAT------ATTCCGGTTATTCTGGCGCTGATTGGCATC 966

RESULT 65
AAD49594
ID   AAD49594 standard; cDNA; 5299 BP.
XX
AC   AAD49594;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human cytoskeleton-associated protein, CSAP-5 cDNA.
XX
KW   Human; cytoskeleton-associated protein; CSAP-5; atherosclerosis; cancer;
KW   gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             305. .5089
FT                   /*tag=  a
FT                   /product= "Human CSAP-5"
XX
PN   WO200279404-A2.
XX
PD   10-OCT-2002.
XX
PF   25-MAR-2002; 2002WO-US009288.
XX
PR   29-MAR-2001; 2001US-0280508P.
PR   03-APR-2001; 2001US-0281323P.
PR   13-APR-2001; 2001US-0283769P.
PR   04-MAY-2001; 2001US-0288609P.
PR   10-MAY-2001; 2001US-0290518P.
PR   18-MAY-2001; 2001US-0291870P.
PR   29-MAY-2001; 2001US-0294451P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hafalia AJA,  Tang TY,  Yue H,  Khan FA,  Ison CH,  Baughn MR;
PI   Warren BA,  Duggan BM,  Thangavelu K,  Honchell CD,  Azimzai Y;
PI   Elliott VS,  Burford N,  Ding L,  Yue H,  Becha S,  Emerling BM;
PI   Richardson TW,  Lee SY,  Bandman O,  Lal PG,  Lee S,  Gietzen KJ;
PI   Walia NK,  Griffin JA,  Lee EA,  Swarnakar A,  Ring HZ,  Jones KA;
XX
DR   WPI; 2003-092894/08.
DR   P-PSDB; AAE32107.
XX
PT   New human cytoskeleton-associated proteins, useful for preparing a
PT   composition for diagnosing or treating a disease or condition associated
PT   with decreased expression or overexpression of functional CSAP e.g.,
PT   cancer.
XX
PS   Claim 5; Page 207-208; 233pp; English.
XX
CC   The invention relates to new human cytoskeleton-associated protein (CSAP)
CC   and its polynucleotide. The polypeptide is useful for preparing a
CC   composition for diagnosing or treating a disease or condition associated
CC   with decreased expression or overexpression of functional CSAP e.g.
CC   atherosclerosis or cancer. The present sequence is human CSAP-5 cDNA. The
CC   invention is useful in gene therapy
XX
SQ   Sequence 5299 BP; 1772 A; 1043 C; 1255 G; 1229 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.115          Length:          5299
Score:                    117.50         Matches:         76
Percent Similarity:       34.7%          Conservative:    48
Best Local Similarity:    21.3%          Mismatches:      104
Query Match:              6.2%           Indels:          129
DB:                       8              Gaps:            18

US-10-757-745-2 (1-362) x AAD49594 (1-5299)

Qy         10 GlyArgGluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCys 29

Db       3419 GGTCTTCAGGCTGCCCAGAAAGAGGCTCAAATTAAAGTACTGGAAGGTCGACTCAAACAA 3478

Qy         30 ValGluPheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeu------ 47

Db       3479 ACAGAAATAACCAGT--------------GCTACCCAAAACCAGCTCTTATTCCATATG 3523

Qy         48 ------------AlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGlu 63

Db       3524 TTGAAAGAGAAGGCAGAATTAAATCCTGAGCTAGATGCTTTACTAGGCCATGCTTTACAA 3583

Qy         64 ProProValGluGluSerAlaLeuGluArgArgPro----------------------- 75

Db       3584 GATAATGTAGAGGATAGTACTGATGAGGATGCTCCTTTAAACAGCCCAGGATCAGAAGGA 3643

Qy         76 GluThrIleSer------------------------------GluProLysThr------ 83

Db       3644 AGCACGCTGTCTTCAGATCTCATGAAGCTTTGTGGTGAAGTGAAACCTAAGAACAAGGCC 3703

Qy         84 --------------------------------TyrValAspLeuThrAsnGluGluThr 92

Db       3704 CGAAGGAGAACCACCACTCAGATGGAATTGCTGTATGCAGAT---AGCAGTGAACTAGCT 3760

Qy         93 ThrAspSerThrThrSerLysIleSer---ProSerGluAspThrGlnGlnGluAsnGly 111

Db       3761 TCAGACACTAGTACAGGAGATGCCTCCTTGCCTGGCCCTCTCACACCTGTTGCAGAAGGG 3820

Qy        112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131

Db       3821 CAAGAGATTGGAATGAATACA----------------GAGACAAGTGGTACTTCTGCT 3862

Qy        132 ArgAlaArg-----------------GlyValCysSerTyrLeuAlaLeuTyrSerPro 145

Db       3863 AGGGAAAAAGAGCTCTCTCCCCCACCTGGCTTACCTTCTAAGATAGGCAGCATTTCCAGG 3922

Qy        146 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 165

Db       3923 CAGTCATCTCTATCAGAAAAAAAAATTCCAGAGCCTTCTCCTGTAACAAGGAGAAAGGCA 3982

Qy        166 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 185

Db       3983 ------TATGAG-----------------------------------------AAAGCAGAA 3997
```

```
Qy        186 LysSerArgValLysLeuLysSerGlnGluIleIle---ProPheProSerThrLysMet 204
              ||||||:::   |||   :::   |||  ||||||   ||||||||:::::::|||
Db       3998 AAATCAAAGGCCAAGGAACAAAAGCAGGGCATAATCAACCCATTTCCTGCTTCAAAAGGA 4057

Qy        205 MetArg------AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu 222
              :::|||         |||   |||:::|||:::
Db       4058 ATCAGAGCTTTTCCACTTCAGTGTATTCACATA-------------------------- 4090

Qy        223 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 242
                                          ||||||
Db       4091 --------------------GCTGAAGGGCATACA---------------------- 4105

Qy        243 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 262
                   |||  :::             :::      :::::|||   |||
Db       4106 ---------AAAGCTGTGCTCTGTGTGGATTCTACTGATGATCTCCTCTTCACTGGATCA 4156

Qy        263 ThrAsn----------------------LeuArgAspArgGluValThrArgCysGly 274
                 :::                  :::      :::|||:::          |||
Db       4157 AAAGATCGTACTTGTAAAGTATGGAATCTGGTGACTGGGCAGGAAATAATGTCACTGGGG 4216

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294
              |||   |||||||||:::|||   |||                  |||:::|||
Db       4217 GGTCATCCCAACAATGTCGTGTCTGTA-------------------AAATACTGTAAT 4255

Qy        295 TyrThr----------------------------------------TrpAsp 298
              ||||||                                        ||||||
Db       4256 TATACCAGTTTGGTCTTCACTGTATCAACATCTTATATTAAGGTGTGGGAT 4306
```

RESULT 66
ACF62734/c
ID ACF62734 standard; DNA; 96960 BP.
XX
AC ACF62734;
XX
DT 08-OCT-2003 (first entry)
XX
DE Cancer based on CYP3A5 related polynucleotide SEQ ID NO:662.
XX
KW Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;
KW cytochrome p450; subfamily IIIA; nifedipine oxidase; polypeptide 5;
KW cytostatic; gene; ds.
XX
OS Unidentified.
XX
PN WO2003013534-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002WO-EP008219.
XX
PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-268144/26.
XX
PT New use of irinotecan for preparation of compositions for treating cancer
PT in subject having genome with variant allele comprising cytochrome p450,
PT subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.
XX
PS Disclosure; SEQ ID NO 662; 86pp; English.
XX
CC The present invention describes the use of irinotecan (I) or its
CC derivative for the preparation of a pharmaceutical composition for
CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC cancer, or malignant glioma in a subject having a genome with a variant
CC allele which comprises a cytochrome p450, subfamily IIIA (nifedipine
CC oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have
CC cytostatic activity. The therapeutic applications of (I) is improved,
CC since it is possible to individually treat a subject with an appropriate
CC dosage and/or an appropriate derivative of (I). Therefore, undesirable,
CC harmful or toxic effects are efficiently avoided. Unnecessary and
CC potentially harmful treatment of those subjects who do not respond to the
CC treatment with substances (nonresponders), as well as the development of
CC drug resistances due to suboptimal drug dosing can be avoided. ACF62200
CC to ACF62751 and ABM34912 to ABM35013 represent sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 96960 BP; 27407 A; 19882 C; 20450 G; 29221 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9.01	Length:	96960
Score:	117.00	Matches:	59
Percent Similarity:	44.4%	Conservative:	41
Best Local Similarity:	26.2%	Mismatches:	91
Query Match:	6.2%	Indels:	35
DB:	8	Gaps:	9

US-10-757-745-2 (1-362) x ACF62734 (1-96960)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                |||:::
Db      34966 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGAATCA 34908

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                  |||      :::::::::|||   ||||||:::|||       |||||:::
Db      34907 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 34854

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||  :::   |||:::::             |||   |||      :::
Db      34853 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 34797

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                :::::|||      |||   :::::
Db      34796 CAGGAA----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 34755

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||     ::   |||                 ||||||:::   |||
Db      34754 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 34695

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||      ||||||:::   |||:::
Db      34694 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 34635

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
                  ::::::   |||   |||:::          ||||||   ::::::      :::
Db      34634 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 34575

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                 :::|||                |||   |||          :::|||   ||||||
Db      34574 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 34530

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266
                 :::|||       :::|||   |||:::|||      ||||||   |||
Db      34529 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 34473

Qy        267 AspArgGluValThrArgCysGlyGlyLeuProAsnAsn---------------------- 279
                    :::   |||       |||          |||
Db      34472 CTGTCAACATTAGACAGATCAACGGGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 34413

Qy        280 ----------------------IleValAspValTrpGluPheLeuGlyLysProLysHis 292
                                   :::::::|||::::::      |||      ||||||
Db      34412 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 34356

Qy        293 CysGlnTyrThrTrp 297
                ::::|||||||:::
Db      34355 ACAGAATATACATTT 34341
```

RESULT 67
ADB20849/c
ID ADB20849 standard; DNA; 96960 BP.
XX
AC ADB20849;

```
XX
DT   20-NOV-2003  (first entry)
XX
DE   MRP1 based cancer related nucleic acid SEQ ID NO:662.
XX
KW   irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW   lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW   variant allele; multidrug resistance protein 1; MRP1; cytostatic; gene;
KW   ds.
XX
OS   Unidentified.
XX
PN   WO2003013533-A2.
XX
PD   20-FEB-2003.
XX
PF   23-JUL-2002; 2002WO-EP008200.
XX
PR   23-JUL-2001; 2001EP-00117608.
PR   24-MAY-2002; 2002EP-00011710.
XX
PA   (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI   Heinrich G,  Kerb R;
XX
DR   WPI; 2003-354397/33.
XX
PT   Use of irinotecan or its derivative for preparation of a pharmaceutical
PT   composition for treating cancer in a subject having a genome with a
PT   variant allele comprising a multidrug resistance protein 1
PT   polynucleotide.
XX
PS   Disclosure; SEQ ID NO 662; 100pp; English.
XX
CC   The present invention describes a method for the use of irinotecan (I) or
CC   its derivative for the preparation of a pharmaceutical composition for
CC   treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC   cancer, or malignant glioma in a subject having a genome with a variant
CC   allele which comprises a multidrug resistance protein 1 (MRP1)
CC   polynucleotide (II). (I) has cytostatic activity. (I) or its derivative
CC   can be used for the preparation of a pharmaceutical composition for
CC   treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC   cancer, or malignant glioma in a subject, where the subject is a human
CC   (preferably African or Asian) or a mouse. The present sequence represents
CC   a sequence which is used in the exemplification of the present invention.
XX
SQ   Sequence 96960 BP; 27407 A; 19882 C; 20450 G; 29221 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              9.01          Length:          96960
Score:                  117.00        Matches:         59
Percent Similarity:     44.4%         Conservative:    41
Best Local Similarity:  26.2%         Mismatches:      91
Query Match:            6.2%          Indels:          35
DB:                     8             Gaps:            9

US-10-757-745-2 (1-362) x ADB20849 (1-96960)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                ||||::
Db      34966 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGAATCA 34908

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||      ::::::::::|||   ||||||:::|||   .    ||||||:::
Db      34907 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 34854

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::          |||   |||      :::
Db      34853 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 34797

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  ::::::|||      |||   ::::::
Db      34796 CAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 34755

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                ||||||:::   |||
Db      34754 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 34695

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::          |||      ||||||:::   |||:::
Db      34694 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 34635

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::             ||||||   ::::::      :::
Db      34634 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 34575

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||   |||         :::|||   ||||||
Db      34574 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 34530

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266
              :::|||          :::|||   |||:::|||      ||||||   |||
Db      34529 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 34473

Qy        267 AspArgGluValThrArgCysGlyGlyLeuProAsnAsn---------- 279
                    :::   |||      |||      |||
Db      34472 CTGTCAACATTAGACAGATCAACGGGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 34413

Qy        280 ----------IleValAspValTrpGluPheLeuGlyLysProLysHis 292
                       ::::::|||::::::      |||      ||||||
Db      34412 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 34356

Qy        293 CysGlnTyrThrTrp 297
                ::::||||||:::
Db      34355 ACAGAATATACATTT 34341

RESULT 68
ADB87938/c
ID   ADB87938 standard; DNA; 96960 BP.
XX
AC   ADB87938;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human UGT1A1 gene sequence SEQ ID NO:662.
XX
KW   irinotecan; cancer; UGT1A1; cytostatic; topoisomerase I inhibitor;
KW   colorectal cancer; cervical cancer; gastric cancer; lung cancer;
KW   ovarian cancer; pancreatic cancer; malignant glioma;
KW   uridine diphosphate glycosyltransferasel member A1; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003013536-A2.
XX
PD   20-FEB-2003.
XX
PF   23-JUL-2002; 2002WO-EP008217.
XX
PR   23-JUL-2001; 2001EP-00117608.
PR   24-MAY-2002; 2002EP-00011710.
XX
PA   (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI   Heinrich G,  Kerb R;
XX
DR   WPI; 2003-289896/28.
XX
PT   Use of irinotecan to treat cancer patient by determining if patient has
PT   variant alleles of UGT1A1 gene, administering increased/decreased amounts
PT   of irinotecan based on increased/decreased levels of UGT1A1 gene product.
XX
PS   Disclosure; SEQ ID NO 662; 107pp; English.
XX
CC   The invention relates to the novel use of irinotecan to treat a patient
CC   suffering from cancer. This involves determining if the patient has one
```

```
CC   or more variant alleles of the UGT1A1 gene, and if the patient has one or
CC   more of such variant alleles, irinotecan is administered in an increased
CC   or decreased amount in comparison to the amount that is administered
CC   without regard to the patient's alleles in the UGT1A1 gene. The invention
CC   has cytostatic activity. A composition of the invention acts as a
CC   topoisomerase I inhibitor. The method is useful for treating a patient,
CC   an animal e.g. mouse or a human, preferably African or Asian, suffering
CC   from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
CC   pancreatic cancer or malignant glioma. The present sequence is udes in
CC   the exemplification of the invention.
XX
SQ   Sequence 96960 BP; 27407 A; 19882 C; 20450 G; 29221 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.01          Length:          96960
Score:                    117.00        Matches:         59
Percent Similarity:       44.4%         Conservative:    41
Best Local Similarity:    26.2%         Mismatches:      91
Query Match:              6.2%          Indels:          35
DB:                       10            Gaps:            9

US-10-757-745-2 (1-362) x ADB87938 (1-96960)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||  ||||||      :::   ||||||:::                ||||::
Db       34966 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGAATCA 34908

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||::|||      ||||||:::
Db       34907 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 34854

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||  |||   :::  |||::::::              |||   |||      :::
Db       34853 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 34797

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||      |||   ::::::
Db       34796 CAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 34755

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                ||||||:::  |||
Db       34754 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 34695

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::      :::         |||      ||||||:::   |||:::
Db       34694 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 34635

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                  ::::::   |||   |||:::          ||||||   ::::::      :::
Db       34634 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 34575

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::|||          |||  |||          :::|||   ||||||
Db       34574 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 34530

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266
               :::|||          :::|||   |||:::|||      ||||||   |||
Db       34529 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 34473

Qy         267 AspArgGluValThrArgCysGlyGlyLeuProAsnAsn---------------------- 279
                  :::   |||   |||         |||      |||
Db       34472 CTGTCAACATTAGACAGATCAACGGGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 34413

Qy         280 ---------------------IleValAspValTrpGluPheLeuGlyLysProLysHis 292
                                   ::::::|||::::::      |||      ||||||
Db       34412 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 34356

Qy         293 CysGlnTyrThrTrp 297
                  :::||||||:::
Db       34355 ACAGAATATACATTT 34341

RESULT 69
ADB96921/c
ID   ADB96921 standard; DNA; 96960 BP.
XX
AC   ADB96921;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human MDR1 related DNA sequence SEQ ID NO:662.
XX
KW   irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW   lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW   multidrug resistance 1; MDR1; cytostatic; human; Cyp3A5; MRP1; MDR1;
KW   TOP1; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003013537-A2.
XX
PD   20-FEB-2003.
XX
PF   23-JUL-2002; 2002WO-EP008218.
XX
PR   23-JUL-2001; 2001EP-00117608.
PR   24-MAY-2002; 2002EP-00011710.
XX
PA   (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI   Heinrich G,  Kerb R;
XX
DR   WPI; 2003-268145/26.
XX
PT   New use of irinotecan for preparation of pharmaceutical compositions for
PT   treating cancer in subject having genome with variant allele comprising
PT   multidrug resistance 1 polynucleotide.
XX
PS   Disclosure; SEQ ID NO 662; 130pp; English.
XX
CC   The invention relates to the novel use of irinotecan or its derivative
CC   for the preparation of pharmaceutical compositions for treating
CC   colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or
CC   malignant glioma in a subject having a genome with a variant allele which
CC   comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition
CC   of the invention has cytostatic activity. The invention is useful for the
CC   preparation of pharmaceutical compositions for treating colorectal,
CC   cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC   glioma in a subject (preferably human, more preferably African or Asian)
CC   or a mouse. The present sequence is used in the exemplification of the
CC   invention.
XX
SQ   Sequence 96960 BP; 27407 A; 19882 C; 20450 G; 29221 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.01          Length:          96960
Score:                    117.00        Matches:         59
Percent Similarity:       44.4%         Conservative:    41
Best Local Similarity:    26.2%         Mismatches:      91
Query Match:              6.2%          Indels:          35
DB:                       10            Gaps:            9

US-10-757-745-2 (1-362) x ADB96921 (1-96960)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||  ||||||      :::   ||||||:::                ||||::
Db       34966 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGAATCA 34908

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||::|||      ||||||:::
Db       34907 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 34854

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||  |||   :::  |||::::::              |||   |||      :::
Db       34853 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 34797

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
```

```
Db      34796 CAGGAA----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 34755

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db      34754 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 34695

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db      34694 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 34635

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226

Db      34634 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 34575

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db      34574 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 34530

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266

Db      34529 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 34473

Qy        267 AspArgGluValThrArgCysGlyGlyLeuProAsnAsn---------------------- 279

Db      34472 CTGTCAACATTAGACAGATCAACGGGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 34413

Qy        280 ----------------------IleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db      34412 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 34356

Qy        293 CysGlnTyrThrTrp 297

Db      34355 ACAGAATATACATTT 34341

RESULT 70
ADB92112/c
ID   ADB92112 standard; DNA; 96960 BP.
XX
AC   ADB92112;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human MDR1 related DNA sequence SEQ ID NO:662.
XX
KW   irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW   lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW   multidrug resistance 1; MDR1; cytostatic; human; UGT1A1; MRP1; TOP1; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003013535-A2.
XX
PD   20-FEB-2003.
XX
PF   23-JUL-2002; 2002WO-EP008220.
XX
PR   23-JUL-2001; 2001EP-00117608.
PR   24-MAY-2002; 2002EP-00011710.
XX
PA   (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI   Heinrich G,  Kerb R;
XX
DR   WPI; 2003-342400/32.
XX
PT   New use of irinotecan for preparation of pharmaceutical compositions for
PT   treating cancer in subject having genome with variant allele comprising
PT   multidrug resistance 1 polynucleotide.
XX
PS   Disclosure; SEQ ID NO 662; 104pp; English.
XX
CC   The invention relates ro a novel use of irinotecan or its derivative for
CC   the preparation of a pharmaceutical composition for treating colorectal,
CC   cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC   glioma in a subject having a genome with a variant allele which comprises
CC   a multidrug resistance 1 (MDR1) polynucleotide. A composition of the
CC   invention has cytostatic activity. The present sequence is used in the
CC   exemplification of the invention.
XX
SQ   Sequence 96960 BP; 27407 A; 19882 C; 20450 G; 29221 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.01          Length:          96960
Score:                    117.00        Matches:         59
Percent Similarity:       44.4%         Conservative:    41
Best Local Similarity:    26.2%         Mismatches:      91
Query Match:              6.2%          Indels:          35
DB:                       10            Gaps:            9

US-10-757-745-2 (1-362) x ADB92112 (1-96960)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110

Db      34966 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGAATCA 34908

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db      34907 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 34854

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db      34853 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 34797

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db      34796 CAGGAA----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 34755

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db      34754 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 34695

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db      34694 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 34635

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226

Db      34634 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 34575

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db      34574 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 34530

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266

Db      34529 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 34473

Qy        267 AspArgGluValThrArgCysGlyGlyLeuProAsnAsn---------------------- 279

Db      34472 CTGTCAACATTAGACAGATCAACGGGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 34413

Qy        280 ----------------------IleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db      34412 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 34356

Qy        293 CysGlnTyrThrTrp 297

Db      34355 ACAGAATATACATTT 34341

RESULT 71
ADR46193
ID   ADR46193 standard; DNA; 5496 BP.
XX
AC   ADR46193;
XX
```

```
DT   18-NOV-2004  (first entry)
XX
DE   Retrotransposable element, optimised codon usage and polyA site removal.
XX
KW   Retrotransposable element L1; LINE-1; long interspersed element; human;
KW   gene therapy; transposon; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO2004072290-A1.
XX
PD   26-AUG-2004.
XX
PF   06-FEB-2004; 2004WO-US003552.
XX
PR   07-FEB-2003; 2003US-0445945P.
XX
PA   (TULA ) TULANE EDUCATIONAL FUND.
XX
PI   Deininger P,  Belancio VP;
XX
DR   WPI; 2004-625881/60.
XX
PT   New retrotransposable element where one or more polyadenylation sites
PT   have been removed from within the coding regions of its genes, useful in
PT   gene insertion and delivery applications e.g., gene therapy.
XX
PS   Claim 8; SEQ ID NO 2; 52pp; English.
XX
CC   The present sequence is that of retrotransposable element LINE-1 (long
CC   interspersed element-1 or L1) following removal of polyadenylation
CC   signals within the coding regions and optimisation of codon usage for
CC   expression in mammalian cells. This an example of a high efficiency
CC   retrotransposable element of the invention that is characterised by a DNA
CC   sequence that has been engineered to have fewer polyadenylation sites
CC   within the coding regions of its genes compared to the endogenous
CC   element. Retrotransposable elements of the invention may be further
CC   modified to add restriction sites for ease of use ADR46200. A method for
CC   increasing the expression of a gene in a retrotransposable element
CC   comprises identifying one or more putative polyadenylation sites in the
CC   element and mutagenise these to increase production of mRNA transcripts
CC   that comprise sequence located 3' to the mutagenesis polyadenylation
CC   site(s), while maintaining the original amino acid coding within the
CC   genes of the retrotransposable element. The retrotransposable element is
CC   useful for gene insertion and delivery applications, particularly for
CC   insertion of DNA into the genome of a cell or animal and for gene
CC   therapy. It is also useful for gene tagging and mutagenesis in mice. The
CC   retrotransposition potential of a modified retrotransposable element was
CC   compared with that of the wild-type using a standard retrotransposition
CC   experiment cassette in HeLa cells. The experiments supported a modest but
CC   statistically significant stimulation of retrotransposition by the
CC   polyadenylation mutations.
XX
SQ   Sequence 5496 BP; 1504 A; 1755 C; 1525 G; 712 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0.178          Length:          5496
Score:                  116.00         Matches:         66
Percent Similarity:     43.5%          Conservative:    48
Best Local Similarity:  25.2%          Mismatches:      104
Query Match:            6.1%           Indels:          45
DB:                     13             Gaps:            10

US-10-757-745-2 (1-362) x ADR46193 (1-5496)

Qy         64 ProProValGluGluSerAlaLeuGluArgArgProGlu---------------------- 76
              ||||||       |||   |||   |||       ||||||
Db       1249 CCGCCCCGCCCTGAAGGAGCTGCTGAAGGAGGCCCTGAACATGGAGCGCAACAACCGCTA 1308

Qy         77 ---ThrIleSerGluProLysThrTyrVal--AspLeuThrAsnGluGluThrThrAspS 95
                 :::||||||:::     |||  |||       :::   |||   :::   :
Db       1309 CCAGCCCCTGCAGAACCACGCCAAGATGTAAAGACCATCAAGACTAGGAAAGATCTGCAG 1368

Qy         95 erThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheS 115
              ::|||      :::   ||||||:::       |||          :::   |||      :
Db       1369 CAACTAATGAGCAAAA-TCACCAGCTAACACCACCATGACCGGCAGCAACAGCCACATCA 1427

Qy        115 erLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgG 135
              ::::::::|||   ||||||:::|||      ||||||:::      :::|||   |||
Db       1428 CCATCCTGACCCTGAACATCAACGGC------CTGAACAGCGCCATCAAGCGGCACCGG- 1480

Qy        135 lyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleP 155
               :::   |||::::::            |||   |||       :::||||||
Db       1481 --CTGGCCAGCTGGATCAAGAGCCAGGACCCCAGCGTGTGCTGCATCCAGGAG------- 1531

Qy        155 roProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisG 175
                      ::::::|||       |||   ::::::          |||   |||
Db       1532 ----------ACCCACCTGACCTGCCGGGACACCCACCGGCTGAAGATCAAGGGCTGGC 1580

Qy        175 luGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnG 195
               :::   |||              ||||||:::   |||   :::      :
Db       1581 GGAAGATCTACCAGGCCAACGGCAAGCAGAAGAAGGCCGGCGTGGCCATCCTGGTGAGCG 1640

Qy        195 luIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnV 215
               ::       |||       |||||:::   |||:::             ::::::   |
Db       1641 ACAAGACCGACTTCAAGCCCACCAAGATCAAGCGGGACAAGGAGGGCCACTACATCATGG 1700

Qy        215 alSerGlyAsn------------GluLeuCysLeuMetThrSerHisLeuGluSerThrA 231
               ||   |||:::           ||||||   ::::::      :::       :::|||
Db       1701 TGAAGGGCAGCATCCAGCAGGAGGAGCTGACCATCCTGAACATCTACGCCCCCAACACC- 1759

Qy        231 rgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluA 251
                        |||   |||           :::|||   ||||||       :::|||
Db       1760 --------GGCGCCCCCCGGTTC------ATCAAGCAGGTGCTGAGCGACCTGCAGCGGG 1805

Qy        251 laProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu---------------- 265
                  :::|||   |||:::|||       ||||||   |||
Db       1806 ACCTGGACAGCCACACCCTGATCATG---GGCGACTTCAACACCCCCCTGAGCACCCTGG 1862

Qy        266 -------------------------ArgAspArgGluValThrArgCysGlyGlyLeuP 277
                                        :::|||   :::
Db       1863 ACCGGAGCACCCGGCAGAAGGTGAACAAGGACACCCAGGAGCTGAACAGCGCCCTGCACC 1922

Qy        277 roAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrT 297
                 :::::::::|||:::::      |||      ||||||      :::||||||:
Db       1923 AGGCCGACCTGATCGACATCTACCGGACCCTG---CACCCCAAGAGCACCGAGTACACCT 1979

Qy        297 rp 297
               ::
Db       1980 TC 1981

RESULT 72
ADR46200
ID   ADR46200 standard; DNA; 5496 BP.
XX
AC   ADR46200;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Retrotransposable element, optimised codon usage and polyA site removal.
XX
KW   Retrotransposable element L1; LINE-1; long interspersed element; human;
KW   gene therapy; transposon; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO2004072290-A1.
XX
PD   26-AUG-2004.
XX
PF   06-FEB-2004; 2004WO-US003552.
XX
PR   07-FEB-2003; 2003US-0445945P.
XX
```

```
PA   (TULA ) TULANE EDUCATIONAL FUND.
XX
PI   Deininger P,  Belancio VP;
XX
DR   WPI; 2004-625881/60.
XX
PT   New retrotransposable element where one or more polyadenylation sites
PT   have been removed from within the coding regions of its genes, useful in
PT   gene insertion and delivery applications e.g., gene therapy.
XX
PS   Claim 8; SEQ ID NO 9; 52pp; English.
XX
CC   The present sequence is that of retrotransposable element LINE-1 (long
CC   interspersed element-1 or L1) following removal of polyadenylation
CC   signals within the coding regions, optimisation of codon usage for
CC   expression in mammalian cells, and addition of restriction sites for ease
CC   of use. The sequence retains sequences at the start of ORF2 that are
CC   thought to be important for translation of the gene. It is an example of
CC   a high efficiency retrotransposable element of the invention that is
CC   characterised by a DNA sequence that has been engineered to have fewer
CC   polyadenylation sites within the coding regions of its genes compared to
CC   the endogenous element. A method for increasing the expression of a gene
CC   in a retrotransposable element comprises identifying one or more putative
CC   polyadenylation sites in the element and mutagenise these to increase
CC   production of mRNA transcripts that comprise sequence located 3' to the
CC   mutagenesis polyadenylation site(s), while maintaining the original amino
CC   acid coding within the genes of the retrotransposable element. The
CC   retrotransposable element is useful for gene insertion and delivery
CC   applications, particularly for insertion of DNA into the genome of a cell
CC   or animal and for gene therapy. It is also useful for gene tagging and
CC   mutagenesis in mice. The retrotransposition potential of a modified
CC   retrotransposable element was compared with that of the wild-type using a
CC   standard retrotransposition experiment cassette in HeLa cells. The
CC   experiments supported a modest but statistically significant stimulation
CC   of retrotransposition by the polyadenylation mutations.
XX
SQ   Sequence 5496 BP; 1511 A; 1744 C; 1521 G; 720 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.295          Length:          5496
Score:                    114.00         Matches:         66
Percent Similarity:       43.5%          Conservative:    48
Best Local Similarity:    25.2%          Mismatches:      104
Query Match:              6.0%           Indels:          45
DB:                       13             Gaps:            10

US-10-757-745-2 (1-362) x ADR46200 (1-5496)

Qy         64 ProProValGluGluSerAlaLeuGluArgArgProGlu---------------------- 76
              ||||||      |||   |||   |||     ||||||
Db       1249 CCGCCCCGCCCTGAAGGAGCTGCTGAAGGAGGCCCTGAACATGGAGCGCAACAACCGCTA 1308

Qy         77 ---ThrIleSerGluProLysThrTyrVal--AspLeuThrAsnGluGluThrThrAspS 95
                 :::||||||:::       |||  |||       :::   |||   :::   :
Db       1309 CCAGCCCCTGCAGAACCACGCCAAGATGTAAAGACCATCAAGACTAGGAAAGATCTGCAG 1368

Qy         95 erThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheS 115
              ::|||      :::   ||||||:::       |||          :::   |||      :
Db       1369 CAACTAATGAGCAAAA-TCACCAGCTAACACCACCATGACAGGATCAACTTCACACATAA 1427

Qy        115 erLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgG 135
              ::::::::|||   ||||||:::|||      ||||||:::       :::|||   |||
Db       1428 CAATATTAACTTTAAATATAAATGGC------CTGAACAGCGCCATCAAGCGGCACCGG- 1480

Qy        135 lyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleP 155
                 :::   |||::::::              |||   |||       :::||||||
Db       1481 --CTGGCCAGCTGGATCAAGAGCCAGGACCCCAGCGTGTGCTGCATCCAGGAG------- 1531

Qy        155 roProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisG 175
                       ::::::|||       |||   ::::::          |||   |||
Db       1532 -----------ACCCACCTGACCTGCCGGGACACCCACCGGCTGAAGATCAAGGGCTGGC 1580

Qy        175 luGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnG 195
                 :::   |||                        ||||||:::   |||   :::         :
Db       1581 GGAAGATCTACCAGGCCAACGGCAAGCAGAAGAAGGCCGGCGTGGCCATCCTGGTGAGCG 1640

Qy        195 luIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnV 215
              ::          |||       ||||||:::   |||:::             ::::::   |
Db       1641 ACAAGACCGACTTCAAGCCCACCAAGATCAAGCGGGACAAGGAGGGCCACTACATCATGG 1700

Qy        215 alSerGlyAsn-----------GluLeuCysLeuMetThrSerHisLeuGluSerThrA 231
              ||   |||:::            ||||||   ::::::      :::      :::|||
Db       1701 TGAAGGGCAGCATCCAGCAGGAGGAGCTGACCATCCTGAACATCTACGCCCCCAACACC- 1759

Qy        231 rgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluA 251
                       |||   |||          :::|||   ||||||      :::|||
Db       1760 --------GGCGCCCCCCGGTTC------ATCAAGCAGGTGCTGAGCGACCTGCAGCGGG 1805

Qy        251 laProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu---------------- 265
                    :::|||   |||:::|||       |||||   |||
Db       1806 ACCTGGACAGCCACACCCTGATCATG---GGCGACTTCAACACCCCCCTGAGCACCCTGG 1862

Qy        266 -------------------------ArgAspArgGluValThrArgCysGlyGlyLeuP 277
                                        :::|||   :::
Db       1863 ACCGGAGCACCCGGCAGAAGGTGAACAAGGACACCCAGGAGCTGAACAGCGCCCTGCACC 1922

Qy        277 roAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrT 297
                    :::::::::|||:::::       |||       ||||||      :::|||||:
Db       1923 AGGCCGACCTGATCGACATCTACCGGACCCTG---CACCCCAAGAGCACCGAGTACACCT 1979

Qy        297 rp 297
                ::
Db       1980 TC 1981

RESULT 73
AAD49592
ID   AAD49592 standard; cDNA; 5322 BP.
XX
AC   AAD49592;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human cytoskeleton-associated protein, CSAP-3 cDNA.
XX
KW   Human; cytoskeleton-associated protein; CSAP-3; atherosclerosis; cancer;
KW   gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             327. .5171
FT                   /*tag=  a
FT                   /product= "Human CSAP-3"
XX
PN   WO200279404-A2.
XX
PD   10-OCT-2002.
XX
PF   25-MAR-2002; 2002WO-US009288.
XX
PR   29-MAR-2001; 2001US-0280508P.
PR   03-APR-2001; 2001US-0281323P.
PR   13-APR-2001; 2001US-0283769P.
PR   04-MAY-2001; 2001US-0288609P.
PR   10-MAY-2001; 2001US-0290518P.
PR   18-MAY-2001; 2001US-0291870P.
PR   29-MAY-2001; 2001US-0294451P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hafalia AJA,  Tang TY,  Yue H,  Khan FA,  Ison CH,  Baughn MR;
PI   Warren BA,  Duggan BM,  Thangavelu K,  Honchell CD,  Azimzai Y;
PI   Elliott VS,  Burford N,  Ding L,  Yue H,  Becha S,  Emerling BM;
PI   Richardson TW,  Lee SY,  Bandman O,  Lal PG,  Lee S,  Gietzen KJ;
PI   Walia NK,  Griffin JA,  Lee EA,  Swarnakar A,  Ring HZ,  Jones KA;
XX
```

```
DR   WPI; 2003-092894/08.
DR   P-PSDB; AAE32105.
XX
PT   New human cytoskeleton-associated proteins, useful for preparing a
PT   composition for diagnosing or treating a disease or condition associated
PT   with decreased expression or overexpression of functional CSAP e.g.,
PT   cancer.
XX
PS   Claim 5; Page 205-206; 233pp; English.
XX
CC   The invention relates to new human cytoskeleton-associated protein (CSAP)
CC   and its polynucleotide. The polypeptide is useful for preparing a
CC   composition for diagnosing or treating a disease or condition associated
CC   with decreased expression or overexpression of functional CSAP e.g.
CC   atherosclerosis or cancer. The present sequence is human CSAP-3 cDNA. The
CC   invention is useful in gene therapy
XX
SQ   Sequence 5322 BP; 1763 A; 1064 C; 1274 G; 1221 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.469          Length:         5322
Score:                    112.00         Matches:        74
Percent Similarity:       32.9%          Conservative:   45
Best Local Similarity:    20.4%          Mismatches:     111
Query Match:              5.9%           Indels:         132
DB:                       8              Gaps:           17

US-10-757-745-2 (1-362) x AAD49592 (1-5322)

Qy         10 GlyArgGluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCys 29
              |||   :::||||||:::::::|||   :::   :::|||   :::   ||||||
Db       3480 GGTCTTCAGGCTGCCCAGAAAGAGGCTCAAATTAAAGTACTGGAAGGTCGACTCAAACAA 3539

Qy         30 ValGluPheAlaSerValAla---------------------------------------- 36
                 |||     |||
Db       3540 ACAGAAATAACCAGTGCTACCCAAAACCAGCTCTTATTCCATATGTTGAAAGAGAAGGCA 3599

Qy         37 ------------SerCysAspAlaAlaValAlaGlnCysPheLeu--------------- 47
                         ||||||   :::
Db       3600 GAATTAAATCCTGAGCTAGATGCTTTACTAGGCCATGCTTTACAAGATCTAGATAGCGTA 3659

Qy         48 ------------AlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGlu 63
                        |||:::      :::   :::      |||||||||
Db       3660 CCATTAGAAAATGTAGAGGATAGTACTGATGAGGATGCTCCTTTAAACAGC--------- 3710

Qy         64 ProProValGluGluSerAlaLeuGlu------------------------------- 72
              |||     |||   |||   |||
Db       3711 CCAGGATCAGAAGGAAGCACGCTGTCTTCAGATCTCATGAAGCTTTGTGGTGAAGTGAAA 3770

Qy         73 ---------------ArgArgProGluThrIleSerGluProLysThrTyrValAspLeu 87
                           ||||||      |||          |||   |||
Db       3771 CCTAAGAACAAGGCCCGAAGGAGAACCACCACTCAGATGGAATTGCTGTATGCAGAT--- 3827

Qy         88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSer---ProSerGluAspThr 106
              ::::::|||      :::|||::::::|||        `|||   |||         |||
Db       3828 AGCAGTGAACTAGCTTCAGACACTAGTACAGGAGATGCCTCCTTGCCTGGCCCTCTCACA 3887

Qy        107 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 126
                 |||              :::   |||                        :::
Db       3888 CCTGTTGCAGAAGGGCAAGAGATTGGAATGAATACA-----------------GAGACA 3929

Qy        127 AsnAsnLeuSerGluArgAlaArg-----------------GlyValCysSerTyrLeu 140
              :::     |||   |||  :::                 |||:::   |||   :::
Db       3930 AGTGGTACTTCTGCTAGGGAAAAAGAGCTCTCTCCCCCACCTGGCTTACCTTCTAAGATA 3989

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
                 |||                  ::::::   ||||||
Db       3990 GGCAGCATTTCCAGGCAGTCATCTCTATCAGAAAAAAAAATTCCAGAGCCTTCTCCTGTA 4049

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
               ::::::::::::      ||||||
Db       4050 ACAAGGAGAAAGGCA------TATGAG---------------------------------- 4070

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle---ProPhe 199
                       :::||||||:::   |||   :::   |||   ||||||   ||||||
Db       4071 ------AAAGCAGAAAAATCAAAGGCCAAGGAACAAAAGCAGGGCATAATCAACCCATTT 4124

Qy        200 ProSerThrLysMetMetArg------AsnLeuLeuCysValHisValAsnValSerGly 217
              |||:::::::|||   :::|||          |||   |||:::|||:::
Db       4125 CCTGCTTCAAAAGGAATCAGAGCTTTTCCACTTCAGTGTATTCACATA------------ 4172

Qy        218 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 237
                                                        ||||||
Db       4173 ----------------------------------GCTGAAGGGCATACA--------- 4187

Qy        238 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
                                        |||   :::                :::      :::
Db       4188 ----------------------AAAGCTGTGCTCTGTGTGGATTCTACTGATGATCTC 4223

Qy        258 IlePheAlaGlyAspThrAsn----------------------LeuArgAspArgGlu 269
              :::|||   |||     :::                    :::      :::|||
Db       4224 CTCTTCACTGGATCAAAAGATCGTACTTGTAAAGTATGGAATCTGGTGACTGGGCAGGAA 4283

Qy        270 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLys 289
              :::      ||||||   |||||||||:::|||   |||
Db       4284 ATAATGTCACTGGGGGGTCATCCCAACAATGTCGTGTCTGTA------------------ 4325

Qy        290 ProLysHisCysGlnTyrThr---------------------------------------- 296
                 |||:::|||   ||||||
Db       4326 ---AAATACTGTAATTATACCAGTTTGGTCTTCACTGTATCAACATCTTATATTAAGGTG 4382

Qy        297 TrpAsp 298
              ||||||
Db       4383 TGGGAT 4388

RESULT 74
ACA01197
ID   ACA01197 standard; DNA; 1602 BP.
XX
AC   ACA01197;
XX
DT   03-JUN-2003  (first entry)
XX
DE   C. glutamicum derived ORF SEQ ID 1188.
XX
KW   Coryneform; nucleic acid array; fermentation; culture; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   DE10128510-A1.
XX
PD   19-DEC-2002.
XX
PF   13-JUN-2001; 2001DE-01028510.
XX
PR   13-JUN-2001; 2001DE-01028510.
XX
PA   (DEGS ) DEGUSSA AG.
XX
PI   Farwick M,  Moeckel B,  Pfefferle W,  Bathe B,  Huthmacher K;
XX
DR   WPI; 2003-279970/28.
XX
PT   New nucleic acid array useful for monitoring mRNA expression of
PT   Corynebacterium glutamicum during fermentation, comprising nucleic acid
PT   from Corynebacterium glutamicum.
XX
PS   Claim 1; Page 429; 709pp; German.
XX
CC   This invention describes a novel nucleic acid array involving
CC   Corynebacterium glutamicum polynucleotides. The arrays are used to
CC   analyse C. glutamicum, particularly for monitoring a fermentation process
CC   to determine expression levels of C. glutamicum cellular mRNA. Such
CC   monitoring particularly differentiates between expression levels of
CC   different strains of C. glutamicum and allows the adjustment of different
CC   culture and fermentation conditions. ACA00010-ACA02188 represent C.
```

```
CC   glutamicum derived polynucleotides described in the disclosure of the
CC   invention
XX
SQ   Sequence 1602 BP; 394 A; 453 C; 408 G; 347 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.105          Length:          1602
Score:                    111.00         Matches:         99
Percent Similarity:       36.2%          Conservative:    55
Best Local Similarity:    23.2%          Mismatches:      135
Query Match:              5.8%           Indels:          138
DB:                       8              Gaps:            24

US-10-757-745-2 (1-362) x ACA01197 (1-1602)

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPhe---AlaSerValAlaSerCysAsp 39
              |||||||||||||||   ||||||                  ||||||:::      ||:::
Db          7 GAAGTGAAGAAGCGTAATCTCCTCGTAGCTCCCCTCACTGCTTCCCTGGTGTTCTGTAAC 66

Qy         40 AlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsn 59
                 ||||||              |||  |||     |||:::|||
Db         67 TTGGCTGTT--------------GCAGCAAACGCCGTTGAAGTTGAGGCC--------- 102

Qy         60 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 79
                    |||   ||||||      :::   :::|||          :::   :::
Db        103 ---------GAATCACCAGTTGTCATCAATGAAGTTGAATCCAACAGCGACCCAGTTGGT 153

Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr------ 97
              :::         :::|||:::|||   |||   :::      :::|||
Db        154 GAC---------TGGGTGGAGTTGGCTAACACCGACAACAACAACTCCATCGACATTTCC 204

Qy         98 ------SerLysIleSerProSerGluAsp---------------------------- 105
                    |||   :::         ||||||
Db        205 GGTTGGTCCTTAGTCGATGACAAGGAAGACCTGGAAAATGCCCTCGTCCTTCCTGAAGGC 264

Qy        106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
              |||:::   |||:::|||      |||   :::   |||   :::   |||   :::
Db        265 ACTGAGATTGAGTCCGGTGGATACTTTGTTATCTACACCGACTCTGCTGATTACGTGCCT 324

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
                 ||||||              |||         |||       |||
Db        325 ACCAACAAC------------ACCTTTGGTGGCCAGGAATACTTCGGCCTCGGCAAAGAT 372

Qy        146 AspValIlePheLeuGln------------GluValIleProProTyr---TyrSerTyr 160
              |||   :::   |||:::            ||||||:::      |||   :::
Db        373 GACACTGTTACTCTGCGCAACGCTGAAGGCGAAGTAGTTGCTACCTATTCCTGGAAGGAT 432

Qy        161 LeuLysLysArgSerSerAsn---TyrGluIleIle---------------------- 171
              |||   :::   :::   |||   |||       |||
Db        433 CTGGGCGAGCACGCAGAAAACACCTATGGTCGCATCCCAGATATGACTGGTGATTTCGCA 492

Qy        172 ---------------------------------------------------ThrGlyHisGlu 175
                                                                :::|||   |||
Db        493 AACACCGGCGTTCCAACCCCAGGTGCAAAGAATGTTGCTGCTGAAGGCTCCGGCGAAGAA 552

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||                :::::      :::::::::         :::   |||
Db        553 GAAGGC--------------GTTGTTGCAAACGCCCAGCTTCCATTCCACAACGTTGAA 597

Qy        196 IleIlePro-PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVa 215
              |||   ||| |||   |||                                 |||   ::
Db        598 ATCACCCCAATTCACCTCG-----------------------------GTGGAGAT 624

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
              :|||      ::::::|||                 ::::::   |||       |||   ::
Db        625 TTCACCGGTGAAGATATGTCCGGCGTTGATTTCGATGCAA---ACAGCACCGCATGGATC 681

Qy        235 aGluArgMetAsn------------------------------------------------ 239
              :     |||
Db        682 GCCAACAATGACATTGGAAAGATCTACTCCCTCGCCCACGACATAGCTAACAACACCTAC 741

Qy        240 ------------GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
                          |||         :::|||      |||   |||::::::

Db        742 AAGCTGACTGGCGAATGGGAAACCGGCTACCCAGAAGGCGGCGGAGAG-CCAGACGCTGA 800

Qy        255 aThrValIlePheAla------GlyAsp------------------ThrAsnLeuArgAs 267
                  ::::::   |||      ||||||                  |||      ||
Db        801 AGGCATCGTCGCAGCTACCAACGGTGACATCTACCTGTCCACCGAGCGCAACAACGCTGA 860

Qy        267 pArgGluValThrArgCysGlyGlyLeu------------------------------Pr 277
              |::: ||||:::|||      |||
Db        861 CAAGAACGTCTCTCGCCCATCCATCCTGCGTTTTGCTACCCCAACTGGCAAGACTGGCGT 920

Qy        277 oAsnAsnIleValAspValTrp---------GluPheLeuGlyLysProLysHisCysGl 294
                 |||   |||      |||         ||||||:::|||
Db        921 ACAAAACGCAGTTCAGGAATGGGACCTGTCTGAGTTCGTCGGC----------------- 963

Qy        294 nTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuAr 314
                    |||   |||   |||          |||:::   |||     :::|||
Db        964 ----------GACATTCAGCCCAATGGT------GGTCTTGAGGCAATCGCGCAGCTCGA 1007

Qy        314 gPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLe 334
                 |||   ||||||          |||   |||   :::|||         ::
Db       1008 G---GACAACATCTTCGTTGTCGGTGTCGAAGAGACAGGTGATGTCATC------GTTGT 1058

Qy        334 uAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProProSerAspHisTrpGlyLe 354
              :||||||          :::   :::          |||:::   |||          |||::
Db       1059 TGATCTTTCCGCTGACCAGCCAGTTCTGGTTCAAAGGTACGAATCTTCCTTCGACGGTGT 1118

Qy        354 uLeuCysAsnLeuAsp 359
              ::::   :::||||||
Db       1119 CATG---TCGCTTGAT 1131

RESULT 75
AAH67526
ID   AAH67526 standard; DNA; 1608 BP.
XX
AC   AAH67526;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum coding sequence fragment SEQ ID NO: 2561.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-00127688.
XX
PR   16-DEC-1999;   99JP-00377484.
PR   07-APR-2000; 2000JP-00159162.
PR   03-AUG-2000; 2000JP-00280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S,  Mizoguchi H,  Ando S,  Hayashi M,  Ochiai K,  Yokoi H;
PI   Tateishi N,  Senoh A,  Ikeda M,  Ozaki A;
XX
DR   WPI; 2001-376931/40.
DR   P-PSDB; AAG92307.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analyzing
PT   expression profile or pattern of a gene and identifying homologous gene.
XX
PS   Claim 8; SEQ ID NO 2561; 246pp + Sequence Listing; English.
XX
CC   The present invention provides a number of nucleotide and protein
CC   sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC   are useful for identifying the mutation point of a gene derived from a
CC   mutant of coryneform bacterium, measuring expression amount and analysing
```

CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 1608 BP; 397 A; 453 C; 409 G; 349 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.106	Length:	1608
Score:	111.00	Matches:	99
Percent Similarity:	36.2%	Conservative:	55
Best Local Similarity:	23.2%	Mismatches:	135
Query Match:	5.8%	Indels:	138
DB:	5	Gaps:	24

US-10-757-745-2 (1-362) x AAH67526 (1-1608)

```
Qy         21 GluValLysLysArgArgLeuLeuCysValGluPhe---AlaSerValAlaSerCysAsp 39
              |||||||||||||||  ||||||          ||||||:::          ||:::
Db         13 GAAGTGAAGAAGCGTAATCTCCTCGTAGCTCCCCTCACTGCTTCCCTGGTGTTCTGTAAC 72

Qy         40 AlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsn 59
                 ||||||              |||   |||     |||:::|||
Db         73 TTGGCTGTT--------------GCAGCAAACGCCGTTGAAGTTGAGGCC--------- 108

Qy         60 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 79
                     |||   ||||||      :::   :::|||          :::   :::
Db        109 ---------GAATCACCAGTTGTCATCAATGAAGTTGAATCCAACAGCGACCCAGTTGGT 159

Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr------ 97
              :::         :::|||:::|||   |||   :::          ::|||
Db        160 GAC---------TGGGTGGAGTTGGCTAACACCGACAACAACAACTCCATCGACATTTCC 210

Qy         98 ------SerLysIleSerProSerGluAsp------------------------------ 105
                   |||   :::          ||||||
Db        211 GGTTGGTCCTTAGTCGATGACAAGGAAGACCTGGAAAATGCCCTCGTCCTTCCTGAAGGC 270

Qy        106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
              |||:::   |||:::|||       |||   :::   |||   :::   |||   :::
Db        271 ACTGAGATTGAGTCCGGTGGATACTTTGTTATCTACACCGACTCTGCTGATTACGTGCCT 330

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
                 ||||||                 |||       |||       |||
Db        331 ACCAACAAC------------ACCTTTGGTGGCCAGGAATACTTCGGCCTCGGCAAAGAT 378

Qy        146 AspValIlePheLeuGln------------GluValIleProProTyr---TyrSerTyr 160
              |||   :::   |||:::             ||||||:::      |||   :::
Db        379 GACACTGTTACTCTGCGCAACGCTGAAGGCGAAGTAGTTGCTACCTATTCCTGGAAGGAT 438

Qy        161 LeuLysLysArgSerSerAsn---TyrGluIleIle------------------------ 171
              |||   :::   :::   |||   |||      |||
Db        439 CTGGGCGAGCACGCAGAAAACACCTATGGTCGCATCCCAGATATGACTGGTGATTTCGCA 498

Qy        172 ------------------------------------------------ThrGlyHisGlu 175
                                                             :::|||   |||
Db        499 AACACCGGCGTTCCAACCCCAGGTGCAAAGAATGTTGCTGCTGAAGGCTCCGGCGAAGAA 558

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||                ::::::      :::::::::          :::   |||
Db        559 GAAGGC--------------GTTGTTGCAAACGCCCAGCTTCCATTCCACAACGTTGAA 603

Qy        196 IleIlePro-PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVa 215
              |||   ||| |||   |||                                 |||   ::
Db        604 ATCACCCCAATTCACCTCG------------------------------GTGGAGAT 630

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
              :|||      ::::::|||                ::::::   |||      |||   ::
Db        631 TTCACCGGTGAAGATATGTCCGGCGTTGATTTCGATGCAA---ACAGCACCGCATGGATC 687

Qy        235 aGluArgMetAsn---------------------------------------------- 239
              :      |||
Db        688 GCCAACAATGACATTGGAAAGATCTACTCCCTCGCCCACGACATAGCTAACAACACCTAC 747

Qy        240 ------------GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
                          |||         :::|||      |||   |||::::::
Db        748 AAGCTGACTGGCGAATGGGAAACCGGCTACCCAGAAGGCGGCGGAGAG-CCAGACGCTGA 806

Qy        255 aThrValIlePheAla------GlyAsp------------------ThrAsnLeuArgAs 267
                  ::::::   |||      ||||||                    |||      ||
Db        807 AGGCATCGTCGCAGCTACCAACGGTGACATCTACCTGTCCACCGAGCGCAACAACGCTGA 866

Qy        267 pArgGluValThrArgCysGlyGlyLeu----------------------------Pr 277
              |:::   |||:::|||          |||
Db        867 CAAGAACGTCTCTCGCCCATCCATCCTGCGTTTTGCTACCCCAACTGGCAAGACTGGCGT 926

Qy        277 oAsnAsnIleValAspValTrp---------GluPheLeuGlyLysProLysHisCysGl 294
                  |||   |||      |||         ||||||:::|||
Db        927 ACAAAACGCAGTTCAGGAATGGGACCTGTCTGAGTTCGTCGGC----------------- 969

Qy        294 nTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuAr 314
                    |||   |||   |||          |||:::   |||      :::|||
Db        970 ----------GACATTCAGCCCAATGGT------GGTCTTGAGGCAATCGCGCAGCTCGA 1013

Qy        314 gPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLe 334
                  |||   ||||||                |||   |||   :::|||          ::
Db       1014 G---GACAACATCTTCGTTGTCGGTGTCGAAGAGACAGGTGATGTCATC------GTTGT 1064

Qy        334 uAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLe 354
              :||||||          :::   :::          |||:::   |||          |||::
Db       1065 TGATCTTTCCGCTGACCAGCCAGTTCTGGTTCAAAGGTACGAATCTTCCTTCGACGGTGT 1124

Qy        354 uLeuCysAsnLeuAsp 359
              ::::   :::||||||
Db       1125 CATG---TCGCTTGAT 1137
```

Search completed: March 10, 2006, 15:39:04
Job time : 1693.78 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:40:54 ; Search time 2868.4 Seconds
(without alignments)
1724.091 Million cell updates/sec

Title: US-10-757-745-2_COPY_54_140
Perfect score: 448
Sequence: 1 MERALNSYFEPPVEESALER.....IDGLDLNNLSERARGVCSYL 87

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170327_4011/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_5954_@runat_09032006_170327_4011 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	448	100.0	483	6	BD076937	BD076937 5' EST of
2	448	100.0	1261	8	AF201687	AF201687 Homo sapi
3	448	100.0	1920	6	BD205490	BD205490 CD40-Inte
4	448	100.0	1920	6	AR594294	AR594294 Sequence
5	448	100.0	1920	6	AX011599	AX011599 Sequence
6	448	100.0	1921	8	HSA269473	AJ269473 Homo sapi
7	448	100.0	1936	8	AF223469	AF223469 Homo sapi
8	448	100.0	1948	6	AR339398	AR339398 Sequence
9	448	100.0	1952	8	BC017553	BC017553 Homo sapi
10	448	100.0	2499	6	E23195	E23195 Topoisomera
11	439	98.0	752	6	BD150065	BD150065 Primer fo
12	439	98.0	752	6	AX870003	AX870003 Sequence
13	439	98.0	1898	6	BD157138	BD157138 Primer fo
14	439	98.0	1898	6	AX878304	AX878304 Sequence
15	439	98.0	1898	8	AK002168	AK002168 Homo sapi
16	424	94.6	858	6	CQ726600	CQ726600 Sequence
c 17	332	74.1	150344	14	AC152027	AC152027 Dasypus n
18	288.5	64.4	2151	8	HSA420495	AJ420495 Homo sapi
19	288	64.3	20938	8	AY613922	AY613922 Homo sapi
20	288	64.3	102200	8	HS30M3	AL031775 Human DNA
21	288	64.3	167847	8	CR925830	CR925830 Human DNA
22	288	64.3	176819	8	CR942205	CR942205 Human DNA
c 23	269	60.0	411	10	HS133N16S	AL110544 H.sapiens
24	211.5	47.2	1312	6	BD205491	BD205491 CD40-Inte
25	211.5	47.2	1312	6	AR594295	AR594295 Sequence
26	211.5	47.2	1312	6	AX011601	AX011601 Sequence
27	211.5	47.2	1943	9	MMU251328	AJ251328 Mus muscu
28	201	44.9	1236	5	CR761829	CR761829 Xenopus t
29	173	38.6	1451	5	BC083404	BC083404 Danio rer
30	160.5	35.8	256608	9	AL589699	AL589699 Mouse DNA
c 31	157.5	35.2	245394	14	AC125756	AC125756 Rattus no
32	153	34.2	1455	5	BC097117	BC097117 Danio rer
33	148.5	33.1	200076	5	BX511258	BX511258 Zebrafish
34	145.5	32.5	198752	5	BX957279	BX957279 Zebrafish
35	145	32.4	379	6	BD112459	BD112459 EST and e
36	145	32.4	379	6	AR416906	AR416906 Sequence
37	145	32.4	379	6	AX977600	AX977600 Sequence
38	113	25.2	258	6	BD077628	BD077628 5'EST of
39	113	25.2	258	6	AX939201	AX939201 Sequence
40	101	22.5	60	6	CQ544973	CQ544973 Sequence
41	90	20.1	10013	1	DQ073436	DQ073436 Thermotog
c 42	90	20.1	64027	14	AC018192	AC018192 Drosophil
43	90	20.1	84549	14	AC006243	AC006243 Drosophil
44	90	20.1	190616	2	AC007884	AC007884 Drosophil
45	90	20.1	343833	2	AE003465	AE003465 Drosophil
46	88	19.6	229070	14	AC157293	AC157293 Bos tauru
c 47	85	19.0	229070	14	AC157293	AC157293 Bos tauru
48	84	18.8	180129	14	AC074173	AC074173 Mus muscu
49	84	18.8	184800	14	AC074144	AC074144 Mus muscu
50	84	18.8	240203	14	CR974423	CR974423 Mus muscu
51	82.5	18.4	176379	8	AC016751	AC016751 Homo sapi
52	81.5	18.2	3321	6	BD004082	BD004082 Apoptosis
53	81.5	18.2	6763	6	BD004101	BD004101 Apoptosis
54	81.5	18.2	6763	9	AF132726	AF132726 Mus muscu
55	81.5	18.2	138031	9	AL805973	AL805973 Mouse DNA
c 56	80	17.9	233372	14	AC109433	AC109433 Rattus no
57	80	17.9	239241	14	AC128484	AC128484 Rattus no
58	79.5	17.7	110000	1	CP000099_47	Continuation (48 o
c 59	79.5	17.7	143171	15	CLEGCGA	X70810 Euglena gra
60	79.5	17.7	214994	14	AC152575	AC152575 Bos tauru
c 61	79	17.6	479	8	AY228365	AY228365 Homo sapi
c 62	79	17.6	1006	6	CQ728084	CQ728084 Sequence
c 63	79	17.6	1093	6	CQ728085	CQ728085 Sequence
c 64	79	17.6	1473	6	AX571858	AX571858 Sequence
c 65	79	17.6	1597	8	BC101094	BC101094 Homo sapi
c 66	79	17.6	1597	8	BC101095	BC101095 Homo sapi
c 67	79	17.6	1720	8	BC101092	BC101092 Homo sapi
c 68	79	17.6	1720	8	BC101093	BC101093 Homo sapi
c 69	79	17.6	2268	8	AY162215	AY162215 Homo sapi
c 70	79	17.6	2456	6	CQ850796	CQ850796 Sequence
c 71	79	17.6	2456	8	AK127978	AK127978 Homo sapi
c 72	79	17.6	3558	6	CQ845983	CQ845983 Sequence
c 73	79	17.6	3558	8	AK131483	AK131483 Homo sapi
74	78	17.4	3672	8	HUMIGFBP5X	L27560 Homo sapien
c 75	78	17.4	4001	6	CQ787212	CQ787212 Sequence

ALIGNMENTS

```
RESULT 1
BD076937
LOCUS       BD076937                483 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  5' EST of secretory protein expressed in prostate.
ACCESSION   BD076937
VERSION     BD076937.1  GI:22622540
KEYWORDS    JP 2001512013-A/184.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
  TITLE     5' EST of secretory protein expressed in prostate
  JOURNAL   Patent: JP 2001512013-A 184 21-AUG-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001512013-A/184
            PD   21-AUG-2001
            PF   31-JUL-1998 JP 2000505291
            PR   01-AUG-1997 US    08/905144
            PI   JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO  PI
            LACROIX
            PC   C12N15/09,C07K14/47,C12P21/02,C12Q1/02,C12Q1/68,C12N15/00 CC
            blastn
            CC   identity 99
            CC   region 1. .310
            CC   id T26956
            CC   est
            CC   blastn
            CC   identity 98
            CC   region 1. .315
            CC   id T31666
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 137. .267
            CC   id R14990
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 63. .137
            CC   id R14990
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 1. .50
            CC   id R14990
            CC   est
            CC   Von Heijne matrix
            CC   score 5.2
            CC   seq LCVEPASVASCDA/AV
            FH   Key             Location/Qualifiers
            FT   misc_feature    105. .414
            FT   misc_feature    45. .359
            FT   misc_feature    202. .332
            FT   misc_feature    127. .201
            FT   misc_feature    65. .114
            FT   sig_peptide     1. .120.
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              8.23e-46        Length:         483
Score:                  448.00          Matches:        87
```

```
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BD076937 (1-483)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGTTCCTACTTA 420
```

```
RESULT 2
AF201687
LOCUS       AF201687                1261 bp    mRNA    linear   PRI 16-MAY-2003
DEFINITION  Homo sapiens ETS1-associated protein 2 (EAP2) mRNA, complete cds.
ACCESSION   AF201687
VERSION     AF201687.1  GI:11493668
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1261)
  AUTHORS   Pei,H., Yordy,J.S., Leng,Q., Zhao,Q., Watson,D.K. and Li,R.
  TITLE     EAPII interacts with ETS1 and modulates its transcriptional
            function
  JOURNAL   Oncogene 22 (18), 2699-2709 (2003)
   PUBMED   12743594
REFERENCE   2  (bases 1 to 1261)
  AUTHORS   Li,R., Pei,H. and Papas,T.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1999) Center for Molecular and Structural
            Biology, Department of Medicine, and Hollings Cancer Center,
            Medical University of South Carolina, 86 Jonathan Lucas St.,
            Charleston, SC 29425, USA
FEATURES             Location/Qualifiers
     source          1. .1261
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22.1-p22.3"
                     /cell_line="Jurkat"
     gene            1. .1261
                     /gene="EAP2"
     CDS             26. .1114
                     /gene="EAP2"
                     /codon_start=1
                     /product="ETS1-associated protein 2"
                     /protein_id="AAG35600.1"
                     /db_xref="GI:11493669"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEPASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
```

PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.: 2.42e-45 Length: 1261
Score: 448.00 Matches: 87
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x AF201687 (1-1261)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db 185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy 81 ArgGlyValCysSerTyrLeu 87
Db 425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 3
BD205490
LOCUS BD205490 1920 bp DNA linear PAT 17-JUL-2003
DEFINITION CD40-Interacting and TRAF-interacting protein.
ACCESSION BD205490
VERSION BD205490.1 GI:33015260
KEYWORDS JP 2002512796-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
TITLE CD40-Interacting and TRAF-interacting protein
JOURNAL Patent: JP 2002512796-A 1 08-MAY-2002;
VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT OS Homo sapiens (human)
PN JP 2002512796-A/1
PD 08-MAY-2002
PF 28-APR-1999 JP 2000546003
PR 29-APR-1998 EP 98201392.2
PI STEFAN MARIA CHRISTIAAN PYPE,
PI JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI DANNY
FRANCOIS EVELINE HUYLEBROECK
PC C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
A61P37/02,
PC A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
PC G01N33/566,
PC C12N15/00,A61K37/02
CC CD40-Interacting and TRAF-interacting protein. FH Key
Location/Qualifiers
FT CDS (20). .(1108).
FEATURES Location/Qualifiers
source 1. .1920
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.: 3.88e-45 Length: 1920
Score: 448.00 Matches: 87
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BD205490 (1-1920)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db 179 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 238

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 298

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 299 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 358

Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 359 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 418

Qy 81 ArgGlyValCysSerTyrLeu 87
Db 419 CGAGGGGTGTGTTCCTACTTA 439

RESULT 4
AR594294
LOCUS AR594294 1920 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 1 from patent US 6812203.
ACCESSION AR594294
VERSION AR594294.1 GI:56643900
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
TITLE CD40-Interacting and TRAF-interacting proteins
JOURNAL Patent: US 6812203-A 1 02-NOV-2004;
Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
Zwijnaarde;
WOX;
FEATURES Location/Qualifiers
source 1. .1920
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.: 3.88e-45 Length: 1920
Score: 448.00 Matches: 87
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x AR594294 (1-1920)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db 179 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 238

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 298

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 358

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 418

Qy         81 ArgGlyValCysSerTyrLeu 87
                 |||||||||||||||||||||
Db        419 CGAGGGGTGTGTTCCTACTTA 439

RESULT 5
AX011599
LOCUS       AX011599               1920 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 1 from Patent WO9955859.
ACCESSION   AX011599
VERSION     AX011599.1  GI:9998123
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 1 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             20. .1108
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07555.1"
                     /db_xref="GI:9998124"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:              3.88e-45        Length:         1920
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AX011599 (1-1920)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 238

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 298

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 358

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 418

Qy         81 ArgGlyValCysSerTyrLeu 87
                 |||||||||||||||||||||
Db        419 CGAGGGGTGTGTTCCTACTTA 439

RESULT 6
HSA269473
LOCUS       HSA269473              1921 bp    mRNA    linear   PRI 27-MAY-2000
DEFINITION  Homo sapiens mRNA for TRAF and TNF receptor associated protein
            (ttrap gene).
ACCESSION   AJ269473
VERSION     AJ269473.1  GI:8247253
KEYWORDS    TRAF and TNF receptor associated protein; ttrap gene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van
            Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,
            Huylebroeck,D. and Remacle,J.E.
  TITLE     TTRAP, a novel protein that associates with CD40, tumor necrosis
            factor (TNF) receptor-75 and TNF receptor-associated factors
            (TRAFs), and that inhibits nuclear factor-kappa B activation
  JOURNAL   J. Biol. Chem. 275 (24), 18586-18593 (2000)
  PUBMED    10764746
REFERENCE   2  (bases 1 to 1921)
  AUTHORS   Pype,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1999) Pype S., VIB07, Dept. Cell Growth,
            Differentiation and Development, Flanders Interuniversity Institute
            for Biotechnology, CELGEN, K.U.Leuven, Gasthuisberg Campus,
            Herestraat 49, B-3000 Leuven, BELGIUM
COMMENT     Related sequence: AL031775.
FEATURES             Location/Qualifiers
     source          1. .1921
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .1921
                     /gene="ttrap"
     5'UTR           <1. .21
                     /gene="ttrap"
     CDS             22. .1110
                     /gene="ttrap"
                     /function="putative role in signal transduction"
                     /codon_start=1
                     /product="TRAF and TNF receptor associated protein"
                     /protein_id="CAB92966.1"
                     /db_xref="GI:8247254"
                     /db_xref="GOA:O95551"
                     /db_xref="UniProt/TrEMBL:O95551"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
     3'UTR           1111. .1921
                     /gene="ttrap"
     polyA_signal    1900. .1905
                     /gene="ttrap"
ORIGIN

Alignment Scores:
Pred. No.:              3.88e-45        Length:         1921
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
```

```
DB:                    8              Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x HSA269473 (1-1921)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 240

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 300

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 360

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 420

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        421 CGAGGGGTGTGTTCCTACTTA 441
```

```
RESULT 7
AF223469
LOCUS       AF223469               1936 bp    mRNA    linear   PRI 17-APR-2000
DEFINITION  Homo sapiens AD022 protein (AD022) mRNA, complete cds.
ACCESSION   AF223469
VERSION     AF223469.1  GI:7578788
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1936)
  AUTHORS   Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
  TITLE     A novel gene expressed in human adrenal gland
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1936)
  AUTHORS   Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) Chinese National Human Genome Center at
            Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
            Shanghai 201203, China
FEATURES             Location/Qualifiers
     source          1. .1936
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="adrenal gland"
     gene            1. .1936
                     /gene="AD022"
     CDS             17. .1105
                     /gene="AD022"
                     /codon_start=1
                     /product="AD022 protein"
                     /protein_id="AAF64144.1"
                     /db_xref="GI:7578789"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVRFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:              3.92e-45       Length:        1936
Score:                  448.00         Matches:       87
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            100.0%         Indels:        0
DB:                     8              Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x AF223469 (1-1936)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 235

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 295

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 355

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 415

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        416 CGAGGGGTGTGTTCCTACTTA 436
```

```
RESULT 8
AR339398
LOCUS       AR339398               1948 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 889 from patent US 6569662.
ACCESSION   AR339398
VERSION     AR339398.1  GI:33726255
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Tang,Y.T., Zhou,P. and Drmanac,R.T.
  TITLE     Nucleic acids and polypeptides
  JOURNAL   Patent: US 6569662-A 889 27-MAY-2003;
            Hyseq, Inc.; Sunnyvale, CA
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              3.95e-45       Length:        1948
Score:                  448.00         Matches:       87
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            100.0%         Indels:        0
DB:                     6              Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x AR339398 (1-1948)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437
```

```
Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 9
BC017553
LOCUS       BC017553               1952 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein, mRNA (cDNA
            clone MGC:9099 IMAGE:3920790), complete cds.
ACCESSION   BC017553
VERSION     BC017553.2  GI:34782842
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1952)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1952)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Sep 16, 2003 this sequence version replaced gi:17028464.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 15 Row: m Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 23510347.
FEATURES             Location/Qualifiers
     source          1. .1952
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:9099 IMAGE:3920790"
                     /tissue_type="Skin, melanotic melanoma."
                     /clone_lib="NIH_MGC_72"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1952
                     /gene="TTRAP"
                     /note="synonyms: AD022, EAP2, MGC9099, dJ30M3.3"
                     /db_xref="GeneID:51567"
                     /db_xref="MIM:605764"
     CDS             38. .1126
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TRAF and TNF receptor-associated protein"
                     /protein_id="AAH17553.1"
                     /db_xref="GI:17028465"
                     /db_xref="GeneID:51567"
                     /db_xref="MIM:605764"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                3.96e-45        Length:          1952
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BC017553 (1-1952)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 256

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 316

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 376

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 436

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        437 CGAGGGGTGTGTTCCTACTTA 457

RESULT 10
E23195
LOCUS       E23195                 2499 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Topoisomerase binding protein.
ACCESSION   E23195
VERSION     E23195.1  GI:13024277
KEYWORDS    JP 1999075856-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2499)
  AUTHORS   Takashi,T. and Kazuhiko,Y.
  TITLE     Topoisomerase binding protein
```

```
  JOURNAL    Patent: JP 1999075856-A 1 23-MAR-1999;
             TAKASHI TSURUO,CHUGAI PHARMACEUT CO LTD
COMMENT      OS   Unidentified
             PN   JP 1999075856-A/1
             PD   23-MAR-1999
             PF   17-SEP-1997 JP 1997251544
             PR
             PI   TAKASHI TSURUO,KAZUHIKO YAMANE
             PC   C12N15/09,C07K14/47,C12N9/90,C12N15/00
             CC   Strandedness: Double;
             CC   Topology: Linear;
             FH   Key             Location/Qualifiers
             FT   source          1. .2499
             FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2499
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                5.22e-45       Length:         2499
Score:                    448.00         Matches:        87
Percent Similarity:       100.0%         Conservative:   0
Best Local Similarity:    100.0%         Mismatches:     0
Query Match:              100.0%         Indels:         0
DB:                       6              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x E23195 (1-2499)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 700

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 760

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 820

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 880

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        881 CGAGGGGTGTGTTCCTACTTA 901

RESULT 11
BD150065
LOCUS       BD150065                752 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD150065
VERSION     BD150065.1  GI:27855823
KEYWORDS    JP 2002191363-A/4908.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 4908 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/4908
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                     Location/Qualifiers
            FT   source          1. .752
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                1.74e-44       Length:         752
Score:                    439.00         Matches:        86
Percent Similarity:       98.9%          Conservative:   0
Best Local Similarity:    98.9%          Mismatches:     1
Query Match:              98.0%          Indels:         0
DB:                       6              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BD150065 (1-752)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGCTCCTACTTA 420

RESULT 12
AX870003
LOCUS       AX870003                752 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 4908 from Patent EP1074617.
ACCESSION   AX870003
VERSION     AX870003.1  GI:40024866
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 4908 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              1.74e-44        Length:         752
Score:                  439.00          Matches:        86
Percent Similarity:     98.9%           Conservative:   0
Best Local Similarity:  98.9%           Mismatches:     1
Query Match:            98.0%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AX870003 (1-752)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGCTCCTACTTA 420

RESULT 13
BD157138
LOCUS       BD157138                1898 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD157138
VERSION     BD157138.1  GI:27862896
KEYWORDS    JP 2002191363-A/11981.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1898)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 11981 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/11981
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                    Location/Qualifiers
            FT   CDS             (1). .(1086).
FEATURES              Location/Qualifiers
     source           1. .1898
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              4.94e-44        Length:         1898
Score:                  439.00          Matches:        86
Percent Similarity:     98.9%           Conservative:   0
Best Local Similarity:  98.9%           Mismatches:     1
Query Match:            98.0%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BD157138 (1-1898)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGCTCCTACTTA 420

RESULT 14
AX878304
LOCUS       AX878304                1898 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 13209 from Patent EP1074617.
ACCESSION   AX878304
VERSION     AX878304.1  GI:40033040
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 13209 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES              Location/Qualifiers
     source           1. .1898
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
     CDS              1. .1089
                      /note="unnamed protein product"
                      /codon_start=1
                      /protein_id="CAE90095.1"
                      /db_xref="GI:40033041"
                      /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                      CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                      SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                      LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                      ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                      PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                      DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:              4.94e-44        Length:         1898
Score:                  439.00          Matches:        86
Percent Similarity:     98.9%           Conservative:   0
Best Local Similarity:  98.9%           Mismatches:     1
Query Match:            98.0%           Indels:         0
DB:                     6               Gaps:           0
```

US-10-757-745-2_COPY_54_140 (1-87) x AX878304 (1-1898)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGCTCCTACTTA 420
```

RESULT 15
AK002168
LOCUS AK002168 1898 bp mRNA linear PRI 30-JAN-2004
DEFINITION Homo sapiens cDNA FLJ11306 fis, clone PLACE1010031.
ACCESSION AK002168
VERSION AK002168.1 GI:7023882
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
TITLE Complete sequencing and characterization of 21,243 full-length human cDNAs
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
PUBMED 14702039
REFERENCE 2
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1898)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES Location/Qualifiers
source 1. .1898
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1010031"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="cloning vector: pME18SFL3"
CDS 1. .1089
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA92119.1"
/db_xref="GI:7023883"
/translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:

Pred. No.:	4.94e-44	Length:	1898
Score:	439.00	Matches:	86
Percent Similarity:	98.9%	Conservative:	0
Best Local Similarity:	98.9%	Mismatches:	1
Query Match:	98.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x AK002168 (1-1898)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGCTCCTACTTA 420
```

RESULT 16
CQ726600

```
LOCUS       CQ726600                 858 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 12534 from Patent WO02068579.
ACCESSION   CQ726600
VERSION     CQ726600.1  GI:42290170
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 12534 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .858
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                 1.43e-42        Length:          858
Score:                     424.00          Matches:         83
Percent Similarity:        100.0%          Conservative:    0
Best Local Similarity:     100.0%          Mismatches:      0
Query Match:               94.6%           Indels:          0
DB:                        6               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x CQ726600 (1-858)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyVal 83
              ||||||||
Db        400 CGAGGGGTG 408

RESULT 17
AC152027/c
LOCUS       AC152027               150344 bp    DNA     linear   HTG 11-NOV-2004
DEFINITION  Dasypus novemcinctus clone VMRC5-12D21, WORKING DRAFT SEQUENCE, 5
            ordered pieces.
ACCESSION   AC152027
VERSION     AC152027.2  GI:55700150
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Dasypus novemcinctus (nine-banded armadillo)
  ORGANISM  Dasypus novemcinctus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Xenarthra; Dasypodidae; Dasypus.
REFERENCE   1  (bases 1 to 150344)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
            Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
            Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Laric,P.,
            Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L.,
            Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A.,
            Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,
            Shah,K., Sison,C., Stantripop,S., Stephen,E., Thomas,J.W.,
            Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
            Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625
            Fishers Lane, Rockville, MD 20852, USA
COMMENT     On Nov 11, 2004 this sequence version replaced gi:54312159.
            ---------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            ---------- Project Information
                 Center project name: ifi
                 Center clone name: 012D21

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            ---------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 149480 bases at least Q40
                 Consensus quality: 149754 bases at least Q30
                 Consensus quality: 149909 bases at least Q20
                 Insert size: 155000; agarose-fp
                 Insert size: 149944; sum-of-contigs
                 Quality coverage: 9.52x in Q20 bases; agarose-fp
                 Quality coverage: 9.84x in Q20 bases; sum-of-contigs
            ----------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    36273: contig of 36273 bp in length
            *    36274    36373: gap of unknown length
            *    36374    55049: contig of 18676 bp in length
            *    55050    55149: gap of unknown length
            *    55150   101939: contig of 46790 bp in length
            *   101940   102039: gap of unknown length
            *   102040   114806: contig of 12767 bp in length
            *   114807   114906: gap of unknown length
            *   114907   150344: contig of 35438 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150344
                     /organism="Dasypus novemcinctus"
                     /mol_type="genomic DNA"
```

```
                     /db_xref="taxon:9361"
                     /clone="VMRC5-12D21"
                     /clone_lib="VMRC5"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .36273
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    1. .23436
                     /note="clone overlaps with GenBank Accession Number
                     AC152029 clone VMRC5-168B11 (center project name ifj)"
     gap             36274. .36373
                     /estimated_length=unknown
     misc_feature    36374. .55049
                     /note="assembly_fragment"
     gap             55050. .55149
                     /estimated_length=unknown
     misc_feature    55150. .101939
                     /note="assembly_fragment"
     gap             101940. .102039
                     /estimated_length=unknown
     misc_feature    102040. .114806
                     /note="assembly_fragment"
     gap             114807. .114906
                     /estimated_length=unknown
     misc_feature    114907. .150344
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:              1.06e-28        Length:         150344
Score:                  332.00          Matches:        66
Percent Similarity:     83.9%           Conservative:   7
Best Local Similarity:  75.9%           Mismatches:     14
Query Match:            74.1%           Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AC152027 (1-150344)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||   ||||||||||||||||||   ||||||||||||||||   |||
Db      53438 ATGGAAAGAGCGCAGAACTCCTACTTTGAGCCGTGCGTGGAGGAGAGCGCCCCGGAGAGC 53379

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||   |||      :::|||:::      |||||||||:::|||||||:::||||||
Db      53378 CGGCCTCGGACCCCCTGCCAGCCCGAGGCCTGTGTTGACCTAAGCAACGAAGATACGACT 53319

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||      ||||||||||:::   ||||||:::   ||||||||||:::||||||||||
Db      53318 GATTCCAATTGTTCTAAAATCACCTCATCTGAAAATATTCAGCAAGAAGATGGCAGTATG 53259

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db      53258 TTCTCTTTAATTACCTGGAATATTGATGGATTGGATCTAAACAACCTGCAAGAGAGGGCT 53199

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||
Db      53198 CGAGGGGTGTGTTCTTACTTA 53178

RESULT 18
HSA420495
LOCUS       HSA420495               2151 bp    mRNA    linear   PRI 23-NOV-2001
DEFINITION  Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1761756.
ACCESSION   AJ420495
VERSION     AJ420495.1  GI:17066359
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.

REFERENCE   1
  AUTHORS   Auffray,C., Ansorge,W., Ballabio,A., Estivill,X., Gibson,K.,
            Lehrach,H., Poustka,A. and Lundeberg,J.
  TITLE     The European IMAGE consortium for integrated Molecular analysis of
            human gene transcripts
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2151)
  AUTHORS   Persson,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2001) Persson A., Center for Molecular
            Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagsvagen
            30B, 106 91 Stockholm, SWEDEN
COMMENT     This clone is available royalty-free through IMAGE Consortium
            Distributors. IMPORTANT: This sequence represents the full insert
            of this IMAGE cDNA clone. No attempt has been made to verify
            whether this corresponds to the full-length of the original mRNA
            from which it was derived.
FEATURES             Location/Qualifiers
     source          1. .2151
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="EUROIMAGE 1761756"
                     /clone_lib="NCI_CGAP_Kid3"
     polyA_signal    2105. .2110
ORIGIN

Alignment Scores:
Pred. No.:              2.07e-25        Length:         2151
Score:                  288.50          Matches:        76
Percent Similarity:     35.5%           Conservative:   0
Best Local Similarity:  35.5%           Mismatches:     2
Query Match:            64.4%           Indels:         137
DB:                     8               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x HSA420495 (1-2151)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 GAGGCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAG 65

Qy         29 ------------------------------------------------------------ 29

Db         66 ACCTAGTAAGTGATGAGGGGAGCGGGAGCCAGTCCGCACGGAGGCAGTTTGATACCAGGT 125

Qy         29 ------------------------------------------------------------ 29

Db        126 AGCCCTGCGCGTCCTGGGACCTGCCGTGGAGCTGGGCGGACGATGCTGAGGCAGATCTCC 185

Qy         29 ------------------------------------------------------------ 29

Db        186 AAGCTGCGTAGCCAGAGGAGCAGGGCTTGCGTTGCTGCTGCTGCGCGGAGTTGCGCCCCA 245

Qy         29 ------------------------------------------------------------ 29

Db        246 GCAGCGGATCTAACACTGCACCAGACTGTCGCTTCGAACGTGATTTGGGGTCCAGGGAGG 305

Qy         29 ------------------------------------------------------------ 29

Db        306 AAAAGATGGAAATGTAATGTTTAACTCAAATGAAAGGAAATAAGTAAGCGAGAGGTTTGA 365

Qy         29 ------------------------------------------------------------ 29

Db        366 AACAAATGAACGAAGGCTAAACCTTCTATCGAAAAATGGCAACTTGGAAGTCTCCTTGCT 425

Qy         30 ----------------------------------------------ThrTyrValAsp 33
                                                            |||   ||||||
Db        426 TCCTGTATTGTTTATTTATTCTAGGCCCTAGGCTGACCGGCCAGAATAACAGT-GTTGAC 484

Qy         34 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 CTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACT 544
```

```
Qy         54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 73
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 CAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTA 604

Qy         74 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||||||||||||||||||||||||
Db        605 AACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 646

RESULT 19
AY613922
LOCUS       AY613922               20938 bp    DNA     linear   PRI 11-MAY-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein (TTRAP) gene,
            complete cds.
ACCESSION   AY613922
VERSION     AY613922.1  GI:47060296
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 20938)
  AUTHORS   Rieder,M.J., Daniels,R.L., da Ponte,S.H., Hastings,N.C.,
            Ahearn,M.O., Rajkumar,N., Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2004) Genome Sciences, University of Washington,
            1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  SeattleSNPs. NHLBI HL66682 Program
            for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.gs.washington.edu).
FEATURES             Location/Qualifiers
     source          1. .20938
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     repeat_region   1. .32
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     variation       81
                     /frequency="0.67"
                     /replace=""
     repeat_region   87. .368
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     repeat_region   369. .673
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       449
                     /frequency="0.96"
                     /replace=""
     variation       540
                     /frequency="0.01"
                     /replace="t"
     variation       586
                     /frequency="0.01"
                     /replace="a"
     variation       650
                     /frequency="0.03"
                     /replace="g"
     variation       658. .660
                     /frequency="0.01"
                     /replace=""
     repeat_region   683. .989
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       695
                     /frequency="0.18"
                     /replace="t"
     variation       922
                     /frequency="0.02"
                     /replace="t"
     variation       1319
                     /frequency="0.15"
                     /replace="t"
     variation       1456
                     /frequency="0.19"
                     /replace="a"
     variation       1532
                     /frequency="0.01"
                     /replace="c"
     variation       1711. .1713
                     /frequency="0.04"
                     /replace=""
     variation       1795
                     /frequency="0.10"
                     /replace="a"
     variation       1821
                     /frequency="0.02"
                     /replace="t"
     variation       1830. .1862
                     /frequency="0.04"
                     /replace=""
     variation       1913. .1914
                     /frequency="0.13"
                     /replace=""
     gene            <2005. .18938
                     /gene="TTRAP"
     mRNA            join(<2005. .2194,2281. .2366,10172. .10345,11003. .11094,
                     14376. .14494,15753. .15923,17846. .18938)
                     /gene="TTRAP"
                     /product="TRAF and TNF receptor associated protein"
     CDS             join(2030. .2194,2281. .2366,10172. .10345,11003. .11094,
                     14376. .14494,15753. .15923,17846. .18127)
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TRAF and TNF receptor associated protein"
                     /protein_id="AAT09764.1"
                     /db_xref="GI:47060297"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
     variation       2381
                     /gene="TTRAP"
                     /frequency="0.08"
                     /replace="a"
     variation       2767
                     /gene="TTRAP"
                     /frequency="0.08"
                     /replace="g"
     variation       2800
                     /gene="TTRAP"
                     /frequency="0.79"
                     /replace=""
     repeat_region   2919. .3323
                     /rpt_family="L2"
                     /rpt_type=dispersed
     variation       2927
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="g"
     variation       3004
                     /gene="TTRAP"
                     /frequency="0.04"
                     /replace="t"
     variation       3061
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="t"
     variation       3062
                     /gene="TTRAP"
                     /frequency="0.13"
```

```
                /replace="t"
     variation  3114
                /gene="TTRAP"
                /frequency="0.21"
                /replace="c"
     variation  3418
                /gene="TTRAP"
                /frequency="0.03"
                /replace="c"
     variation  3781
                /gene="TTRAP"
                /frequency="0.15"
                /replace="g"
     variation  3988
                /gene="TTRAP"
                /frequency="0.01"
                /replace="a"
     variation  4191
                /gene="TTRAP"
                /frequency="0.01"
                /replace="c"
     variation  4393
                /gene="TTRAP"
                /frequency="0.97"
                /replace=""
     variation  4414
                /gene="TTRAP"
                /frequency="0.03"
                /replace="c"
     variation  4437
                /gene="TTRAP"
                /frequency="0.07"
                /replace="g"
     repeat_region 4446. .4530
                /rpt_family="MER1_type"
                /rpt_type=dispersed
     repeat_region 4559. .4609
                /rpt_family="MER1_type"
                /rpt_type=dispersed
     variation  4565
                /gene="TTRAP"
                /frequency="0.52"
                /replace=""
     repeat_region 4625. .4760
                /rpt_family="MER1_type"
                /rpt_type=dispersed
     variation  4658
                /gene="TTRAP"
                /frequency="0.01"
                /replace="a"
     variation  4861
                /gene="TTRAP"
                /frequency="0.02"
                /replace="a"
     variation  4941
                /gene="TTRAP"
                /frequency="0.07"
                /replace="g"
     variation  4983
                /gene="TTRAP"
                /frequency="0.07"
                /replace="t"
     variation  5024
                /gene="TTRAP"
                /frequency="0.07"
```

Alignment Scores:

Pred. No.:	3.08e-24	Length:	20938
Score:	288.00	Matches:	56
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	64.3%	Indels:	0
DB:	8	Gaps:	0

```
US-10-757-745-2_COPY_54_140 (1-87) x AY613922 (1-20938)

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10173 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 10232

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10233 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 10292

Qy         72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||||||||||||||||||||||||||||
Db      10293 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 10340

RESULT 20
HS30M3
LOCUS       HS30M3                102200 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP1-30M3 on chromosome 6p22.1-22.3
            Contains the 5' end of gene KIAA0319, the gene for TRAF and TNF
            receptor-associated protein (TTRAP)(EAP2, AD022, MGC9099), gene
            FLJ20501, the 3' end of gene FLJ12619 (DKFZP564G182) and two CpG
            islands, complete sequence.
ACCESSION   AL031775
VERSION     AL031775.1  GI:4071041
KEYWORDS    HTG; AD022; CpG island; DKFZP564G182; FLJ12619; FLJ20501; KIAA0319;
            TNF; TRAF; TTRAP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 102200)
  AUTHORS   Phillips,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 29, 1998 this sequence version replaced gi:4007165.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP1-30M3 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .102200
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704M0330"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="p22.1-22.3"
```

```
                /clone="RP1-30M3"
                /clone_lib="RPCI-1"
     gene       join(complement(AL133264.10:4611. .6907),
                complement(AL133264.10:2219. .2395),
                complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-001"
     mRNA       join(complement(AL133264.10:4611. .6907),
                complement(AL133264.10:2219. .2395),
                complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-001"
                /product="chromosome 6 open reading frame 62"
                /note="match: ESTs: AA084220 AA093151 AA168907 AA197971
                AA200765 AA237701 AA292313 AA295166 AA402388 AA478637
                AA591228 AA612170 AA638148 AA647423 AA647435 AA647742
                AA672357 AA798123 AA798467 AA815944 AA833812 AA863774
                AI053561 AI054330 AI122327 AI156738 AI191279 AI194277
                AI284633 AI316639 AI390324 AI458673 AI688647 AI694146
                AI694199 AI788240 AI802319 AI807413 AI831412 AI932363
                AI935446 AI982078 AJ396992 AJ398211 AL529211 AL563388.1
                AL601897.1 AU051650 AU125496 AU135776 AU137791 AU138464
                AV656408 AV686324 AV698092 AV703346 AV708342 AV708873
                AV709145 AV709526 AV739602 AV742411 AV743612 AV744587
                AV745018 AW050065 AW063931 AW270998 AW300107 AW300121
                AW301984 AW302318 AW302718 AW302841 AW302861 AW303029
                AW355281 AW822881 AW859606 AW959243 BB067326 BB123906
                BB609972 BB624305 BB628344 BB654156 BB823708 BE007675
                BE049029 BE049182 BE120619 BE178692 BE178693 BE244058
                BE246922 BE287246 BE289710 BE311050 BE369398 BE377181
                BE386264 BE465032 BE484573 BE554360 BE571717 BE625094
                BE766876 BE766886 BE770856 BE785982.1 BE813002 BE817780
                BE829341 BE829344 BE829355 BE829359 BE893564 BE895461
                BE896141 BE908143 BE913030 BE945891 BE991271 BE992722
                BF015474 BF021751 BF022025 BF029248 BF037603 BF060909
                BF064161 BF109576 BF111215 BF122334 BF137583 BF139381
                BF140373 BF145038 BF148549 BF150174 BF165619 BF207643
                BF326037 BF326041 BF326042 BF326045 BF334390 BF341149
                BF376614 BF378829 BF380539 BF380550 BF465841 BF535383
                BF578793 BF582307 BF748822 BF748938 BF748963 BF765346
                BF766403 BF784199 BF789335 BF789696 BF885392 BF896088
                BF940327 BF984391 BG030354 BG035327 BG088552 BG116557
                BG148023 BG167018 BG168325 BG247693 BG252842 BG295272
                BG295445 BG342500 BG610013 BG655426 BG667600 BG677384.1
                BG678705 BG692520 BG717389 BG801183 BG802707 BG803992
                BG863302 BG914583 BG917993 BG940215 BG943615 BG975946
                BI080618 BI092275 BI106992 BI107793 BI108344 BI222784
                BI258149 BI326830 BI651746 BI659175 BI666111 BI688353
                BI689042 BI694119 BI696541 BI696575 BI696638 BI733610
                BI753721 BI754669 BI851772 BI856689.1 BI872006 BI872428
                BI903238 BI903250 BI960823 BI986171 BM016286 BM069622
                BM069876 BM147734 BM237580 BM238719 BM238848 BM423047
                BM423107 F11344 F20959 H17162 H98057 N56057 R25494 T11219
                T51727 W47233
                match: cDNAs: AB046018.1 AJ420475.1 AK022681.1 AL136632.1
                BC002217.1 BC005537.1"
     gene       join(complement(AL133264.10:101. .211),5119. .5253,
                7768. .8194)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-002"
     mRNA       join(complement(AL133264.10:101. .211),5119. .5253,
                7768. .8194)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-002"
                /product="chromosome 6 open reading frame 62"
                /note="match: ESTs: BG772628 BG772628.1"
     gene       join(complement(AL133264.10:5931. .6069),
                complement(AL133264.10:2219. .2395),
                complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-003"
     mRNA       join(complement(AL133264.10:5931. .6069),
                complement(AL133264.10:2219. .2395),
                complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-003"
                /product="chromosome 6 open reading frame 62"
                /note="match: ESTs: BG342500.1 BG692520.1 BG803992.1"
     CDS        join(complement(AL133264.10:4611. .4739),
                complement(AL133264.10:2219. .2395),
                complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-003"
                /standard_name="OTTHUMP00000016093"
                /note="match: proteins: O95550 Q961D0 Q99K03 Q99LU8 Q9GZU0
                Q9H1Z2 Q9N0B2 Q9QEJ3 Q9W384 Q9XWI7"
                /codon_start=1
                /product="chromosome 6 open reading frame 62"
                /protein_id="CAI20135.1"
                /db_xref="GI:56202469"
                /translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
                KKSALFEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
                DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEFHHGDYEKQFLHV
                LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
                IEDHLRPYMPE"
     CDS        join(complement(AL133264.10:5931. .5972),
                complement(AL133264.10:2219. .2395),
                complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-003"
                /standard_name="OTTHUMP00000016094"
                /codon_start=1
                /product="chromosome 6 open reading frame 62"
                /protein_id="CAI20136.1"
                /db_xref="GI:56202470"
                /db_xref="UniProt/TrEMBL:Q5JVZ2"
                /translation="MSENLSDPVSPVVRKKKSALFEVSEVIPVMTNNYEENILKGVRD
                SSYSLESSLELLQKDVVQLHAPRYQSMRRDVIGCTQEMDFILWPRNDIEKIVCLLFSR
                WKESDEPFRPVQAKFEFHHGDYEKQFLHVLSRKDKTGIVVNNPNQSVFLFIDRQHLQT
                PKNKATIFKLCSICLYLPQEQLTHWAVGTIEDHLRPYMPE"
     polyA_signal 9124. .9129
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-001"
     gene       complement(join(12316. .12571,15963. .16147,46686. .46959))
                /gene="THEM2"
                /locus_tag="RP1-30M3.2-001"
     mRNA       complement(join(12316. .12571,15963. .16147,46686. .46959))
                /gene="THEM2"
                /locus_tag="RP1-30M3.2-001"
                /product="thioesterase superfamily member 2"
                /note="match: ESTs: AA614605 AA928745 AI017897 AI052396
                BF407527 BF475709 BF594161
                match: cDNAs: AF155649.1 AF220186.1 AF274952.1 AK000508.1
                AK002261.1 AK008624.1 BC000894.1"
     gene       complement(join(12316. .12571,15963. .16147,26337. .26628))
                /gene="THEM2"
                /locus_tag="RP1-30M3.2-002"
     mRNA       complement(join(12316. .12571,15963. .16147,26337. .26628))
                /gene="THEM2"
                /locus_tag="RP1-30M3.2-002"
                /product="thioesterase superfamily member 2"
                /note="match: ESTs: BG219890 BG256069"
     polyA_site 12318
     polyA_signal 12336. .12341
     CDS        complement(join(12415. .12571,15963. .16147,46686. .46766))
                /gene="THEM2"
                /locus_tag="RP1-30M3.2-001"
                /standard_name="OTTHUMP00000016090"
                /note="match: proteins: AAA28024 AAH18165 AAL48890 O95549
                P34419 P87304 P93828 P95914 Q18187 Q9ASQ1 Q9NPJ3 Q9VZZ5
                Q9VZZ6"
                /codon_start=1
                /product="thioesterase superfamily member 2"
                /protein_id="CAD92509.1"

Alignment Scores:
```

```
Pred. No.:              1.83e-23        Length:         102200
Score:                  288.00          Matches:        56
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            64.3%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x HS30M3 (1-102200)

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55297 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 55356

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      55357 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 55416

Qy         72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db      55417 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 55464

RESULT 21
CR925830
LOCUS       CR925830               167847 bp    DNA     linear   PRI 08-JAN-2005
DEFINITION  Human DNA sequence from clone RP11-195J19 on chromosome 6, complete
            sequence.
ACCESSION   CR925830
VERSION     CR925830.3  GI:56798018
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 167847)
  AUTHORS   Palmer,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 24, 2004 this sequence version replaced gi:56542150.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-195J19 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES              Location/Qualifiers
     source           1. .167847
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="6"
                      /clone="RP11-195J19"
                      /clone_lib="RPCI-11.1"
ORIGIN

Alignment Scores:
Pred. No.:              3.2e-23         Length:         167847
Score:                  288.00          Matches:        56
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            64.3%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x CR925830 (1-167847)

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      35540 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 35599

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      35600 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 35659

Qy         72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db      35660 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 35707

RESULT 22
CR942205
LOCUS       CR942205               176819 bp    DNA     linear   PRI 29-APR-2005
DEFINITION  Human DNA sequence from clone RP11-948M1 on chromosome 6, complete
            sequence.
ACCESSION   CR942205
VERSION     CR942205.4  GI:62989869
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 176819)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:62551221.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
```

```
          The following abbreviations are used to associate primary accession
          numbers given in the feature table with their source databases:
          Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
          on the WORMPEP database can be found at
          http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
          was generated from part of bacterial clone contigs of human
          chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
          Group.  Further information can be found at
          http://www.sanger.ac.uk/HGP/Chr6
          RP11-948M1 is from the library RPCI-11.4 constructed by the group
          of Pieter de Jong. For further details see
          http://www.chori.org/bacpac/home.htm
          VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .176819
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-948M1"
                     /clone_lib="RPCI-11.4"

ORIGIN

Alignment Scores:
Pred. No.:                3.4e-23        Length:          176819
Score:                    288.00         Matches:         56
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              64.3%          Indels:          0
DB:                       8              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x CR942205 (1-176819)

Qy          32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       29519 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 29578

Qy          52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       29579 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 29638

Qy          72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||||||||||||||||||||||||||||
Db       29639 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 29686

RESULT 23
HS133N16S/c
LOCUS       HS133N16S                411 bp    DNA     linear   STS 02-SEP-1999
DEFINITION  H.sapiens STS from genomic clone 133N16, sequence tagged site.
ACCESSION   AL110544
VERSION     AL110544.1  GI:5830260
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   Mungall,A.J., Huckle,E., Taylor,R. and Hunt,S.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-1999) E-mail contact: humquery@sanger.ac.uk
COMMENT     Marker stdJ133N16SP6 (Primer A : TGTTCCCTCTTTGCTGCAG; Primer B :
            AACAGTTCAAGGTTGTGGGC; amplimer size : 127 bp) is from sequence
            generated from the SP6 end of PAC 133N16. 133N16 is part of the
            bacterial clone contigs constructed by the Chromosome 6 Mapping
            Group.(http://www.sanger.ac.uk/HGP/Chr6/) 133N16 is from the
            library  constructed at the Roswell Park Cancer Institute by the
            group of Pieter de Jong.
            For further details see http://bacpac.med.buffalo.edu/.
FEATURES             Location/Qualifiers
     source          1. .411
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="133N16"
ORIGIN

Alignment Scores:
Pred. No.:                8.18e-24       Length:          411
Score:                    269.00         Matches:         52
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              60.0%          Indels:          0
DB:                       10             Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x HS133N16S (1-411)

Qy          36 AsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGln 55
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         411 AATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAA 352

Qy          56 GluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsn 75
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         351 GAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAAT 292

Qy          76 LeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
               ||||||||||||||||||||||||||||||||||||
Db         291 CTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 256

RESULT 24
BD205491
LOCUS       BD205491                 1312 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  CD40-Interacting and TRAF-interacting protein.
ACCESSION   BD205491
VERSION     BD205491.1  GI:33015261
KEYWORDS    JP 2002512796-A/2.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting protein
  JOURNAL   Patent: JP 2002512796-A 2 08-MAY-2002;
            VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT     OS   Mus musculus (mouse)
            PN   JP 2002512796-A/2
            PD   08-MAY-2002
            PF   28-APR-1999 JP 2000546003
            PR   29-APR-1998 EP   98201392.2
            PI   STEFAN MARIA CHRISTIAAN PYPE,
            PI   JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI   DANNY
            FRANCOIS EVELINE HUYLEBROECK
            PC   C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
              A61P37/02,
            PC   A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
            PC   G01N33/566,
            PC   C12N15/00,A61K37/02
            CC   CD40-Interacting and TRAF-interacting protein. FH   Key
                 Location/Qualifiers
            FT   CDS              (122). .(1234).
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
ORIGIN

Alignment Scores:
Pred. No.:                3.73e-16       Length:          1312
Score:                    211.50         Matches:         45
Percent Similarity:       66.3%          Conservative:    12
Best Local Similarity:    52.3%          Mismatches:      28
Query Match:              47.2%          Indels:          1
```

```
DB:                    6              Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x BD205491 (1-1312)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::||||||::::::|||||||||   |||      :::            |||:::
Db        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||   |||       :::   :::   ||||||||||||||||||||:::       |||
Db        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436

Qy         42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
              :::|||      :::   |||||||||      |||      |||:::   |||
Db        437 ACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGCAGCACTATT 493

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
              |||   ||||||||||||||||||||||||      ||||||   ||||||||||||
Db        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553

Qy         82 GlyValCysSerTyrLeu 87
              ||||||||||||   |||
Db        554 GGGGTGTGTTCCTGCCTA 571

RESULT 25
AR594295
LOCUS       AR594295               1312 bp    DNA     linear   PAT 15-DEC-2004
DEFINITION  Sequence 3 from patent US 6812203.
ACCESSION   AR594295
VERSION     AR594295.1  GI:56643901
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting proteins
  JOURNAL   Patent: US 6812203-A 3 02-NOV-2004;
            Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
            Zwijnaarde;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              3.73e-16        Length:         1312
Score:                  211.50          Matches:        45
Percent Similarity:     66.3%           Conservative:   12
Best Local Similarity:  52.3%           Mismatches:     28
Query Match:            47.2%           Indels:         1
DB:                     6               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x AR594295 (1-1312)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::||||||::::::|||||||||   |||      :::            |||:::
Db        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||   |||       :::   :::   ||||||||||||||||||||:::       |||
Db        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436

Qy         42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
              :::|||      :::   |||||||||      |||      |||:::   |||
Db        437 ACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGCAGCACTATT 493

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
              |||   ||||||||||||||||||||||||      ||||||   ||||||||||||
Db        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553

Qy         82 GlyValCysSerTyrLeu 87
              ||||||||||||   |||
Db        554 GGGGTGTGTTCCTGCCTA 571

RESULT 26
AX011601
LOCUS       AX011601               1312 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 3 from Patent WO9955859.
ACCESSION   AX011601
VERSION     AX011601.1  GI:9998125
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 3 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
     CDS             122. .1234
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07556.1"
                     /db_xref="GI:9998126"
                     /translation="MASGSSSDAAEPAGPAGRAASAPEAAQAEEDRVKRRRLQCLGFA
                     LVGGCDPTMVPSVLRENDWQTQKALSAYFELPENDQGWPRQPPTSFKSEAYVDLTNED
                     ANDTTILEASPSGTPLEDSSTISFITWNIDGLDGCNLPERARGVCSCLALYSPDVVFL
                     QEVIPPYCAYLKKRAASYTIITGNEEGYFTAILLKKGRVKFKSQEIIPFPNTKMMRNL
                     LCVNVSLGGNEFCLMTSHLESTREHSAERIRQLKTVLGKMQEAPDSTTVIFAGDTNLR
                     DQEVIKCGGLPDNVFDAWEFLGKPKHCQYTWDTKANNNLRIPAAYKHRFDRIFFRAEE
                     GHLIPQSLDLVGLEKLDCGRFPSDHWGLLCTLNVVL"
ORIGIN

Alignment Scores:
Pred. No.:              3.73e-16        Length:         1312
Score:                  211.50          Matches:        45
Percent Similarity:     66.3%           Conservative:   12
Best Local Similarity:  52.3%           Mismatches:     28
Query Match:            47.2%           Indels:         1
DB:                     6               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x AX011601 (1-1312)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::||||||::::::|||||||||   |||      :::            |||:::
Db        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||   |||       :::   :::   ||||||||||||||||||||:::       |||
Db        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436

Qy         42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
              :::|||      :::   |||||||||      |||      |||:::   |||
Db        437 ACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGCAGCACTATT 493

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
              |||   ||||||||||||||||||||||||      ||||||   ||||||||||||
Db        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553

Qy         82 GlyValCysSerTyrLeu 87
              ||||||||||||   |||
Db        554 GGGGTGTGTTCCTGCCTA 571

RESULT 27
```

```
MMU251328
LOCUS       MMU251328               1943 bp    mRNA    linear   ROD 27-MAY-2000
DEFINITION  Mus musculus mRNA for putative TRAF and TNF receptor associated
            protein (TTRAP gene).
ACCESSION   AJ251328
VERSION     AJ251328.1  GI:8247282
KEYWORDS    ttrap gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van
            Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,
            Huylebroeck,D. and Remacle,J.E.
  TITLE     TTRAP, a novel protein that associates with CD40, tumor necrosis
            factor (TNF) receptor-75 and TNF receptor-associated factors
            (TRAFs), and that inhibits nuclear factor-kappa B activation
  JOURNAL   J. Biol. Chem. 275 (24), 18586-18593 (2000)
   PUBMED   10764746
REFERENCE   2  (bases 1 to 1943)
  AUTHORS   Pype,S.M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-NOV-1999) Pype S.M.C., VIB07, Department Cell Growth,
            Differentiation and Developme, Flanders Interuniversity Institute
            for Biotechnology, K.U.Leuven, Campus Gasthuisberg, Herestraat 49,
            3000 Leuven, BELGIUM
COMMENT     Related sequence: AJ269473.
FEATURES             Location/Qualifiers
     source          1. .1943
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /note="IMAGE cDNA EST clone 1262914"
     gene            1. .1943
                     /gene="TTRAP"
     CDS             48. .1160
                     /gene="TTRAP"
                     /codon_start=1
                     /product="putative TRAF and TNF receptor associated
                     protein"
                     /protein_id="CAB92971.1"
                     /db_xref="GI:8247283"
                     /db_xref="GOA:Q9JJX7"
                     /db_xref="UniProt/TrEMBL:Q9JJX7"
                     /translation="MASGSSSDAAEPAGPAGRAASAPEAAQAEEDRVKRRRLQCLGFA
                     LVGGCDPTMVPSVLRENDWQTQKALSAYFELPENDQGWPRQPPTSFKSEAYVDLTNED
                     ANDTTILEASPSGTPLEDSSTISFITWNIDGLDGCNLPERARGVCSCLALYSPDVVFL
                     QEVIPPYCAYLKKRAASYTIITGNEEGYFTAILLKKGRVKFKSQEIIPFPNTKMMRNL
                     LCVNVSLGGNEFCLMTSHLESTREHSAERIRQLKTVLGKMQEAPDSTTVIFAGDTNLR
                     DQEVIKCGGLPDNVFDAWEFLGKPKHCQYTWDTKANNNLRIPAAYKHRFDRIFFRAEE
                     GHLIPQSLDLVGLEKLDCGRFPSDHWGLLCTLNVVL"

ORIGIN

Alignment Scores:
Pred. No.:                5.79e-16        Length:          1943
Score:                    211.50          Matches:         45
Percent Similarity:       66.3%           Conservative:    12
Best Local Similarity:    52.3%           Mismatches:      28
Query Match:              47.2%           Indels:          1
DB:                       9               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x MMU251328 (1-1943)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::||||||::::::|||||||||   |||      :::              |||:::
Db        243 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 302

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||   |||      :::   :::   ||||||||||||||||||||::::      |||
Db        303 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 362

Qy         42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
              :::|||      :::   |||||||||      |||      |||:::   |||
Db        363 ACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGCAGCACTATT 419

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
              |||   ||||||||||||||||||||||||||||      ||||||   ||||||||||||
Db        420 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 479

Qy         82 GlyValCysSerTyrLeu 87
              ||||||||||||   |||
Db        480 GGGGTGTGTTCCTGCCTA 497

RESULT 28
CR761829
LOCUS       CR761829                1236 bp    mRNA    linear   VRT 15-SEP-2004
DEFINITION  Xenopus tropicalis finished cDNA, clone TGas102g02.
ACCESSION   CR761829
VERSION     CR761829.1  GI:51967738
KEYWORDS    .
SOURCE      Xenopus tropicalis (Silurana tropicalis)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1236)
  AUTHORS   Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
            Francis,M.D., Garrett,N., Gilchrist,M.J., Grafham,D.V.,
            McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
            Voigt,J. and Zorn,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
COMMENT     Sanger Xenopus tropicalis EST/cDNA project.
            This sequence is from a Xenopus Gene Collection (XGC) library, from
            a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
            extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
            ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'
            end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1-blue.
FEATURES             Location/Qualifiers
     source          1. .1236
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TGas102g02"
                     /clone_lib="XGC-gastrula"
                     /dev_stage="gastrula (stage 10.5-13 mixed)"

ORIGIN

Alignment Scores:
Pred. No.:                6.87e-15        Length:          1236
Score:                    201.00          Matches:         44
Percent Similarity:       61.8%           Conservative:    19
Best Local Similarity:    43.1%           Mismatches:      21
Query Match:              44.9%           Indels:          18
DB:                       5               Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x CR761829 (1-1236)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||:::|||||:::|||||||||   ||||||   :::      :::
Db        205 ATGGAAAGGGCGATTAATTCTTTCTTTGAGCCTGGGGTGGAGTCAGCCTTGCAAAACAAA 264

Qy         21 ArgProGluThrIleSerGluPro--------------------------------Lys 29
                  |||::::::|||
Db        265 GCTGCAGCTGATATAGCAGATCCTTTAAAGCAAGAGACTATGTCTGGAACTGCTTCCGAT 324

Qy         30 ThrTyrValAspLeuThrAsnGluGlu------------ThrThrAspSerThrThrSer 45
              :::   :::||||||||||   ::::::               :::::::::||||||||||
Db        325 TCATGTATTGACCTCACCGGCGATGACTTAGTGGTTACCAAATCAGAAGCAACTACCAGC 384

Qy         46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 65
```

```
                    |||         :::||||||:::   |||  ||||||   :::|||
Db        385 AAT---------TCCTCCACTGTGAAGCAAGAAGATGAGAGCCATTTTTCCTTTCTGACT 435

Qy         66 TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 85
              ||||||||||||||||||   :::|||::::::||||||||||||||||||||||||||
Db        436 TGGAATATAGATGGCCTTGATGAATCAAATGTCGCAGAAAGGGCTCGTGGTGTGTGTTCC 495

Qy         86 TyrLeu 87
              ||||||
Db        496 TATTTG 501
```

RESULT 29
BC083404
LOCUS BC083404 1451 bp mRNA linear VRT 27-MAY-2005
DEFINITION Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7238334), partial cds.
ACCESSION BC083404
VERSION BC083404.1 GI:53734140
KEYWORDS .
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 1451)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
CONSRTM Mammalian Gene Collection Program Team
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 1451)
AUTHORS .
CONSRTM NIH MGC Project
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: John Postlethwait, University of Oregon
cDNA Library Preparation: Dr. Sumio Sugano
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 192 Row: c Column: 18.
FEATURES Location/Qualifiers
source 1. .1451
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7238334"
/tissue_type="Ovary"
/clone_lib="NIH_ZGC_5"
/lab_host="DH10B"
/note="Vector: pME18S-FL3"
gene <1. .1451
/gene="TTRAP"
/db_xref="GeneID:553516"
CDS <1. .1141
/gene="TTRAP"
/codon_start=2
/product="TTRAP protein"
/protein_id="AAH83404.1"
/db_xref="GI:53734141"
/db_xref="GeneID:553516"
/translation="FLSGNQTDQQMSALEESRTRLCDQFAFVSGSDSAVAQCYLAENE
WDMERALNSFFEAHMDSVFDEEAEKTEVTGNKRKDDTAEASGTKKKLKTDNADCIDLT
AEEPTCSITVNSKENQAENGTAKSEVEDSKLSIISWNVDGLDTLNLADRARGLCSYLA
LYTPDVVFLQELIPAYVQYLKKRAVSYLFFEGSDDGYFTGIMLRKSRVKFLESEIICF
PTTQMMRNLLIAQVTFSGQKLYLMTSHLESCKNQSQERTKQLRVVLQKIKEAPEDAIV
IFAGDTNLRDAEVANVGGLPAGVCDVWEQLGKQEHCRYTWDTKANSNKTVPYVSRCRF
DRIFLRSAKTGPPVTPDHMALIGMEKLDCGRYTSDHWGIYCTFNT"
ORIGIN

Alignment Scores:

Pred. No.:	2.34e-11	Length:	1451
Score:	173.00	Matches:	45
Percent Similarity:	50.9%	Conservative:	14
Best Local Similarity:	38.8%	Mismatches:	25
Query Match:	38.6%	Indels:	32
DB:	5	Gaps:	4

US-10-757-745-2_COPY_54_140 (1-87) x BC083404 (1-1451)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGlu---------------------------- 10
              |||||||||||||||||||||||:::|||||||
Db        140 ATGGAGAGAGCTTTAAACTCTTTCTTTGAGGCTCATATGGATTCAGTCTTTGATGAGGAA 199

Qy         11 ---ProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
                    |||   :::     :::|||:::   :::|||
Db        200 GCAGAGAAAACTGAAGTGACTGGGAATAAGAGAAAAGATGACACTGCTGAAGCTTCAGGG 259

Qy         30 Thr----------------------------TyrValAspLeuThrAsnGluGluThr 39
              |||                          :::|||||||||   ||||||
Db        260 ACGAAGAAGAAATTGAAGACGGACAATGCGGATTGCATTGACTTGACAGCAGAGGAGCCC 319

Qy         40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly--- 58
              |||   |||   |||               |||::::::   |||   |||||||||
Db        320 ACCTGCTCTATTACTGTAAAC---------TCTAAAGAAAATCAGGCTGAAAATGGCACA 370

Qy         59 --------------------SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
                                  |||      |||:::|||:::||||||:::||||||||||
Db        371 GCCAAATCTGAAGTGGAGGACAGCAAGCTCTCCATCATCAGCTGGAATGTGGACGGTTTG 430

Qy         72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||     ||||||::::::||||||||||||:::||||||||||||
Db        431 GACACATTAAATCTTGCAGATCGTGCCAGGGGCTTGTGCTCATATCTA 478
```

RESULT 30
AL589699
LOCUS AL589699 256608 bp DNA linear ROD 17-DEC-2004
DEFINITION Mouse DNA sequence from clone RP23-92G13 on chromosome 13 Contains the 3' end of the gene for a novel protein similar to KIAA0386, five novel genes, the Gmnn gene for geminin, gene PNAS-27, the Ttrap gene for Traf and Tnf receptor associated protein, the gene for a novel protein similar to KIAA0319, the gene for the ortholog of human and rat aldehyde dehydrogenase 5A1 (succinate-semialdehyde dehydrogenase) (ALDH5A1) (SSADH), the 5' end of the Gpld1 gene for glycosylphosphatidylinositol specific phospholipase D1 and four CpG islands, complete sequence.
ACCESSION AL589699

```
VERSION     AL589699.4  GI:13810107
KEYWORDS    HTG; Aldh5a1; CpG island; geminin; Gmnn; Gpld1; KIAA0319; KIAA0386;
            phospholipase; SSADH; Ttrap.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 256608)
  AUTHORS   Pelan,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 26, 2001 this sequence version replaced gi:13751568.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep ----------
            Genome Center
            Center: UK Medical Research Council
            Center code: UK-MRC
            Web site: http://mrcseq.har.mrc.ac.uk
            Contact: mouseq@har.mrc.ac.uk
            ----------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence.  Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
            RP23-92G13 is from the RPCI-23 Mouse BAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .256608
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="13"
                     /clone="RP23-92G13"
                     /clone_lib="RPCI-23"
     gene            join(AL513014.9:11808. .79376,4770. .28457)
                     /locus_tag="RP23-461O24.1-003"
     mRNA            join(AL513014.9:11808. .12037,AL513014.9:43061. .43187,
                     AL513014.9:44856. .45011,AL513014.9:46133. .46211,
                     AL513014.9:47748. .47771,AL513014.9:48876. .48929,
                     AL513014.9:50825. .50974,AL513014.9:53729. .53792,
                     AL513014.9:62567. .62610,AL513014.9:66867. .66992,
                     AL513014.9:67218. .67366,AL513014.9:68674. .68803,
                     AL513014.9:74163. .74831,AL513014.9:79192. .79376,
                     4770. .4932,8263. .8398,9931. .10092,12188. .12346,
                     16272. .16474,18350. .18433,19587. .19677,26226. .28457)
                     /locus_tag="RP23-461O24.1-003"
                     /note="match: cDNAs: AK007145.1 AB002384.1
                     match: ESTs: AW822704.1 BB674120.1 AA041998.1 BM144846.1
                     Z42557.1 R20596.1"
     CDS             join(AL513014.9:11989. .12037,AL513014.9:43061. .43187,
                     AL513014.9:44856. .45011,AL513014.9:46133. .46211,
                     AL513014.9:47748. .47771,AL513014.9:48876. .48929,
                     AL513014.9:50825. .50974,AL513014.9:53729. .53792,
                     AL513014.9:62567. .62610,AL513014.9:66867. .66992,
                     AL513014.9:67218. .67366,AL513014.9:68674. .68803,
                     AL513014.9:74163. .74831,AL513014.9:79192. .79376,
                     4770. .4932,8263. .8398,9931. .10092,12188. .12346,
                     16272. .16474,18350. .18433,19587. .19677,26226. .26326)
                     /locus_tag="RP23-461O24.1-003"
                     /standard_name="OTTMUSP00000000566"
                     /note="match: proteins: Q9D9C2 Q9Y4F9"
                     /codon_start=1
                     /protein_id="CAI26080.1"
                     /db_xref="GI:56237839"
                     /translation="MPPGTKRLRALGAFSAGLPTRLPEIMLVGSQSFSPGGPNGIIRS
                     QSFAGFSGLQERRSRCNSFIENASALKKPQAKLKKMHNLGHKNNNTPKEPQPKRVEEV
                     YRALKNGLDEYLEFHQTELDKLTAQLKDMKRNSRLGVLYDLDKQIKTIERYMRRLEFH
                     ISKVDELYEAYCIQRRLQDGASKMKQAFATSPASKAARESLSEINRSYKEYTENMCAI
                     EAELESLLGEFSIKMKGLAGFARLCPGDQYEIFMKYGRQRWKLKGKIEVNGKQSWDGA
                     ETVFLPLIVGFISIKVTELKGLATHILVGSVTCETKELFAARPQVVAVDINDLGTIKL
                     NLEITWYPFDVEDTTPSSGPGNKTAALQRRMSMYSQGTPETPTFKDQSFFSNLPDDIF
                     ESGKAAEEKRPLSLSFSDLQDGDCVFTSSSATSPSSSHSAHPEITITPAELTHSSLSS
                     QNEGTEDSSSASSRNSLGEDHEPKSHPKSDTVEPGKPGVATRSGTESLFLESSVAEAL
                     LQESDEASELKPVELDTFEGNITKQLVKRLTSAEGPVTTDKLFFEGSVGSESEAGRSF
                     LDGSLEDAFNGLFLALDPHKKQYKEFQDLNQEVTHLDDVLKCKPAGSRSRSSSLSLTV
                     ESALESFDFLNTSDFDEEEEDGDDVCHVGGGADSVFSDTETEKSGSVHPEARGHLSEA
                     LTEDTGVGTSVAGSPLPLTTGNESLDITIVKHLQYCTQLIQQIVFSSKTPFVARSLLE
                     KLSRQVLVMQKLAAVSDENLGNITSVVEAIPEFHKKLSLLAFWTKCCSPSGVYHSSAA
                     RLIKQLEASFARSINKDYPGLAEPVFRTLVSQILDRAEPLLSSSLSSEVITVFQYYSF
                     FTSHGVSDLETYLGQLTRQVAMVQTLQSLRDEKLLQTMSDLAPSNLPAQQEVLRTLAL
                     LLTRDDNEVSEAVTLYLAAASKNEHFREKALLYYCEALTKANLQLQKAACLALKSLEA
                     TESIKMLVTLCQSDTEEIRTVASETLLSLGEDGRLAYEQLDKFPRDCVKVGGRHGTEV
                     ATAF"
     polyA_signal    26424. .26429
                     /locus_tag="RP23-461O24.1-003"
     polyA_site      26446
                     /locus_tag="RP23-461O24.1-003"
     polyA_site      26447
                     /locus_tag="RP23-461O24.1-003"
     polyA_site      26457
                     /locus_tag="RP23-461O24.1-003"
     gene            join(35873. .36201,40700. .41233)
                     /locus_tag="RP23-92G13.3-001"
     mRNA            join(35873. .36201,40700. .41233)
                     /locus_tag="RP23-92G13.3-001"
                     /note="match: cDNAs: AK006029.1
                     match: ESTs: AV045211 BF018417 BB012811 BB014319 BB020000
                     AV263251 BF319523 AV206925 AV041162 AV259155 AV041846"
     polyA_signal    39211. .39216
                     /locus_tag="RP23-92G13.3-001"
     polyA_site      39233
                     /locus_tag="RP23-92G13.3-001"
     polyA_site      complement(44486)
     polyA_site      complement(44487)
     polyA_site      complement(44488)
     polyA_signal    complement(44511. .44516)
     polyA_signal    complement(44515. .44520)
     gene            complement(join(46486. .46871,47989. .48099,48299. .48381,
                     51241. .51376,52237. .52314,54801. .54858,56454. .56564))
                     /gene="Gmnn"
                     /locus_tag="RP23-92G13.2-001"
     mRNA            complement(join(46486. .46871,47989. .48099,48299. .48381,
                     51241. .51376,52237. .52314,54801. .54858,56454. .56564))
                     /gene="Gmnn"
                     /locus_tag="RP23-92G13.2-001"
                     /note="match: ESTs: BI789536.1 BQ178448.1 BF722692.1
                     AA510015.1 BG797776.1 BB868504.1 AA014584.1 AA032471.1
                     BI454155.1 BE448331.1 W13399.1 BF556131.1 BG085437.1
                     AI528457.1 AA277821.1 AA681520.1 W99871.1 AA004051.1"
     gene            complement(join(46486. .46871,47989. .48099,48299. .48381,
                     51241. .51376,52237. .52314,54801. .54858,56228. .56502))
                     /gene="Gmnn"
                     /locus_tag="RP23-92G13.2-002"
     mRNA            complement(join(46486. .46871,47989. .48099,48299. .48381,
                     51241. .51376,52237. .52314,54801. .54858,56228. .56502))
                     /gene="Gmnn"
                     /locus_tag="RP23-92G13.2-002"
                     /note="match: cDNAs: AF068780.1 BC020061.1
                     match: ESTs: BF160648.1 AA437903.1"
     CDS             complement(join(46710. .46871,47989. .48099,48299. .48381,
                     51241. .51376,52237. .52314,54801. .54851))
```

```
                     /gene="Gmnn"
                     /locus_tag="RP23-92G13.2-001"
                     /standard_name="OTTMUSP00000000569"
                     /note="match: proteins: O88513 O75496"
                     /codon_start=1
                     /protein_id="CAI26081.1"
                     /db_xref="GI:56237840"
                     /translation="MNLSMKQKQEGAQENVKNSPVPRRTLKMIQPSADGSLVGRENEL
                     PKGLFKRKLWDDQLASQTSSCGPEANENKDVGDLTQEAFDLISKENPSSQYWKEVAEQ
                     RRKALYEALKENEKLHKEIEQKDSEIARLRKENKDLAEVAEHVQYMAEVIERLSNEPL
                     DNFESPDSQEFDSEEEAVEYSELEDSGAGTCAEETVSSSTDARPCT"
     gene            96664. .108654
                     /gene="RP23-92G13.4"
                     /locus_tag="RP23-92G13.4-001"
     mRNA            join(96664. .98150,99423. .99599,100503. .100625,
                     104527. .104661,107278. .108654)
                     /gene="RP23-92G13.4"
                     /locus_tag="RP23-92G13.4-001"
                     /note="match: cDNAs: BC002217.1 BC005537.1 AL136632.1
                     AK022681.1 AJ420475.1
                     match: ESTs: BI078093.1 BG295445.1 BI696541.1 BG717389.1
                     BI108344.1 BG975946.1 BI452750.1 BG168325.1 BG295139.1
                     BG914583.1 BI753721.1"
     CDS             join(98022. .98150,99423. .99599,100503. .100625,
                     104527. .104661,107278. .107403)
                     /gene="RP23-92G13.4"
                     /locus_tag="RP23-92G13.4-001"
                     /standard_name="OTTMUSP00000000563"
                     /note="match: proteins: Q9H1Z2 Q99LU8 Q9GZU0 O95550"
                     /codon_start=1
                     /protein_id="CAI26082.1"
                     /db_xref="GI:56237841"
                     /translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
                     KKSALFEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
                     DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEFHHGDYEKQFLHV
                     LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
                     IEDHLRPYMPE"
     polyA_signal    106607. .106612
                     /gene="RP23-92G13.4"
                     /locus_tag="RP23-92G13.4-001"
     polyA_site      106654
                     /gene="RP23-92G13.4"
                     /locus_tag="RP23-92G13.4-001"

Alignment Scores:
Pred. No.:              2.74e-07        Length:         256608
Score:                  160.50          Matches:        36
Percent Similarity:     69.0%           Conservative:   4
Best Local Similarity:  62.1%           Mismatches:     17
Query Match:            35.8%           Indels:         1
DB:                     9               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x AL589699 (1-256608)

Qy          30 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 49
               |||   |||||||||||||||||:::      |||:::|||      :::   ||||||
Db      130652 ACCAGTGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCA 130711

Qy          50 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 69
               |||     |||      |||:::  |||      |||  ||||||||||||||||||
Db      130712 TCT---GGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGAT 130768

Qy          70 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
               |||||||||      ||||||  ||||||||||||||||||||||||   |||
Db      130769 GGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 130822

RESULT 31
AC125756/c
LOCUS       AC125756               245394 bp    DNA     linear   HTG 13-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-42H3, WORKING DRAFT SEQUENCE, 2
            unordered pieces.
ACCESSION   AC125756
VERSION     AC125756.5  GI:30578987
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 245394)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 245394)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 245394)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 13, 2003 this sequence version replaced gi:23265926.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
```

```
          (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
          in the feature table below represents a scaffold in the Atlas
          assembly (a 'contig-scaffold'). Within each contig-scaffold,
          individual sequence contigs are ordered and oriented, and separated
          by sized gaps filled with Ns to the estimated size. The sequence
          may extend beyond the ends of the clone and there may be sequence
          contigs within a contig-scaffold that consist entirely of whole
          genome shotgun sequence reads. Both end sequences and whole genome
          shotgun sequence only contigs will be indicated in the feature
          table.
          -------------- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
          -------------- Project Information
              Center project name: GJVU
              Center clone name: CH230-42H3
          -------------- Summary Statistics
              Assembly program: Atlas 3.0;
              Consensus quality: 238380 bases at least Q40
              Consensus quality: 239660 bases at least Q30
              Consensus quality: 240760 bases at least Q20
              Estimated insert size: 255581; sum-of-contigs estimation
              Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
          --------------
          * NOTE: Estimated insert size may differ from sequence length
          *    (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
          * NOTE: This sequence may represent more than one clone.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 2 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *         1   242009: contig of 242009 bp in length
          *    242010   242109: gap of unknown length
          *    242110   245394: contig of 3285 bp in length.
FEATURES             Location/Qualifiers
     source          1. .245394
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-42H3"
     misc_feature    1. .1025
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    1076. .2080
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    3723. .4536
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH348771"
     misc_feature    128656. .128715
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH260088"
     misc_feature    236171. .238412
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    239605. .242009
                     /note="wgs_end_extension
                     clone_end:Sp6"
     gap             242010. .242109
                     /estimated_length=unknown
ORIGIN

Alignment Scores:

Pred. No.:              6.11e-07        Length:         245394
Score:                  157.50          Matches:        35
Percent Similarity:     69.6%           Conservative:   4
Best Local Similarity:  62.5%           Mismatches:     16
Query Match:            35.2%           Indels:         1
DB:                     14              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x AC125756 (1-245394)

Qy          32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
               ||||||||||||||||||:::     |||:::|||      :::     ||||||||
Db      145740 GTTGATCTAACCAATGAGGATGCAAATGATACCACCATTTTAGAAACCAGTCCATCT--- 145684

Qy          52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
               |||       |||:::   |||      |||   ||||||||||||||||||||||||
Db      145683 GGAACTCCTCTAGAAGATAGCAGCACTATCTCTTTCATTACCTGGAATATTGATGGATTA 145624

Qy          72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
               |||       ||||||   ||||||||||||||||||||||||   |||
Db      145623 GATGGATGCAATCTCCCAGAGAGGGCTCGAGGGGTGTGTTCCTGCCTT 145576

RESULT 32
BC097117
LOCUS       BC097117               1455 bp    mRNA    linear   VRT 23-JUN-2005
DEFINITION  Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7431816), partial
            cds.
ACCESSION   BC097117
VERSION     BC097117.1  GI:66910444
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 1455)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1455)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Will Talbot
            cDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
```

```
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 226 Row: g Column: 18
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 53734140.
FEATURES              Location/Qualifiers
     source           1. .1455
                      /organism="Danio rerio"
                      /mol_type="mRNA"
                      /db_xref="taxon:7955"
                      /clone="IMAGE:7431816"
                      /tissue_type="Embryo, whole, 2-8 hr, mixed wild type, pool
                      of 2500 embryos"
                      /clone_lib="NIH_ZGC_15"
                      /lab_host="DH10B"
                      /note="Vector: pME18S-FL3"
     gene             <1. .1455
                      /gene="TTRAP"
                      /db_xref="GeneID:553516"
     CDS              <1. .438
                      /gene="TTRAP"
                      /codon_start=1
                      /product="TTRAP protein"
                      /protein_id="AAH97117.1"
                      /db_xref="GI:66910445"
                      /db_xref="GeneID:553516"
                      /translation="VFSGNQTDQQMSALGESRTRLCDQFAFVSGSDSAVAQCYLAENE
                      WDMERALNSFFEAHMDSVFDEEAEKTEVTGNKRKDDTAEASGTKKKLKTDNADCIDLT
                      AEEPTCSITVNSKKIRLKMAQPNLKWRTASSPSSAGMWTVWTH"
ORIGIN

Alignment Scores:
Pred. No.:              6.87e-09        Length:         1455
Score:                  153.00          Matches:        45
Percent Similarity:     50.9%           Conservative:   14
Best Local Similarity:  38.8%           Mismatches:     25
Query Match:            34.2%           Indels:         33
DB:                     5               Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x BC097117 (1-1455)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGlu---------------------------- 10
              |||||||||||||||||||:::||||||
Db        139 ATGGAGAGAGCTTTAAACTCTTTCTTTGAGGCTCATATGGATTCAGTCTTTGATGAGGAA 198

Qy         11 ---ProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
                  |||   :::      :::|||:::   :::|||
Db        199 GCAGAGAAAACTGAAGTGACTGGGAATAAGAGAAAAGATGACACTGCTGAAGCTTCAGGG 258

Qy         30 Thr----------------------------TyrValAspLeuThrAsnGluGluThr 39
              |||                          :::||||||||||   ||||||
Db        259 ACGAAGAAGAAATTGAAGACGGACAATGCGGATTGCATTGACTTGACAGCAGAGGAGCCC 318

Qy         40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly--- 58
              |||   |||   |||   |||:::        :::   |||   |||||||||
Db        319 ACCTGCTCTATTACTGT-AAACTC---------TAAGAAAATCAGGCTGAAAATGGCACA 368

Qy         59 --------------------SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
                             |||      |||:::|||:::||||||:::|||||||||
Db        369 GCCAAATCTGAAGTGGAGGACAGCAAGCTCTCCATCATCAGCTGGAATGTGGACGGTTTG 428

Qy         72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||      ||||||:::::|||||||||||||:::||||||||||||
Db        429 GACACATTAAATCTTGCAGATCGTGCCAGGGGCTTGTGCTCATATCTA 476

RESULT 33
BX511258
LOCUS       BX511258              200076 bp    DNA     linear   VRT 02-APR-2004
DEFINITION  Zebrafish DNA sequence from clone CH211-81E5 in linkage group 19,
            complete sequence.
ACCESSION   BX511258
VERSION     BX511258.6  GI:42820869
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 200076)
  AUTHORS   Heath,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 25, 2004 this sequence version replaced gi:42538803.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-81E5
            is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES              Location/Qualifiers
     source           1. .200076
                      /organism="Danio rerio"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:7955"
                      /clone="CH211-81E5"
                      /clone_lib="CHORI-211"
ORIGIN

Alignment Scores:
Pred. No.:              6.26e-06        Length:         200076
Score:                  148.50          Matches:        33
Percent Similarity:     64.1%           Conservative:   8
Best Local Similarity:  51.6%           Mismatches:     12
Query Match:            33.1%           Indels:         11
DB:                     5               Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x BX511258 (1-200076)

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
              :::|||||||||   ||||||   |||   |||   |||                |||:::
```

```
Db      84545 ATTGACTTGACAGCAGAGGAGCCCACCTGCTCTATTACTGTAAAC---------TCTAAA 84595

Qy         52 AspThrGlnGlnGluAsnGly-----------------------SerMetPheSerLeu 63
              :::   |||   |||||||||                       |||      |||:::
Db      84596 GAAAATCAGGCTGAAAATGGCACAGCCAAATCTGAAGTGGAGGACAGCAAGCTATCCATC 84655

Qy         64 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVal 83
              |||:::|||||||:::|||||||||||      ||||||:::|||||||||||||:::
Db      84656 ATCAGCTGGAATGTGGACGGTTTGGACACATTAAATCTTGCAGAACGTGCCAGGGGCTTG 84715

Qy         84 CysSerTyrLeu 87
              ||||||||||||
Db      84716 TGCTCATATCTA 84727

RESULT 34
BX957279
LOCUS       BX957279               198752 bp    DNA     linear   VRT 29-APR-2005
DEFINITION  Zebrafish DNA sequence from clone DKEY-218N20 in linkage group 19,
            complete sequence.
ACCESSION   BX957279
VERSION     BX957279.13  GI:62989971
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 198752)
  AUTHORS   Clark,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:61673710.
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-218N20
            is from a Zebrafish BAC library
            VECTOR: pIndigoBAC-5.
FEATURES             Location/Qualifiers
     source          1. .198752
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-218N20"
                     /clone_lib="DanioKey"
ORIGIN

Alignment Scores:
Pred. No.:              1.46e-05        Length:         198752
Score:                  145.50          Matches:        32
Percent Similarity:     64.1%           Conservative:   9
Best Local Similarity:  50.0%           Mismatches:     12
Query Match:            32.5%           Indels:         11
DB:                     5               Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x BX957279 (1-198752)

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
              :::|||||||||   ||||||   |||   |||   |||                |||:::
Db     143483 ATTGACTTGACAGCAGAGGAGCCCACCTGCTCTATTACTGTAAAC---------TCTAAA 143533

Qy         52 AspThrGlnGlnGluAsnGly-----------------------SerMetPheSerLeu 63
              :::   |||   |||||||||                       |||      |||:::
Db     143534 GAAAATCAGGCTGAAAATGGCACAGCCAAATCTGAAGTGGAGGACAGCAAGCTCTCCATC 143593

Qy         64 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVal 83
              |||:::|||||||:::|||||||||||      ||||||:::::|||||||||||:::
Db     143594 ATCAGCTGGAATGTGGACGGTTTGGACACATTAAATCTTGCAGATCGTGCCAGGGGCTTG 143653

Qy         84 CysSerTyrLeu 87
              ||||||||||||
Db     143654 TGCTCATATCTA 143665

RESULT 35
BD112459
LOCUS       BD112459                379 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  EST and encoded human protein.
ACCESSION   BD112459
VERSION     BD112459.1  GI:23207277
KEYWORDS    JP 2002010789-A/4536.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
  TITLE     EST and encoded human protein
  JOURNAL   Patent: JP 2002010789-A 4536 15-JAN-2002;
            GENSET CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002010789-A/4536
            PD   15-JAN-2002
            PF   07-AUG-2000 JP 2000280989
            PR   05-AUG-1999 US    60/147499
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE  PI
             GIORDANO
            PC   C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
             C12N1/21,
            PC   C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
            C12N15/00
            CC   EST and encoded human protein
            FH   Key             Location/Qualifiers
            FT   source          1. .379
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.47e-08        Length:         379
```

```
Score:                  145.00          Matches:        28
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  96.6%           Mismatches:     0
Query Match:            32.4%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BD112459 (1-379)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 78

Qy         22 ProGluThrIleSerGluProLysThr 30
              ||||||||||||||||||||||||||||
Db         79 CCTGAAACCATCTCTGAGCCCAAGACC 105

RESULT 36
AR416906
LOCUS       AR416906                379 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 8403 from patent US 6639063.
ACCESSION   AR416906
VERSION     AR416906.1  GI:40172016
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
  TITLE     EST's and encoded human proteins
  JOURNAL   Patent: US 6639063-A 8403 28-OCT-2003;
            Genset S.A.;;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              1.47e-08        Length:         379
Score:                  145.00          Matches:        28
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  96.6%           Mismatches:     0
Query Match:            32.4%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AR416906 (1-379)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 78

Qy         22 ProGluThrIleSerGluProLysThr 30
              ||||||||||||||||||||||||||||
Db         79 CCTGAAACCATCTCTGAGCCCAAGACC 105

RESULT 37
AX977600
LOCUS       AX977600                379 bp    DNA     linear   PAT 15-JAN-2004
DEFINITION  Sequence 8403 from Patent EP1104808.
ACCESSION   AX977600
VERSION     AX977600.1  GI:40983740
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
  TITLE     ESTs and encoded human proteins
  JOURNAL   Patent: EP 1104808-A 8403 06-JUN-2001;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.47e-08        Length:         379
Score:                  145.00          Matches:        28
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  96.6%           Mismatches:     0
Query Match:            32.4%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AX977600 (1-379)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 78

Qy         22 ProGluThrIleSerGluProLysThr 30
              ||||||||||||||||||||||||||||
Db         79 CCTGAAACCATCTCTGAGCCCAAGACC 105

RESULT 38
BD077628
LOCUS       BD077628                258 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  5'EST of secreted protein expressed in muscles and other mesodermal
            tissues.
ACCESSION   BD077628
VERSION     BD077628.1  GI:22623231
KEYWORDS    JP 2001512016-A/214.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 258)
  AUTHORS   Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
  TITLE     5'EST of secreted protein expressed in muscles and other mesodermal
            tissues
  JOURNAL   Patent: JP 2001512016-A 214 21-AUG-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001512016-A/214
            PD   21-AUG-2001
            PF   31-JUL-1998 JP 2000505295
            PR   01-AUG-1997 US    08/905134
            PI   JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO  PI
            LACROIX
            PC   C12N15/09,C12N15/09,C07K14/47,C12M1/00,C12N15/00,C12N15/00 CC
             blastn
            CC   identity 97
            CC   region 1. .180
            CC   id T31666
            CC   est
            CC   blastn
            CC   identity 92
            CC   region 88. .141
            CC   id R58665
            CC   est
            CC   blastn
            CC   identity 90
            CC   region 39. .93
            CC   id R58665
            CC   est
            CC   blastn
            CC   identity 98
            CC   region 63. .137
            CC   id R14990
            CC   est
            CC   blastn
```

```
          CC   identity 94
          CC   region 1. .50
          CC   id R14990
          CC   est
          CC   blastn
          CC   identity 99
          CC   region 1. .120
          CC   id T26956
          CC   est
          CC   Von Heijne matrix
          CC   score 6
          CC   seq LCXEFXSVASCDA/AV
          FH   Key             Location/Qualifiers
          FT   misc_feature    75. .254
          FT   misc_feature    73. .126
          FT   misc_feature    23. .77
          FT   misc_feature    157. .231
          FT   misc_feature    95. .144
          FT   misc_feature    135. .254
          FT   sig_peptide     31. .150.
FEATURES             Location/Qualifiers
     source          1. .258
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                8.42e-05        Length:         258
Score:                    113.00          Matches:        22
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              25.2%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BD077628 (1-258)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 249

Qy         21 ArgPro 22
              ||||||
Db        250 CGACCA 255

RESULT 39
AX939201
LOCUS       AX939201                  258 bp    DNA     linear   PAT 07-JAN-2004
DEFINITION  Sequence 220 from Patent EP1367124.
ACCESSION   AX939201
VERSION     AX939201.1  GI:40733556
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   dumas milne Edwards,J.B., Duclert,A. and Lacroix,B.
  TITLE     5' ests for secreted proteins expressed in muscle and other
            mesodermal tissues
  JOURNAL   Patent: EP 1367124-A 220 03-DEC-2003;
            GENSET (FR)
FEATURES             Location/Qualifiers
     source          1. .258
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /tissue_type="KIDNEY"
                     /dev_stage="FETAL"
     sig_peptide     31. .150
ORIGIN

Alignment Scores:
Pred. No.:                8.42e-05        Length:         258
Score:                    113.00          Matches:        22
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              25.2%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AX939201 (1-258)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 249

Qy         21 ArgPro 22
              ||||||
Db        250 CGACCA 255

RESULT 40
CQ544973
LOCUS       CQ544973                  60 bp    DNA     linear   PAT 30-JAN-2004
DEFINITION  Sequence 14608 from Patent WO0210449.
ACCESSION   CQ544973
VERSION     CQ544973.1  GI:41511237
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.
  TITLE     Oligonucleotide library for detecting rna transcripts and splice
            variants that populate a transcriptome
  JOURNAL   Patent: WO 0210449-A 14608 07-FEB-2002;
            Compugen Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                0.000493        Length:         60
Score:                    101.00          Matches:        19
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              22.5%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x CQ544973 (1-60)

Qy         48 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGG 58

RESULT 41
DQ073436
LOCUS       DQ073436                10013 bp    DNA     linear   BCT 12-JUL-2005
DEFINITION  Thermotoga sp. RQ2 RQ2_R12 region, complete sequence.
ACCESSION   DQ073436
VERSION     DQ073436.1  GI:69954024
KEYWORDS    .
SOURCE      Thermotoga sp. RQ2
  ORGANISM  Thermotoga sp. RQ2
            Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
REFERENCE   1  (bases 1 to 10013)
  AUTHORS   Mongodin,E.F., Hance,I.R., Deboy,R.T., Gill,S.R., Daugherty,S.,
            Huber,R., Fraser,C.M., Stetter,K. and Nelson,K.E.
  TITLE     Gene Transfer and Genome Plasticity in Thermotoga maritima, a Model
            Hyperthermophilic Species
  JOURNAL   J. Bacteriol. 187 (14), 4935-4944 (2005)
```

```
   PUBMED   15995209
REFERENCE   2  (bases 1 to 10013)
  AUTHORS   Mongodin,E.F., Hance,I.R., DeBoy,R.T., Gill,S.R., Daugherty,S.,
            Huber,R., Fraser,C.M., Stetter,K. and Nelson,K.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-2005) Microbial Genomics, The Institute for
            Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850,
            USA
FEATURES             Location/Qualifiers
     source          1. .10013
                     /organism="Thermotoga sp. RQ2"
                     /mol_type="genomic DNA"
                     /strain="RQ2"
                     /db_xref="taxon:126740"
     misc_feature    1. .10013
                     /note="RQ2_R12 region"
     CDS             <1. .732
                     /codon_start=1
                     /transl_table=11
                     /product="2-oxoacid ferredoxin oxidoreductase beta
                     subunit"
                     /protein_id="AAZ04346.1"
                     /db_xref="GI:69954025"
                     /translation="VVIVSGIGQAAKMPQYVGVNMFNGLHGRALPVATAIKTTNPNLL
                     VIAESGDGCMYGEGGNHFIHTIRRNPDIVNIVHDNQVYGLTKGQASPTSLKGFKTPVQ
                     VDGVILEPFNPLAVAIALDASFVARTFIGDIEFTKDILKEAMKHKGYALVDILQPCVT
                     FNKLNTYEWYRENTYYLKDHDPTDRELAFKRAIETEKLPLGIFYVNKNKETPEELARK
                     GDRTPLYEYEVNFEKLKELIESKKS"
     CDS             729. .1853
                     /codon_start=1
                     /transl_table=11
                     /product="putative oxygen-independent coproporphyrinogen
                     III oxidase"
                     /protein_id="AAZ04347.1"
                     /db_xref="GI:69954026"
                     /translation="MKLAVYVHIPFCKKKCXYCDFYSVASEDFEEYLSLVFEEIELYR
                     DVLKESEIETVYFGGGTPSLVPPSFLEKILEKLKEVSKKFSPLETTIEVNPESVDREK
                     LRSYRKMGINRISLGVQSCDDGVLKKCGRLYDEKMLKEKAELVLAEFDNVNFDFILGL
                     PGETDETLEKNLHFLEEFTPQHVSLYFLEVDENTPLYSLLKKNLLTLPDSDEVERRHD
                     VFVEFLKSRGFERYEISNFARPERESRHNLFYWRNHDYLGFGPSAGGHLGRFRYTNVS
                     DLSMYAKMVKEKTFPYDYRHENSEDEEALETLFMGLRTKEGVELKRVERLMFLVEKLQ
                     KEYACYVKVKDGKIFLSEEGMDLSKKIFEDLIEWYREAEK"
     CDS             1850. .2416
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAZ04348.1"
                     /db_xref="GI:69954027"
                     /translation="MTHIIHKGLDFFVKPRKISLNLNMKIGSAKVHPEDLKILMKKVP
                     VFMMSYYDSKAFMERELEISSADFPNGTIFFSYYEPVPAELNWDVDKNLLSQLARYFH
                     LYDLVSSMNSLIDESKGLSIGIYEEWLESTMVKVPGENAEELRNMLSKFSLMYTTKIL
                     WKMFHGNFEELKKRTHEIAYKFYEVAGF"
     CDS             2544. .5015
                     /codon_start=1
                     /transl_table=11
                     /product="alpha-glucan phosphorylase"
                     /protein_id="AAZ04349.1"
                     /db_xref="GI:69954028"
                     /translation="MSKKLPENLKHLEELAYNLWWSWXRPAQRLWRKIDPEGWEEHRN
                     PVKILKEVSDERLEELSKDDDFISLYELTLERFKDYMEKEDTWFNVNYPEWDEKIVYM
                     CMEYGLTKALPIYSGGLGILAGDHLKSASDLGLPLIAIGLLYKHGYFTQQIDRDGKQI
                     EIFPDYNPEDLPMKPLKDEKGNQVIVEVPLDSTVVKARVFEVKVGRVSLYLLDPDFEE
                     NEEKYRKICNYLYNPEPDVRISQEILLGIGGMKLLRALNLKPGVIHLNEGHPAFSSLE
                     RIKNYMEEGYSFTXALEIVRQTSVFTTHTPVPAGHDRFPFDLVEKKLSKFFEGFEKKN
                     LLMDLGKDETGSFNMTYLALRTSSFINGVSKLHAEVSRRMFKNVWQGVPVEEIPIEGI
                     TNGVHMGTWINREMRKLYDRYLGRVWRDHTDLEGIWYGVDRIPDEELWQAHLRAKKRF
                     IEYIKESVRRRNERLGIDEDVPNIDENSLIIGFARRFATYKRAVLLLSDLERLKKILN
                     DPERPVYVVYAGKAHPRDDAGKEFLKRIYEVSQMPEFKNRIIVLENYDIGMARLMVSG
                     VDVWLNNPRRPMEASGTSGMKAAANGVLNASVYDGWWVEGYNGRNGWVIGDESVLPET
                     EVDDPRDAEALYDLLENEIIPTYYENKEKWIFMMKESIKSVAPRFSTTRMLKEYTEKF
                     YIKGLVNKEWLERKENAERFGAWKERILRNWSSVSIERIVLEDTRSVEVTVKLGDLSP
                     DDVLVELLIGRGESMEDLEIWKVIQIRKHRREGDLFIYSYVNGALGHLGSPGWFYAVR
                     VLPYHPKLPTRFLPEIPVVWKKVLG"
     CDS             complement(5040. .5651)
                     /codon_start=1
                     /transl_table=11
                     /product="3-oxoacyl-acyl carrier protein reductase"
                     /protein_id="AAZ04350.1"
                     /db_xref="GI:69954029"
                     /translation="MVCSRNEELLKKTGHEYVVCDLKREPDRLFERIKEVDILVLNMG
                     GPKPGYLDDLTDEDFKDAIDGLFLNMVKIVRNYLPDMKRKRWGRIVAITSFSVVSPLE
                     NLYTSNSARMALTGFLKTLSFEVAPYNITVNCVAPGWTETERVKELLDEDKKRQVESQ
                     IPMRRMAKPEEIANVVAFLCSEKASYITGQTVVVDGGLSRFPL"
     CDS             complement(5731. .7179)
                     /codon_start=1
                     /transl_table=11
                     /product="putative response regulator"
                     /protein_id="AAZ04351.1"
                     /db_xref="GI:69954030"
                     /translation="MVDVLLRKNPDVRFEFFDNYSSMMEALKKREISVFLGDDVVARY
                     YLSKEGLLSDFLTLHLETNYLHVAVLKGNESILETVNSGLSLVSEDERNEIIRRYIPL
                     VVVTPPWLMKVVFYGSVLFSTALGISLAFVYLLRRKVEEKTLQLKLKNEELKAQNEEI
                     EALYQEVSANQEEIERLYSEIRELNDRFRHSIKELSKLVFIEDTSKFASELAEIIRDL
                     IEAKNVKVVLDGQVTGEARGETFEISHSKESHGYVVIEGDLGEDEREVLQAFLNIVSA
                     IHTYRELVRKEKKLHRDIVRTWVRALEYYDYYTKGHSEEVAYYAVEIGKMFNVDGEKL
                     EKLYWAGLLHDIGKIYVPQVVLNKTSKLDEREFELIKIHPVKGYELVKEIEGFEDVAI
                     WIRHHHERWDGKGYPDGLKGEEIPFEARVLCVADSYQAMRSDRPYKKGKTVEEAIQEL
                     LRNAGRQFDPVIVERFVEFLRKGGDLGTGHQG"
     CDS             complement(7600. .8148)
                     /codon_start=1
                     /transl_table=11
                     /product="crp family transcription regulator"
                     /protein_id="AAZ04352.1"
                     /db_xref="GI:69954031"
                     /translation="MVRYQGDPIDDILVLLEGSLRTEHVSENGKVLEIDTIRPVQIVA
                     SGLVFSRDPVYPVXVIAKEDSKILSIPKEKFLDLLMKDRNLLLFFLEDVSEHFRIVSE
                     KLFFLTXKTLKEKVIHYLVHHMNEDGELVLPVSVEELSRIFGCARPALSRVFQDLERE
                     GFIEKKGRRIRVLRNPLEDGKI"
     CDS             8294. .9589
                     /codon_start=1
                     /transl_table=11
                     /product="prismane protein"
                     /protein_id="AAZ04353.1"
                     /db_xref="GI:69954032"
                     /translation="MQMFCYQCSQTAKGVGCTEYGVCGKSPTVARLQDNLVFAIKGIS
                     AYYYHARELGYDBPEIAGFLDEALYSTLTNVNFDAQSFVEYALEAGRMNLKAMQLLKK
                     AHIETYGEPTPVEVETGTKKGKGIIVTGHNLKALEELLKQVEGTNVYVYTHSEMLPAH
                     GYPGLRKYKNLVGNLGKAWYDQRKLFAVYSVAILGTSNCVLIPSESYRDRMFTTSIAR
                     LPGVKHIDGYDYTEVIEKAKSLPDLEEKPGSYKLRTGFSTSVVVSLADKIKELVEAGK
                     IKHFLVVGGCDVPFKRNEYYREFVQKLPKETVVITLACGKFRINDLDLGDIEGIPRLI
                     DVGQCNDTIVAIEIAQALAKVFGVEVTELPLTLVLTWMEQKAVAILWTLLALGLKNIY
                     VGPVLPAWVNEDILKVLTAEFGLKTISEPEKDIKEILKV"
     CDS             complement(9629. .>10013)
                     /note="subfamily IA"
                     /codon_start=2
                     /transl_table=11
                     /product="HAD superfamily hydrolase"
                     /protein_id="AAZ04354.1"
                     /db_xref="GI:69954033"
                     /translation="FLPGAEEFLERLKKKDLRMAAVTNGVRFVQEKRSRKLKLDRFFE
                     FVLTSEEAGVEKPDPHIFWMALERMKLKKEEVLYVGDDLSSDLKGARNTGIDFVLFSP
                     DGDSSGDFPVARNFEELEKIVEELL"
ORIGIN

Alignment Scores:
Pred. No.:              3.54          Length:         10013
Score:                  90.00         Matches:        22
Percent Similarity:     56.2%         Conservative:   5
Best Local Similarity:  45.8%         Mismatches:     21
Query Match:            20.1%         Indels:         0
DB:                     1             Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x DQ073436 (1-10013)

Qy         18 LeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGlu 37
```

```
Db       2899 TTGGAATCCTCGCGGGAGACCATCTCAAATCCGCAAGCGATCTTGGACTTCCTCTCATAG 2958

Qy         38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
                 :::      ||||||          ||||||||||||       |||:::   ||||||
Db       2959 CGATCGGACTTCTCTACAAACATGGATATTTCACCCAGCAGATCGACAGAGATGGAAAAC 3018

Qy         58 GlySerMetPheSerLeuIleThr 65
                      |||||||||||||||
Db       3019 AGATAGAGATTTTCCCTGATTACA 3042

RESULT 42
AC018192/c
LOCUS       AC018192               64027 bp    DNA     linear   HTG 09-DEC-1999
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION   AC018192
VERSION     AC018192.1  GI:6552999
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 64027)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10213730 by the submitter.
            For more information on this record e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .64027
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
ORIGIN

Alignment Scores:
Pred. No.:                28.5          Length:         64027
Score:                    90.00         Matches:        21
Percent Similarity:       53.2%         Conservative:   21
Best Local Similarity:    26.6%         Mismatches:     35
Query Match:              20.1%         Indels:         2
DB:                       14            Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x AC018192 (1-64027)

Qy          3 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 22
              :::|||::::::          ||||||   ::::::||||||:::
Db      45264 CAAGCCATCCACCCCCCAAAATCGCCACCAAAACAGCAGTCGGCCATCCGTCCAACGGAG 45205

Qy         23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
                |||:::|||:::|||:::|||         :::          :::|||     :::
Db      45204 CACACAGTCAGTCAACCGCAGACAGGTCCCGAAAACGCCAAACACCAAACACCGCCCACG 45145

Qy         43 ThrThrSerLysIleSerProSerGluAspThrGln------GlnGluAsnGlySerMet 60
              ||||||:::          |||       :::|||          |||   ||||||||:::
Db      45144 ACCACGACCACTACACCACCAGAGATGAACACATCAGACTACCAAAGAAATGGGAGCTTA 45085

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                    ::::::       :::       ::::::|||       :::|||   |||
Db      45084 CATATGAAAGTATCTAATTATTTGCATCGCATTGAACTGCAAGAAGTGTCGGGCAGA 45028

RESULT 43
AC006243
LOCUS       AC006243               84549 bp    DNA     linear   HTG 30-JUL-1999
DEFINITION  Drosophila melanogaster strain y; cn bw sp chromosome 2 clone
            DS04807 (D462) map 60E5-60E8, *** SEQUENCING IN PROGRESS ***, 5
            unordered pieces.
ACCESSION   AC006243
VERSION     AC006243.5  GI:5630030
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 84549)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Sequencing of Drosophila melanogaster
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 84549)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R.,
            Zieran,L.L. and Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-DEC-1998) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Jul 28, 1999 this sequence version replaced gi:4510312.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www.fruitfly.org/sequence/) or send email
            to bdgp@fruitfly.berkeley.edu. All contigs in this submission meet
            the following cutoffs: length >= 200 bases. P1 library location:
            51-7.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1984: contig of 1984 bp in length
            *     1985     2064: gap of unknown length
            *     2065     4642: contig of 2578 bp in length
            *     4643     4722: gap of unknown length
            *     4723    10134: contig of 5412 bp in length
            *    10135    10214: gap of unknown length
            *    10215    16011: contig of 5797 bp in length
            *    16012    16091: gap of unknown length
            *    16092    84549: contig of 68458 bp in length.
FEATURES             Location/Qualifiers
     source          1. .84549
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="2"
                     /map="60E5-60E8"
                     /clone="DS04807 (D462)"
                     /clone_lib="P1 library, partial Sau3A in pNS582tet14Ad10"
     gap             1985. .2064
                     /estimated_length=unknown
     gap             4643. .4722
                     /estimated_length=unknown
     gap             10135. .10214
                     /estimated_length=unknown
     gap             16012. .16091
```

```
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:                39              Length:         84549
Score:                    90.00           Matches:        21
Percent Similarity:       53.2%           Conservative:   21
Best Local Similarity:    26.6%           Mismatches:     35
Query Match:              20.1%           Indels:         2
DB:                       14              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x AC006243 (1-84549)

Qy          3 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 22
              :::|||::::::            ||||||   ::::::||||||:::
Db      43259 CAAGCCATCCACCCCCCAAAATCGCCACCAAAACAGCAGTCGGCCATCCGTCCAACGGAG 43318

Qy         23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
              |||:::|||:::|||:::|||      :::             :::|||      :::
Db      43319 CACACAGTCAGTCAACCGCAGACAGGTCCCGAAAACGCCAAACACCAAACACCGCCCACG 43378

Qy         43 ThrThrSerLysIleSerProSerGluAspThrGln------GlnGluAsnGlySerMet 60
              ||||||:::         |||      :::|||         |||   ||||||||||:::
Db      43379 ACCACGACCACTACACCACCAGAGATGAACACATCAGACTACCAAAGAAATGGGAGCTTA 43438

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                 ::::::      :::      ::::::|||      :::|||   |||
Db      43439 CATATGAAAGTATCTAATTATTTGCATCGCATTGAACTGCAAGAAGTGTCGGGCAGA 43495

RESULT 44
AC007884
LOCUS       AC007884               190616 bp    DNA     linear   INV 28-FEB-2001
DEFINITION  Drosophila melanogaster, chromosome 2R, region 60F-60F, BAC clone
            BACR08I14, complete sequence.
ACCESSION   AC007884
VERSION     AC007884.5  GI:13162482
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 190616)
  AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
            Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
            Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
            Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
            Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
            Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
            Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
            McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
            Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
            Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
            Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
            Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
  TITLE     Sequencing of Drosophila chromosome 2R, region 60F-60F
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 190616)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Feb 28, 2001 this sequence version replaced gi:6587775.
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory, MS 64-121
            Berkeley, CA 94720
            This sequence was assembled using end sequences from a whole genome
            shotgun and from subclones of this BAC and its neighboring clones.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www.fruitfly.org/sequence/) or send email
            to bdgp@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .190616
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="2R"
                     /map="60F-60F"
                     /clone="BACR08I14 (D641)"
                     /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                     Drosophila melanogaster BAC library, partial EcoRI in
                     pBACe3.6)"
ORIGIN

Alignment Scores:
Pred. No.:                97.2            Length:         190616
Score:                    90.00           Matches:        21
Percent Similarity:       53.2%           Conservative:   21
Best Local Similarity:    26.6%           Mismatches:     35
Query Match:              20.1%           Indels:         2
DB:                       2               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x AC007884 (1-190616)

Qy          3 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 22
              :::|||::::::            ||||||   ::::::||||||:::
Db      24564 CAAGCCATCCACCCCCCAAAATCGCCACCAAAACAGCAGTCGGCCATCCGTCCAACGGAG 24623

Qy         23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
              |||:::|||:::|||:::|||      :::             :::|||      :::
Db      24624 CACACAGTCAGTCAACCGCAGACAGGTCCCGAAAACGCCAAACACCAAACACCGCCCACG 24683

Qy         43 ThrThrSerLysIleSerProSerGluAspThrGln------GlnGluAsnGlySerMet 60
              ||||||:::         |||      :::|||         |||   ||||||||||:::
Db      24684 ACCACGACCACTACACCACCAGAGATGAACACATCAGACTACCAAAGAAATGGGAGCTTA 24743

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                 ::::::      :::      ::::::|||      :::|||   |||
Db      24744 CATATGAAAGTATCTAATTATTTGCATCGCATTGAACTGCAAGAAGTGTCGGGCAGA 24800

RESULT 45
AE003465
LOCUS       AE003465               343833 bp    DNA     linear   INV 11-AUG-2005
DEFINITION  Drosophila melanogaster chromosome 2R, section 73 of 74 of the
            complete sequence.
ACCESSION   AE003465 AE002575 AE013599
VERSION     AE003465.3  GI:55380465
KEYWORDS    .
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 343833)
  AUTHORS   Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,G.L.,
            Abril,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D.,
            Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M.,
            Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S.,
            Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P.,
```

```
            Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A.,
            Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B.,
            Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I.,
            Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S.,
            Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C.,
            Ferriera,S., Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
            Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
            Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
            Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
            Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
            McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
            Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
            Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
            Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
            Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
            Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
            Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
            Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
            Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
            Weinstock,G.M., Weissenbach,J., Williams,S.M., WoodageT,
            Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
            Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
            Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
            Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
  TITLE     The genome sequence of Drosophila melanogaster
  JOURNAL   Science 287 (5461), 2185-2195 (2000)
  PUBMED    10731132
REFERENCE   2  (bases 1 to 343833)
  AUTHORS   Celniker,S.E., Wheeler,D.A., Kronmiller,B., Carlson,J.W.,
            Halpern,A., Patel,S., Adams,M., Champe,M., Dugan,S.P., Frise,E.,
            Hodgson,A., George,R.A., Hoskins,R.A., Laverty,T., Muzny,D.M.,
            Nelson,C.R., Pacleb,J.M., Park,S., Pfeiffer,B.D., Richards,S.,
            Sodergren,E.J., Svirskas,R., Tabor,P.E., Wan,K., Stapleton,M.,
            Sutton,G.G., Venter,C., Weinstock,G., Scherer,S.E., Myers,E.W.,
            Gibbs,R.A. and Rubin,G.M.
  TITLE     Finishing a whole-genome shotgun: release 3 of the Drosophila
            melanogaster euchromatic genome sequence
  JOURNAL   Genome Biol. 3 (12), RESEARCH0079 (2002)
  PUBMED    12537568
REFERENCE   3  (bases 1 to 343833)
  AUTHORS   Misra,S., Crosby,M.A., Mungall,C.J., Matthews,B.B., Campbell,K.S.,
            Hradecky,P., Huang,Y., Kaminker,J.S., Millburn,G.H., Prochnik,S.E.,
            Smith,C.D., Tupy,J.L., Whitfied,E.J., Bayraktaroglu,L.,
            Berman,B.P., Bettencourt,B.R., Celniker,S.E., de Grey,A.D.,
            Drysdale,R.A., Harris,N.L., Richter,J., Russo,S., Schroeder,A.J.,
            Shu,S.Q., Stapleton,M., Yamada,C., Ashburner,M., Gelbart,W.M.,
            Rubin,G.M. and Lewis,S.E.
  TITLE     Annotation of the Drosophila melanogaster euchromatic genome: a
            systematic review
  JOURNAL   Genome Biol. 3 (12), RESEARCH0083 (2002)
  PUBMED    12537572
REFERENCE   4  (bases 1 to 343833)
  AUTHORS   Kaminker,J.S., Bergman,C.M., Kronmiller,B., Carlson,J.,
            Svirskas,R., Patel,S., Frise,E., Wheeler,D.A., Lewis,S.E.,
            Rubin,G.M., Ashburner,M. and Celniker,S.E.
  TITLE     The transposable elements of the Drosophila melanogaster
            euchromatin: a genomics perspective
  JOURNAL   Genome Biol. 3 (12), RESEARCH0084 (2002)
  PUBMED    12537573
REFERENCE   5  (bases 1 to 343833)
  AUTHORS   Celniker,S., Carlson,J., Wan,K., Pfeiffer,B., Frise,E., George,R.,
            Hoskins,R., Stapleton,M., Pacleb,J., Park,S., Svirskas,R.,
            Smith,E., Yu,C. and Rubin,G.
  CONSRTM   Berkeley Drosophila Genome Project
  TITLE     Drosophila melanogaster release 4 sequence
  JOURNAL   Unpublished
REFERENCE   6  (bases 1 to 343833)
  AUTHORS   Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
REFERENCE   7  (bases 1 to 343833)
  AUTHORS   .
  CONSRTM   FlyBase
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUL-2005) FlyBase, Harvard University, Biological
            Laboratories, 16 Divinity Ave, Cambridge, MA 02138, USA
COMMENT     On Nov 4, 2004 this sequence version replaced gi:21645071.
FEATURES             Location/Qualifiers
     source          1. .343833
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /chromosome="2R"
     gene            complement(4011. .6144)
                     /gene="ST6Gal"
                     /locus_tag="Dmel_CG4871"
                     /old_locus_tag="CG4871"
                     /map="60D14-60D14"
                     /db_xref="FLYBASE:FBgn0035050"
     mRNA            complement(join(4011. .4308,4372. .4469,4577. .4856,
                     4919. .5016,5271. .5813,5883. .6144))
                     /gene="ST6Gal"
                     /locus_tag="Dmel_CG4871"
                     /old_locus_tag="CG4871"
                     /product="CG4871-RA, transcript variant A"
                     /db_xref="FLYBASE:FBgn0035050"
     mRNA            complement(join(4011. .4308,4372. .4469,4577. .4856,
                     4919. .5016,5271. .6144))
                     /gene="ST6Gal"
                     /locus_tag="Dmel_CG4871"
                     /old_locus_tag="CG4871"
                     /product="CG4871-RB, transcript variant B"
                     /db_xref="FLYBASE:FBgn0035050"
     CDS             complement(join(4138. .4308,4372. .4469,4577. .4856,
                     4919. .5016,5271. .5813,5883. .6048))
                     /gene="ST6Gal"
                     /locus_tag="Dmel_CG4871"
                     /old_locus_tag="CG4871"
                     /note="CG4871 gene product from transcript CG4871-RA"
                     /codon_start=1
                     /product="CG4871-PA, isoform A"
                     /protein_id="AAF47256.1"
                     /db_xref="GI:7291837"
                     /db_xref="FLYBASE:FBgn0035050"
                     /translation="MRQVTRASSGQSQALLSCLIIAVCAALIIQSGHIQGVKAQPGRE
                     PGEVNSRQPGNGNGTTGHRVPRSVFHVRWNPNEKFIVESRENPAINSSKLAPHPRLKV
                     SKNTKLTLSPKLYLCHDKHSELCHNKTQQFRQRIVRAFEKAMVESVNESQANHYNVDY
                     KPVFGDSFEEQYYPSTCLVMEAGVRVLRRKDAPFNKLPFGRLFPRQKLFRNVKDIKTC
                     AIVSSAGSLAGSKLGRFIDTHDIVMRFNHAPTQGHEVDVGSKTTIRVVNSQVVTKPEF
                     DFTRAPIFRNVTIAAWDPGKYNGTLEDWLTSADYDLFSNYELYRRRYPKSRAFLIDPH
                     SVWRLWQSLQMFAGNRPISKNPPSSGFIGLALLLPHCPQVDFVEYVPSTRLNGRCHYY
                     SKEMNSACTFGSWHPLAAEKLMALDMNMAEDDDMSVFQFGILRIRRPDKLLCGFNFFG
                     Y"
     CDS             complement(join(4138. .4308,4372. .4469,4577. .4856,
                     4919. .5016,5271. .6048))
                     /gene="ST6Gal"
                     /locus_tag="Dmel_CG4871"
                     /old_locus_tag="CG4871"
                     /note="CG4871 gene product from transcript CG4871-RB"
                     /codon_start=1
                     /product="CG4871-PB, isoform B"
                     /protein_id="AAM70791.1"
                     /db_xref="GI:21645072"
                     /db_xref="FLYBASE:FBgn0035050"
                     /translation="MRQVTRASSGQSQALLSCLIIAVCAALIIQSGHIQGVKAQPGRE
                     PGEVNSRQPGNGAYENSTGTERRLRKVIWKCIPTGNGTTGHRVPRSVFHVRWNPNEKF
                     IVESRENPAINSSKLAPHPRLKVSKNTKLTLSPKLYLCHDKHSELCHNKTQQFRQRIV
                     RAFEKAMVESVNESQANHYNVDYKPVFGDSFEEQYYPSTCLVMEAGVRVLRRKDAPFN
                     KLPFGRLFPRQKLFRNVKDIKTCAIVSSAGSLAGSKLGRFIDTHDIVMRFNHAPTQGH
                     EVDVGSKTTIRVVNSQVVTKPEFDFTRAPIFRNVTIAAWDPGKYNGTLEDWLTSADYD
                     LFSNYELYRRRYPKSRAFLIDPHSVWRLWQSLQMFAGNRPISKNPPSSGFIGLALLLP
                     HCPQVDFVEYVPSTRLNGRCHYYSKEMNSACTFGSWHPLAAEKLMALDMNMAEDDDMS
```

```
                VFQFGILRIRRPDKLLCGFNFFGY"
     gene       complement(<11940. .>14172)
                /gene="CG13595"
                /locus_tag="Dmel_CG13595"
                /old_locus_tag="CG13595"
                /map="60D14-60D14"
                /db_xref="FLYBASE:FBgn0035051"
     mRNA       complement(join(<11940. .12913,12975. .13132,13201. .13805,
                14134. .>14172))
                /gene="CG13595"
                /locus_tag="Dmel_CG13595"
                /old_locus_tag="CG13595"
                /product="CG13595-RA"
                /db_xref="FLYBASE:FBgn0035051"
     CDS        complement(join(11940. .12913,12975. .13132,13201. .13805,
                14134. .14172))
                /gene="CG13595"
                /locus_tag="Dmel_CG13595"
                /old_locus_tag="CG13595"
                /note="CG13595 gene product from transcript CG13595-RA"
                /codon_start=1
                /product="CG13595-PA"
                /protein_id="AAF47257.1"
                /db_xref="GI:7291838"
                /db_xref="FLYBASE:FBgn0035051"
                /translation="MPIDSSCDYELNEEAYRRWVCSTLVPYFAEPKDVSPQPKDEGQG
                KGKRSARSAASAEATAMADGIKHKSLKLISNNNKSNNGNNMNERQINQKHDKSSNANA
                NANSNASKDIDRSAEEAIDRDLVQTFYDVVGLSKPRHGEEEVLPENEAGLGRSADSEV
                DKGAEPDRKRSRNREAELELDTEAEAEQQSNNFISTANNSDPRATDPVISKLQKRSTR
                KTEFRKRRRIRPCLSVCQTVEQKCPYLLPADRAPALPTQYAGEPTFLCLDQNIPETGA
                QLEKSSYGPNDCCYSYCNGPASGICTVCQDFGQPGTEEREMAGNSTRRVHNITLSLSS
                HPGEHARAKSVALALRNVSQTGSDPDAVETLLDRLPYYASHDGIFYYDEEGEMPHASL
```

Alignment Scores:

Pred. No.:	189	Length:	343833
Score:	90.00	Matches:	21
Percent Similarity:	53.2%	Conservative:	21
Best Local Similarity:	26.6%	Mismatches:	35
Query Match:	20.1%	Indels:	2
DB:	2	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x AE003465 (1-343833)

```
Qy          3 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 22
              :::|||::::::            ||||||   ::::::||||||:::
Db     176868 CAAGCCATCCACCCCCCAAAATCGCCACCAAAACAGCAGTCGGCCATCCGTCCAACGGAG 176927

Qy         23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
              |||:::|||:::|||:::|||      :::           :::|||      :::
Db     176928 CACACAGTCAGTCAACCGCAGACAGGTCCCGAAAACGCCAAACACCAAACACCGCCCACG 176987

Qy         43 ThrThrSerLysIleSerProSerGluAspThrGln------GlnGluAsnGlySerMet 60
              ||||||:::         |||      :::|||         |||   |||||||||:::
Db     176988 ACCACGACCACTACACCACCAGAGATGAACACATCAGACTACCAAAGAAATGGGAGCTTA 177047

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                 ::::::      :::      ::::::|||      :::|||   |||
Db     177048 CATATGAAAGTATCTAATTATTTGCATCGCATTGAACTGCAAGAAGTGTCGGGCAGA 177104
```

RESULT 46
AC157293
LOCUS AC157293 229070 bp DNA linear HTG 01-JUL-2005
DEFINITION Bos taurus clone CH240-74L14, *** SEQUENCING IN PROGRESS ***, 28 unordered pieces.
ACCESSION AC157293
VERSION AC157293.2 GI:68301613
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 229070)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 229070)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 229070)
AUTHORS Cow Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jun 29, 2005 this sequence version replaced gi:59276098.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole

```
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
-------------- Genome Center
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
    Center project name: FECD
    Center clone name: CH240-74L14
-------------- Summary Statistics
    Assembly program: Atlas 3.0;
    Consensus quality: 219647 bases at least Q40
    Consensus quality: 221858 bases at least Q30
    Consensus quality: 223670 bases at least Q20
    Estimated insert size: 226343; sum-of-contigs estimation
    Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
--------------
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1    13691: contig of 13691 bp in length
*    13692    13741: gap of 50 bp
*    13742    41041: contig of 27300 bp in length
*    41042    41091: gap of 50 bp
*    41092    59205: contig of 18114 bp in length
*    59206    59255: gap of 50 bp
*    59256    61031: contig of 1776 bp in length
*    61032    61204: gap of 173 bp
*    61205    68366: contig of 7162 bp in length
*    68367    68416: gap of 50 bp
*    68417    81627: contig of 13211 bp in length
*    81628    82593: gap of 966 bp
*    82594    87131: contig of 4538 bp in length
*    87132    87181: gap of 50 bp
*    87182    97909: contig of 10728 bp in length
*    97910    97959: gap of 50 bp
*    97960   120079: contig of 22120 bp in length
*   120080   120129: gap of 50 bp
*   120130   129376: contig of 9247 bp in length
*   129377   129426: gap of 50 bp
*   129427   134468: contig of 5042 bp in length
*   134469   134518: gap of 50 bp
*   134519   140709: contig of 6191 bp in length
*   140710   140759: gap of 50 bp
*   140760   147068: contig of 6309 bp in length
*   147069   147118: gap of 50 bp
*   147119   165213: contig of 18095 bp in length
*   165214   165263: gap of 50 bp
*   165264   181056: contig of 15793 bp in length
*   181057   181106: gap of 50 bp
*   181107   199945: contig of 18839 bp in length
*   199946   200045: gap of unknown length
*   200046   201667: contig of 1622 bp in length
*   201668   201717: gap of 50 bp
*   201718   203261: contig of 1544 bp in length
*   203262   203361: gap of unknown length
*   203362   205057: contig of 1696 bp in length
*   205058   205157: gap of unknown length
*   205158   206783: contig of 1626 bp in length
*   206784   206883: gap of unknown length
*   206884   207972: contig of 1089 bp in length
*   207973   208072: gap of unknown length
*   208073   209085: contig of 1013 bp in length
*   209086   209185: gap of unknown length
*   209186   211162: contig of 1977 bp in length
*   211163   211262: gap of unknown length
*   211263   214299: contig of 3037 bp in length
*   214300   214399: gap of unknown length
*   214400   216680: contig of 2281 bp in length
*   216681   216780: gap of unknown length
*   216781   220038: contig of 3258 bp in length
*   220039   220138: gap of unknown length
*   220139   224937: contig of 4799 bp in length
*   224938   225037: gap of unknown length
*   225038   229070: contig of 4033 bp in length.
FEATURES             Location/Qualifiers
     source          1. .229070
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9913"
                     /clone="CH240-74L14"
     gap             13692. .13741
                     /estimated_length=50
     gap             41042. .41091
                     /estimated_length=50
     gap             59206. .59255
                     /estimated_length=50
     gap             61032. .61204
                     /estimated_length=173
     gap             68367. .68416
                     /estimated_length=50
     gap             81628. .82593
                     /estimated_length=966
     gap             87132. .87181
                     /estimated_length=50
     gap             97910. .97959
                     /estimated_length=50
     gap             120080. .120129
                     /estimated_length=50
     gap             129377. .129426
                     /estimated_length=50
     gap             134469. .134518
                     /estimated_length=50
     gap             140710. .140759
                     /estimated_length=50
     gap             147069. .147118

Alignment Scores:
Pred. No.:                211             Length:          229070
Score:                    88.00           Matches:         23
Percent Similarity:       47.4%           Conservative:    14
Best Local Similarity:    29.5%           Mismatches:      41
Query Match:              19.6%           Indels:          0
DB:                       14              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x AC157293 (1-229070)

Qy          4 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 23
              :::||||||          :::           |||:::   |||           :::
Db     210554 AGCTTGAACCAAATGAAAAAATAGAAGGCCTCAGAAAATAAGCTCAGTAATGAAATAAAA 210613

Qy         24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
               |||                 ::::::|||      |||:::||||||:::
Db     210614 TATATAGATAACTAAATGGAAATTTTAGAACTCAAAGAAGAAAAAACAACAAACAAACAA 210673

Qy         44 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 63
              |||           ||||||     ||||||:::||||||          :::|||
Db     210674 ACCCATAATAAAAAGCTCAGTGAATGGGATCAACAGCAGAATGGAAAGGAAGGAAATTTA 210733

Qy         64 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
              :::   |||||||||:::         |||:::|||   :::         |||
Db     210734 GTGAACTGGAACATAGAACAAAAGAAATTATCCAATCAGAATAACAATGAGAGA 210787

RESULT 47
AC157293/c
LOCUS       AC157293              229070 bp    DNA     linear   HTG 01-JUL-2005
DEFINITION  Bos taurus clone CH240-74L14, *** SEQUENCING IN PROGRESS ***, 28
```

unordered pieces.
ACCESSION AC157293
VERSION AC157293.2 GI:68301613
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 229070)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 229070)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 229070)
AUTHORS Cow Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Jun 29, 2005 this sequence version replaced gi:59276098.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
-------------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
Center project name: FECD
Center clone name: CH240-74L14
-------------- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 219647 bases at least Q40
Consensus quality: 221858 bases at least Q30
Consensus quality: 223670 bases at least Q20
Estimated insert size: 226343; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 13691: contig of 13691 bp in length
* 13692 13741: gap of 50 bp
* 13742 41041: contig of 27300 bp in length
* 41042 41091: gap of 50 bp
* 41092 59205: contig of 18114 bp in length
* 59206 59255: gap of 50 bp
* 59256 61031: contig of 1776 bp in length
* 61032 61204: gap of 173 bp
* 61205 68366: contig of 7162 bp in length
* 68367 68416: gap of 50 bp
* 68417 81627: contig of 13211 bp in length
* 81628 82593: gap of 966 bp
* 82594 87131: contig of 4538 bp in length
* 87132 87181: gap of 50 bp
* 87182 97909: contig of 10728 bp in length
* 97910 97959: gap of 50 bp
* 97960 120079: contig of 22120 bp in length
* 120080 120129: gap of 50 bp
* 120130 129376: contig of 9247 bp in length
* 129377 129426: gap of 50 bp
* 129427 134468: contig of 5042 bp in length
* 134469 134518: gap of 50 bp
* 134519 140709: contig of 6191 bp in length
* 140710 140759: gap of 50 bp
* 140760 147068: contig of 6309 bp in length
* 147069 147118: gap of 50 bp
* 147119 165213: contig of 18095 bp in length
* 165214 165263: gap of 50 bp
* 165264 181056: contig of 15793 bp in length
* 181057 181106: gap of 50 bp
* 181107 199945: contig of 18839 bp in length
* 199946 200045: gap of unknown length
* 200046 201667: contig of 1622 bp in length
* 201668 201717: gap of 50 bp

```
            *   201718   203261: contig of 1544 bp in length
            *   203262   203361: gap of unknown length
            *   203362   205057: contig of 1696 bp in length
            *   205058   205157: gap of unknown length
            *   205158   206783: contig of 1626 bp in length
            *   206784   206883: gap of unknown length
            *   206884   207972: contig of 1089 bp in length
            *   207973   208072: gap of unknown length
            *   208073   209085: contig of 1013 bp in length
            *   209086   209185: gap of unknown length
            *   209186   211162: contig of 1977 bp in length
            *   211163   211262: gap of unknown length
            *   211263   214299: contig of 3037 bp in length
            *   214300   214399: gap of unknown length
            *   214400   216680: contig of 2281 bp in length
            *   216681   216780: gap of unknown length
            *   216781   220038: contig of 3258 bp in length
            *   220039   220138: gap of unknown length
            *   220139   224937: contig of 4799 bp in length
            *   224938   225037: gap of unknown length
            *   225038   229070: contig of 4033 bp in length.
FEATURES             Location/Qualifiers
     source          1. .229070
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9913"
                     /clone="CH240-74L14"
     gap             13692. .13741
                     /estimated_length=50
     gap             41042. .41091
                     /estimated_length=50
     gap             59206. .59255
                     /estimated_length=50
     gap             61032. .61204
                     /estimated_length=173
     gap             68367. .68416
                     /estimated_length=50
     gap             81628. .82593
                     /estimated_length=966
     gap             87132. .87181
                     /estimated_length=50
     gap             97910. .97959
                     /estimated_length=50
     gap             120080. .120129
                     /estimated_length=50
     gap             129377. .129426
                     /estimated_length=50
     gap             134469. .134518
                     /estimated_length=50
     gap             140710. .140759
                     /estimated_length=50
     gap             147069. .147118

Alignment Scores:
Pred. No.:              494          Length:          229070
Score:                  85.00        Matches:         22
Percent Similarity:     47.4%        Conservative:    15
Best Local Similarity:  28.2%        Mismatches:      41
Query Match:            19.0%        Indels:          0
DB:                     14           Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x AC157293 (1-229070)

Qy           4 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 23
               :::||||||         :::             |||:::   |||          :::
Db      117000 AGCTTGAACCAAATGAAAAAATAGAAGGCCTCAGAAAATAAGCTCAGTAATGAAATAAAA 116941

Qy          24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
                 |||                 ::::::|||      |||:::||||||:::
Db      116940 TATATAGATAACTAAATGGAAATTTTAGAACTCAAAGAAGAAAAAACAACAAACAAACAA 116881

Qy          44 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 63
               |||                |||||      ||||||:::||||||         :::|||

Db      116880 ACCCATAATAAAAAGCTCAGTGAATGGGATCAACAGCAGAATGGAAAGGAAGGAAATTTA 116821

Qy          64 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
               :::   ||||||::::::          |||:::|||   :::         |||
Db      116820 GTGAACTGGAACATGGAACAAAAGAAATTATCCAATCAGAATAACAATGAGAGA 116767

RESULT 48
AC074173
LOCUS       AC074173               180129 bp    DNA     linear   HTG 15-JUL-2000
DEFINITION  Mus musculus strain C57BL6/J chromosome 12 clone RP23-104B17,
            WORKING DRAFT SEQUENCE, 12 unordered pieces.
ACCESSION   AC074173
VERSION     AC074173.1  GI:9211283
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 180129)
  AUTHORS   Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
            Bouffard,G.G., Dietrich,N.L., Eagle,W.O., Gupta,J., Ho,S.-L.,
            Huang,M.C., Idol,J., Lee-Lin,S.-Q., Maduro,Q.L., Maduro,V.B.,
            Mastrian,S.D., McCloskey,J.C., Ojodu,M.A., Pearson,R.,
            Stantripop,S., Summers,T.J., Thomas,J.W., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tran,J.T., Vogt,J.L., Walker,M.A.,
            Wetherby,K.D. and Green,E.D.
  TITLE     NISC Mouse Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 180129)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2000) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     -------------- Genome Center
                Center: NIH Intramural Sequencing Center
                Center code: NISC
                Web site: http://www.nisc.nih.gov
                Contact: nisc_mouse@nhgri.nih.gov
            -------------- Project Information
                Center project name: ur
                Center clone name: 104B17
            -------------- Summary Statistics
                Sequencing vector: plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.990319
                Consensus quality: 174572 bases at least Q40
                Consensus quality: 175790 bases at least Q30
                Consensus quality: 176302 bases at least Q20
                Insert size: 157000; agarose-fp
                Insert size: 212000; pulse-field-gel
                Insert size: 179029; sum-of-contigs
                Quality coverage: 7.99x in Q20 bases; agarose-fp
                Quality coverage: 5.92x in Q20 bases; pulse-field-gel
                Quality coverage: 7.01x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2373: contig of 2373 bp in length
            *     2374     2473: gap of unknown length
            *     2474     5454: contig of 2981 bp in length
            *     5455     5554: gap of unknown length
            *     5555     8773: contig of 3219 bp in length
            *     8774     8873: gap of unknown length
            *     8874    12930: contig of 4057 bp in length
            *    12931    13030: gap of unknown length
            *    13031    18158: contig of 5128 bp in length
```

```
            *    18159    18258: gap of unknown length
            *    18259    27856: contig of 9598 bp in length
            *    27857    27956: gap of unknown length
            *    27957    45444: contig of 17488 bp in length
            *    45445    45544: gap of unknown length
            *    45545    57407: contig of 11863 bp in length
            *    57408    57507: gap of unknown length
            *    57508    75041: contig of 17534 bp in length
            *    75042    75141: gap of unknown length
            *    75142   103108: contig of 27967 bp in length
            *   103109   103208: gap of unknown length
            *   103209   139636: contig of 36428 bp in length
            *   139637   139736: gap of unknown length
            *   139737   180129: contig of 40393 bp in length.
FEATURES             Location/Qualifiers
     source          1. .180129
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL6/J"
                     /db_xref="taxon:10090"
                     /chromosome="12"
                     /clone="RP23-104B17"
                     /clone_lib="RPCI mouse BAC library 23"
     misc_feature    1. .2373
                     /note="assembly_fragment"
     gap             2374. .2473
                     /estimated_length=unknown
     misc_feature    2474. .5454
                     /note="assembly_fragment"
     gap             5455. .5554
                     /estimated_length=unknown
     misc_feature    5555. .8773
                     /note="assembly_fragment"
     gap             8774. .8873
                     /estimated_length=unknown
     misc_feature    8874. .12930
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     gap             12931. .13030
                     /estimated_length=unknown
     misc_feature    13031. .18158
                     /note="assembly_fragment"
     gap             18159. .18258
                     /estimated_length=unknown
     misc_feature    18259. .27856
                     /note="assembly_fragment"
     gap             27857. .27956
                     /estimated_length=unknown
     misc_feature    27957. .45444
                     /note="assembly_fragment"
     gap             45445. .45544
                     /estimated_length=unknown
     misc_feature    45545. .57407
                     /note="assembly_fragment"
     gap             57408. .57507
                     /estimated_length=unknown
     misc_feature    57508. .75041
                     /note="assembly_fragment"
     gap             75042. .75141
                     /estimated_length=unknown
     misc_feature    75142. .103108
                     /note="assembly_fragment"
     gap             103109. .103208
                     /estimated_length=unknown
     misc_feature    103209. .139636
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     gap             139637. .139736
                     /estimated_length=unknown
     misc_feature    139737. .180129
                     /note="assembly_fragment"

ORIGIN

Alignment Scores:
Pred. No.:              501          Length:          180129
Score:                  84.00        Matches:         22
Percent Similarity:     51.7%        Conservative:    23
Best Local Similarity:  25.3%        Mismatches:      36
Query Match:            18.8%        Indels:          6
DB:                     14           Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x AC074173 (1-180129)

Qy          4 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 23
              :::::::      |||||||:::      :::|||      |||   :::|||
Db      93704 AGCATCACCAAGTATTTCCAGATGTCTCTTGAG------GCAGAGAAGCGGGTGAATGCC 93757

Qy         24 ThrIleSerGluProLysThrTyrValAsp---------LeuThrAsnGluGluThrThr 40
              :::   :::::::|||   :::   |||:::          ||||||   :::
Db      93758 TCCACCACAGACCCCAACAGCACTGTGGAGCAGTCTGCCCTCACGCGAGACAGAGTAGAA 93817

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||              :::   ||||||   ::::::   |||:::|||         :::
Db      93818 GATCTGATGTTGGAGCGAGAGTCTCCGTTCAAGGAGCAGCAGGAGGAACAGGCACGCCTC 93877

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                    :::          :::      |||||||||:::   :::::::::
Db      93878 CTGGACGAACTGGCCGGCAAACTGCAAAGTCTCGACCTGTCGGCTGTTGCACAGATGGTG 93937

Qy         81 Arg---GlyValCysSerTyr 86
              |||   ||||||||||||:::
Db      93938 AGGCTTGGCGTTTGCTCTCAT 93958

RESULT 49
AC074144
LOCUS       AC074144               184800 bp    DNA     linear   HTG 15-JUL-2000
DEFINITION  Mus musculus clone RP23-104B17, WORKING DRAFT SEQUENCE, 18
            unordered pieces.
ACCESSION   AC074144
VERSION     AC074144.1  GI:9211254
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 184800)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Mouse
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 184800)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 0
            Center clone name: RPCI-23_104B17
            -------
            Summary Statistics
            Consensus quality: 139801 bases at least Q40
            Consensus quality: 159725 bases at least Q30
            Consensus quality: 167763 bases at least Q20
            Estimated insert size: 150300; agarose-fp estimation
            Estimated insert size: 183100; sum-of-contigs estimation
            Quality coverage:  4.06 in Q20 bases; agarose-fp estimation
            Quality coverage:  3.34 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 18 contigs. The true order of the pieces
```

```
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1407: contig of 1407 bp in length
            *     1408     1507: gap of unknown length
            *     1508     2837: contig of 1330 bp in length
            *     2838     2937: gap of unknown length
            *     2938     4566: contig of 1629 bp in length
            *     4567     4666: gap of unknown length
            *     4667     7167: contig of 2501 bp in length
            *     7168     7267: gap of unknown length
            *     7268    10531: contig of 3264 bp in length
            *    10532    10631: gap of unknown length
            *    10632    12654: contig of 2023 bp in length
            *    12655    12754: gap of unknown length
            *    12755    14976: contig of 2222 bp in length
            *    14977    15076: gap of unknown length
            *    15077    18284: contig of 3208 bp in length
            *    18285    18384: gap of unknown length
            *    18385    21137: contig of 2753 bp in length
            *    21138    21237: gap of unknown length
            *    21238    30604: contig of 9367 bp in length
            *    30605    30704: gap of unknown length
            *    30705    42009: contig of 11305 bp in length
            *    42010    42109: gap of unknown length
            *    42110    60802: contig of 18693 bp in length
            *    60803    60902: gap of unknown length
            *    60903    74828: contig of 13926 bp in length
            *    74829    74928: gap of unknown length
            *    74929    91104: contig of 16176 bp in length
            *    91105    91204: gap of unknown length
            *    91205   109469: contig of 18265 bp in length
            *   109470   109569: gap of unknown length
            *   109570   129686: contig of 20117 bp in length
            *   129687   129786: gap of unknown length
            *   129787   158812: contig of 29026 bp in length
            *   158813   158912: gap of unknown length
            *   158913   184800: contig of 25888 bp in length.
FEATURES             Location/Qualifiers
     source          1. .184800
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-104B17"
                     /clone_lib="RPCI mouse BAC library 23"
     gap             1408. .1507
                     /estimated_length=unknown
     gap             2838. .2937
                     /estimated_length=unknown
     gap             4567. .4666
                     /estimated_length=unknown
     gap             7168. .7267
                     /estimated_length=unknown
     gap             10532. .10631
                     /estimated_length=unknown
     gap             12655. .12754
                     /estimated_length=unknown
     gap             14977. .15076
                     /estimated_length=unknown
     gap             18285. .18384
                     /estimated_length=unknown
     gap             21138. .21237
                     /estimated_length=unknown
     gap             30605. .30704
                     /estimated_length=unknown
     gap             42010. .42109
                     /estimated_length=unknown
     gap             60803. .60902
                     /estimated_length=unknown
     gap             74829. .74928
                     /estimated_length=unknown
     gap             91105. .91204
                     /estimated_length=unknown
     gap             109470. .109569
                     /estimated_length=unknown
     gap             129687. .129786
                     /estimated_length=unknown
     gap             158813. .158912
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:                516             Length:          184800
Score:                    84.00           Matches:         22
Percent Similarity:       51.7%           Conservative:    23
Best Local Similarity:    25.3%           Mismatches:      36
Query Match:              18.8%           Indels:          6
DB:                       14              Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x AC074144 (1-184800)

Qy           4 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 23
               ::::::          |||||||:::     :::|||     |||    :::|||
Db       92785 AGCATCACCAAGTATTTCCAGATGTCTCTTGAG------GCAGAGAAGCGGGTGAATGCC 92838

Qy          24 ThrIleSerGluProLysThrTyrValAsp---------LeuThrAsnGluGluThrThr 40
               :::   ::::::|||   :::   |||:::          ||||||   :::
Db       92839 TCCACCACAGACCCCAACAGCACTGTGGAGCAGTCTGCCCTCACGCGAGACAGAGTAGAA 92898

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               |||          :::   ||||||   ::::::   |||:::|||          :::
Db       92899 GATCTGATGTTGGAGCGAGAGTCTCCGTTCAAGGAGCAGCAGGAGGAACAGGCACGCCTC 92958

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                  :::       :::      |||||||||:::   :::::::::
Db       92959 CTGGACGAACTGGCCGGCAAACTGCAAAGTCTCGACCTGTCGGCTGTTGCACAGATGGTG 93018

Qy          81 Arg---GlyValCysSerTyr 86
               |||   ||||||||||||:::
Db       93019 AGGCTTGGCGTTTGCTCTCAT 93039

RESULT 50
CR974423
LOCUS       CR974423              240203 bp    DNA     linear   HTG 10-JUN-2005
DEFINITION  Mus musculus chromosome 12 clone RP23-182F9, *** SEQUENCING IN
            PROGRESS ***, 9 unordered pieces.
ACCESSION   CR974423
VERSION     CR974423.2  GI:67513836
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 240203)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 11, 2005 this sequence version replaced gi:67513434.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            -------------- Project Information
            Center project name: bM182F9
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 237080 bases at least Q40
            Consensus quality: 237450 bases at least Q30
```

```
            Consensus quality: 237872 bases at least Q20
            Insert size: 239403; sum-of-contigs
            Insert size: 231982; 1.2% error; agarose-fp
            Quality coverage: 12.01x in Q20 bases; sum-of-contigs Quality
            coverage: 12.51x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 9 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    73194: contig of 73194 bp in length
            *    73195    73294: gap of 100 bp
            *    73295    87270: contig of 13976 bp in length
            *    87271    87370: gap of 100 bp
            *    87371   126991: contig of 39621 bp in length
            *   126992   127091: gap of 100 bp
            *   127092   130212: contig of 3121 bp in length
            *   130213   130312: gap of 100 bp
            *   130313   150026: contig of 19714 bp in length
            *   150027   150126: gap of 100 bp
            *   150127   182302: contig of 32176 bp in length
            *   182303   182402: gap of 100 bp
            *   182403   215377: contig of 32975 bp in length
            *   215378   215477: gap of 100 bp
            *   215478   228684: contig of 13207 bp in length
            *   228685   228784: gap of 100 bp
            *   228785   240203: contig of 11419 bp in length.
FEATURES             Location/Qualifiers
     source          1. .240203
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="12"
                     /clone="RP23-182F9"
                     /clone_lib="RPCI-23"
     misc_feature    1. .73194
                     /note="assembly_fragment:02555
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:left"
     misc_feature    73295. .87270
                     /note="assembly_fragment:00393
                     fragment_chain:1"
     misc_feature    87371. .126991
                     /note="assembly_fragment:01937
                     fragment_chain:1"
     misc_feature    127092. .130212
                     /note="assembly_fragment:00012
                     fragment_chain:2"
     misc_feature    130313. .150026
                     /note="assembly_fragment:00616
                     fragment_chain:2"
     misc_feature    150127. .182302
                     /note="assembly_fragment:00940
                     fragment_chain:2"
     misc_feature    182403. .215377
                     /note="assembly_fragment:01429
                     fragment_chain:2"
     misc_feature    215478. .228684
                     /note="assembly_fragment:00028
                     fragment_chain:2"
     misc_feature    228785. .240203
                     /note="assembly_fragment:00196
                     fragment_chain:2
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:

Pred. No.:              693          Length:         240203
Score:                  84.00        Matches:        22
Percent Similarity:     51.7%        Conservative:   23
Best Local Similarity:  25.3%        Mismatches:     36
Query Match:            18.8%        Indels:         6
DB:                     14           Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x CR974423 (1-240203)

Qy          4 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 23
              ::::::      ||||||:::      :::|||      |||   :::|||
Db      36693 AGCATCACCAAGTATTTCCAGATGTCTCTTGAG------GCAGAGAAGCGGGTGAATGCC 36746

Qy         24 ThrIleSerGluProLysThrTyrValAsp---------LeuThrAsnGluGluThrThr 40
              :::   ::::::|||   :::   |||:::         ||||||   :::
Db      36747 TCCACCACAGACCCCAACAGCACTGTGGAGCAGTCTGCCCTCACGCGAGACAGAGTAGAA 36806

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||             :::   ||||||   ::::::   |||:::|||       :::
Db      36807 GATCTGATGTTGGAGCGAGAGTCTCCGTTCAAGGAGCAGCAGGAGGAACAGGCACGCCTC 36866

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                    :::         :::      ||||||||||:::   :::::::::
Db      36867 CTGGACGAACTGGCCGGCAAACTGCAAAGTCTCGACCTGTCGGCTGTTGCACAGATGGTG 36926

Qy         81 Arg---GlyValCysSerTyr 86
              |||   ||||||||||||:::
Db      36927 AGGCTTGGCGTTTGCTCTCAT 36947

RESULT 51
AC016751
LOCUS       AC016751               176379 bp    DNA     linear   PRI 21-APR-2005
DEFINITION  Homo sapiens BAC clone RP11-504O20 from 2, complete sequence.
ACCESSION   AC016751
VERSION     AC016751.7  GI:11415221
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 176379)
  AUTHORS   Maupin,R., Kruchowski,S. and Bolla,S.
  TITLE     The sequence of Homo sapiens BAC clone RP11-504O20
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 176379)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 176379)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-NOV-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 176379)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 176379)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6  (bases 1 to 176379)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-OCT-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
```

```
REFERENCE   7  (bases 1 to 176379)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Nov 29, 2000 this sequence version replaced gi:9799860.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0504O20
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-514D19; the clone sequenced
            to the right is RP11-538A12.  Actual start of this clone is at base
            position 1 of RP11-504O20; actual end is at base position 176379 of
            RP11-504O20.
FEATURES             Location/Qualifiers
     source          1. .176379
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-504O20"
                     /clone_lib="RPCI-11"
     misc_feature    85584. .85868
                     /note="CpG_island (%GC=62.5, o/e=0.86, #CpGs=24)"
ORIGIN

Alignment Scores:
Pred. No.:              749          Length:        176379
Score:                  82.50        Matches:       30
Percent Similarity:     47.7%        Conservative:  12
Best Local Similarity:  34.1%        Mismatches:    37
Query Match:            18.4%        Indels:        9
DB:                     8            Gaps:          4

US-10-757-745-2_COPY_54_140 (1-87) x AC016751 (1-176379)

Qy          6 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle 25
              ::::::::::|||  ||||||  |||  :::       |||  :::
Db      82026 TCGACATTCTTTTGCCCACCTTATGAGACCAATGAGTCTACTGAAGATCCTCCAAGTGCC 82085

Qy         26 SerGluProLysThrTyrValAspLeuThrAsn---------GluGluThrThrAspSer 42
                 |||  ||||||     |||     |||              ||||||
Db      82086 GAGGAAATAAAGACATTGGAGGATGCCCTTAATATTTTCAGAAATGTTACCACTGTAAAG 82145

Qy         43 ThrThrSerLysIleSerProSer-----------GluAspThrGlnGlnGluAsnGly 58
              :::|||||||||||||||  |||           |||     |||   :::
Db      82146 AGCACTTCCAAAATCTCCTTTTCACATATTCTCTGTGAAGTTATGCAACACAAAGGCTTC 82205

Qy         59 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 78
                 |||   ::::::     |||:::     ||||||  ||||||:::
Db      82206 AAAGCAATCTCCCAAGTTTCCCAGAAAATAAATACATCTGATTTGAAAAATCTTACATCG 82265

Qy         79 ArgAlaArgGlyValCysSerTyr 86
                |||||   :::|||  |||
Db      82266 ---GCAAGA---TTGTGCAAATAT 82283

RESULT 52
BD004082
LOCUS       BD004082                3321 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Apoptosis-related protein.
ACCESSION   BD004082
VERSION     BD004082.1  GI:18632043
KEYWORDS    JP 2001000193-A/1.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3321)
  AUTHORS   Yonehara,S. and Imai,Y.
  TITLE     Apoptosis-related protein
  JOURNAL   Patent: JP 2001000193-A 1 09-JAN-2001;
            SANKYO CO LTD
COMMENT     OS   Mus musculus (mouse)
            PN   JP 2001000193-A/1
            PD   09-JAN-2001
            PF   21-APR-2000 JP 2000120207
            PR
            PI   SHIN YONEHARA,YUZURU IMAI
            PC   C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
            C12N5/10,C12Q1/02,
            PC   G01N33/15,G01N33/50,G01N33/53//C12P21/02,C12P21/08,(C12N1/21,
            PC   C12R1:19),
            PC   (C12P21/02,C12R1:19),C12N15/00,C12N5/00
            CC
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(3321).
FEATURES             Location/Qualifiers
     source          1. .3321
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
ORIGIN

Alignment Scores:
Pred. No.:              11.4         Length:        3321
Score:                  81.50        Matches:       23
Percent Similarity:     46.7%        Conservative:  20
Best Local Similarity:  25.0%        Mismatches:    28
Query Match:            18.2%        Indels:        21
DB:                     6            Gaps:          4

US-10-757-745-2_COPY_54_140 (1-87) x BD004082 (1-3321)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSer--------------AlaLeuGlu 19
              |||:::     :::|||||||  |||  |||:::              :::::|||
Db       2320 TTGCATCAAGTATGGGAACCAAAAGTTTCTGAGAACCTTCAAGAGTTGCCTTCAATGGAG 2379

Qy         20 ArgArgProGluThrIleSerGlu------ProLysThrTyrValAspLeuThrAsnGlu 37
              :::  |||  ::::::           |||  |||::::::|||||||   :::
Db       2380 AAAATCCCACACTCTCTTGATAATCATCTTCCTGATACACACATAGATCTAACGAAAGAT 2439

Qy         38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
```

```
                  |||::::::                                  :::
Db       2440 TCAGCCACTGAGACC----------------------------------AAAAGCTTG 2463

Qy         58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn------Asn 75
              |||   :::      :::         |||:::|||   |||:::   :::      |||
Db       2464 GGGGAGCTAATGGAAGTAACAGTTTTAAATGTTGATCACTTGGAATGTTCTCAAACCAAC 2523

Qy         76 LeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||   :::   |||       ||||||   |||
Db       2524 TTAGATCAAGATGCAGAGATAACATGTAGTTCTTTA 2559

RESULT 53
BD004101
LOCUS       BD004101               6763 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Apoptosis-related protein.
ACCESSION   BD004101
VERSION     BD004101.1  GI:18632062
KEYWORDS    JP 2001000193-A/20.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 6763)
  AUTHORS   Yonehara,S. and Imai,Y.
  TITLE     Apoptosis-related protein
  JOURNAL   Patent: JP 2001000193-A 20 09-JAN-2001;
            SANKYO CO LTD
COMMENT     OS   Mus musculus (mouse)
            PN   JP 2001000193-A/20
            PD   09-JAN-2001
            PF   21-APR-2000 JP 2000120207
            PR
            PI   SHIN YONEHARA,YUZURU IMAI
            PC   C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
            C12N5/10,C12Q1/02,
            PC   G01N33/15,G01N33/50,G01N33/53//C12P21/02,C12P21/08,(C12N1/21,
            PC   C12R1:19),
            PC   (C12P21/02,C12R1:19),C12N15/00,C12N5/00
            CC
            FH   Key             Location/Qualifiers
            FT   CDS             (256). .(6141).
FEATURES             Location/Qualifiers
     source          1. .6763
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
ORIGIN

Alignment Scores:
Pred. No.:                25.4          Length:          6763
Score:                    81.50         Matches:         23
Percent Similarity:       46.7%         Conservative:    20
Best Local Similarity:    25.0%         Mismatches:      28
Query Match:              18.2%         Indels:          21
DB:                       6             Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x BD004101 (1-6763)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSer--------------AlaLeuGlu 19
              |||:::     :::|||||||   |||   |||:::              ::::::|||
Db       5140 TTGCATCAAGTATGGGAACCAAAAGTTTCTGAGAACCTTCAAGAGTTGCCTTCAATGGAG 5199

Qy         20 ArgArgProGluThrIleSerGlu------ProLysThrTyrValAspLeuThrAsnGlu 37
              :::   |||   ::::::          |||   |||::::::|||||||||   :::
Db       5200 AAAATCCCACACTCTCTTGATAATCATCTTCCTGATACACACATAGATCTAACGAAAGAT 5259

Qy         38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::                                  :::
Db       5260 TCAGCCACTGAGACC----------------------------------AAAAGCTTG 5283

Qy         58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn------Asn 75
              |||   :::      :::         |||:::|||   |||:::   :::      |||
```

```
Db       5284 GGGGAGCTAATGGAAGTAACAGTTTTAAATGTTGATCACTTGGAATGTTCTCAAACCAAC 5343

Qy         76 LeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||   :::   |||       ||||||   |||
Db       5344 TTAGATCAAGATGCAGAGATAACATGTAGTTCTTTA 5379

RESULT 54
AF132726
LOCUS       AF132726               6763 bp    mRNA    linear   ROD 06-MAY-1999
DEFINITION  Mus musculus FLASH (Flash) mRNA, complete cds.
ACCESSION   AF132726
VERSION     AF132726.1  GI:4754904
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 6763)
  AUTHORS   Imai,Y., Kimura,T., Murakami,A., Yajima,N., Sakamaki,K. and
            Yonehara,S.
  TITLE     The CED-4-homologous protein FLASH is involved in Fas-mediated
            activation of caspase-8 during apoptosis
  JOURNAL   Nature 398 (6730), 777-785 (1999)
   PUBMED   10235259
REFERENCE   2  (bases 1 to 6763)
  AUTHORS   Imai,Y., Kimura,T., Murakami,A., Yajima,N., Sakamaki,K. and
            Yonehara,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAR-1999) Institute for Virus Research, Kyoto
            University, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan
FEATURES             Location/Qualifiers
     source          1. .6763
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_line="WR19L"
                     /tissue_type="T lymphoma"
     gene            1. .6763
                     /gene="Flash"
     CDS             256. .6144
                     /gene="Flash"
                     /note="similar to CED-4"
                     /codon_start=1
                     /product="FLASH"
                     /protein_id="AAD29045.1"
                     /db_xref="GI:4754905"
                     /translation="MAADDDNGDGTGLFDVCPASPLKNNDEGSLDIYAGLDSAVSDST
                     ARSCVSFRNCLDLYEEILTEEGTAKEATYNDLQIEYGKCQQQMKDLMKRFKEIQTQNL
                     NLKNENQSLKKNISALIKTARVEINRKDEEINHLHQRLSEFPHFRNNHKAARTKDSQS
                     TSPHLDDCSKTDHGVKSDVQKDVHPNTAQPNLEKEGKSHSEAQNPLHLSTGVEKHCAN
                     NVWSRSPYQVGEGNSNEDNRRGRSGTRHSQCSRGTDRTQKDLHSSCNDSEPRDKEANS
                     RLQGHPEKHGNSEARTESKISESKSSTGMGYKSERSASSWEKETSRERPHTRVESQHD
                     KNLEKQNERLQNMHRKELPSQDKTERKVDVKFKPAGEEQGHRGRVDRALPPHPKNDVK
                     HYGFNKYHPEERRGREDCKRDRGMNSHGFQDRRCSSFLSSNRNSKYPHSKEVSVAHQW
                     ENTPFKAERHRTEDRRKRERENKEESRHVKSDKKSPPEHLQRTHKDTKKSTADGKRQT
                     EPKHGKGAVSNSELSKGTDSKEGATKVESGPNEAKGKDLKLSFMEKLNLTLSPAKKQP
                     ACQDNPHQITGVPEPSGTCDSRSLETTGTVACLPSGSEHNREETKSELPEPKEALLAT
                     SQLRISIPENKMKEEKRLLFKSVENTVPCELLACGTEISLPAPVEIEQARCLLGSVEV
                     EETCGGARTAASVVMHVLPEHASEDASQELDTKRHDGINACAISEGVKTKVILSPKAA
                     AASESHLAPLVEEPSISLVNCSGDNNPKLEPSLEERPIVETKSCPLESCLPKETFVPS
                     PQKTELIDHKIETGESNSVYQDDDNSVLSIDFNNLRPIPDPISPLNSPVRPVCKVVSM
                     ESSCAIPLYDSSHKDEFPSNSTLSTFKSQSDLNKENEKPVPKFDKCSEADSCKHLSLD
                     ELEEGEIRSDDEESVAQKRLEKSARPRVSAEVQPGKSSPGSRRSTVHVHKDNGRTAVK
                     LPRDRLTWSKRSSESRPSNTERKSKTMSISSLEKILPLILVPSSLWEVMHMLRLLGKH
                     VRKNYMKFKIKFSLTQFHRIIESAILSFTSLIKCLDLSKICKSVSTLQKSLCEVIESN
                     LKQVKKNGIVDRLFEQQQTDMKKKLWKFVDEQLDYLFEKLKKILLKFCDSVNFENENS
                     EGKLGKKYKERTQHSNCQKKKMDNKEIRREKVLKSENTVNFKSSLGCEKSEEKHQDQN
                     KTNASIVKHDVKRTFSTCSDNTKNAECKEQFLEKSCPSTPRPGKDEGHTEEEAQAAQH
                     ASAKSERSFEILTEQQASSLTFNLVSDAQMGEIFKSLLQGSDLLDTSGTEKAEWELKT
                     PEKQLLESLKCESAPACATEELVSEGASLCPKVISDDNWSLLSSEKGPSLSSGLSLPV
                     HPDVLDENCMFEVSSNTALGKDNVYSSEKSKPCISSILLEDLAVSLTVPSPLKSDGHL
                     SFLKPEVLSTSTPEEVISAHFSEDALLEEEDASEQDIHLALESDNSSSKSSCSSWTSR
```

SVASGFQYHPNLPMHAVIMEKSNDHFIVKIRRATPSTSPGLKHGVVAEESLTSLPRTG
KEAGVATEKEPNLFQSTVLKPVKDLENTDKNIDKSKLTHEEQNSIVQTQVPDIYEFLK
DASNKVVHCDQVVDDCFKLHQVWEPKVSENLQELPSMEKIPHSLDNHLPDTHIDLTKD
SATETKSLGELMEVTVLNVDHLECSQTNLDQDAEITCSSLQPDTIDAFIDLTHDASSE
SKNEGSEPVLAVEGMGCQVICIDEDTNKEGKMGRANSPLESIVEETCIDLTSESPGSC
EIKRHNLKSEPPSKLDCLELPETLGNGHKKRKNSPGVSHSSQKKQRKDIDLSSEKTQR
LSPNSDRNGDAHRKQASKKREPAVNETSLSSEASPEVKGSTAVLAASPASLSAKNVIK
KKGEIIVSWTRNDDREILLECQKRMPSLKTFTYLAVKLNKNPNQVSERFQQLKKLFEK
SKCR"

ORIGIN

Alignment Scores:

Pred. No.:	25.4	Length:	6763
Score:	81.50	Matches:	23
Percent Similarity:	46.7%	Conservative:	20
Best Local Similarity:	25.0%	Mismatches:	28
Query Match:	18.2%	Indels:	21
DB:	9	Gaps:	4

US-10-757-745-2_COPY_54_140 (1-87) x AF132726 (1-6763)

```
Qy          5 LeuAsnSerTyrPheGluProProValGluGluSer--------------AlaLeuGlu 19
              |||:::      :::||||||   |||   |||:::             ::::::|||
Db       5140 TTGCATCAAGTATGGGAACCAAAAGTTTCTGAGAACCTTCAAGAGTTGCCTTCAATGGAG 5199

Qy         20 ArgArgProGluThrIleSerGlu------ProLysThrTyrValAspLeuThrAsnGlu 37
              :::   |||   ::::::            |||  |||::::::|||||||||   :::
Db       5200 AAAATCCCACACTCTCTTGATAATCATCTTCCTGATACACACATAGATCTAACGAAAGAT 5259

Qy         38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
                    |||::::::                                    :::
Db       5260 TCAGCCACTGAGACC----------------------------------AAAAGCTTG 5283

Qy         58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn------Asn 75
              |||   :::     :::         |||:::|||   |||:::   :::        |||
Db       5284 GGGGAGCTAATGGAAGTAACAGTTTTAAATGTTGATCACTTGGAATGTTCTCAAACCAAC 5343

Qy         76 LeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||   :::   |||       ||||||   |||
Db       5344 TTAGATCAAGATGCAGAGATAACATGTAGTTCTTTA 5379
```

RESULT 55
AL805973

LOCUS	AL805973 138031 bp DNA linear ROD 25-JUL-2002
DEFINITION	Mouse DNA sequence from clone RP23-69G24 on chromosome 4, complete sequence.
ACCESSION	AL805973
VERSION	AL805973.6 GI:21998271
KEYWORDS	HTG.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 138031)
AUTHORS	Harrison,E.
TITLE	Direct Submission
JOURNAL	Submitted (25-JUL-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT	On Jul 26, 2002 this sequence version replaced gi:21955701. ---------- Genome Center Center: Wellcome Trust Sanger Institute Center code: SC Web site: http://www.sanger.ac.uk Contact: humquery@sanger.ac.uk ---------- During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-69G24 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBACe3.6.
FEATURES	Location/Qualifiers
source	1. .138031 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="4" /clone="RP23-69G24" /clone_lib="RPCI-23"

ORIGIN

Alignment Scores:

Pred. No.:	756	Length:	138031
Score:	81.50	Matches:	23
Percent Similarity:	46.7%	Conservative:	20
Best Local Similarity:	25.0%	Mismatches:	28
Query Match:	18.2%	Indels:	21
DB:	9	Gaps:	4

US-10-757-745-2_COPY_54_140 (1-87) x AL805973 (1-138031)

```
Qy          5 LeuAsnSerTyrPheGluProProValGluGluSer--------------AlaLeuGlu 19
              |||:::      :::||||||   |||   |||:::             ::::::|||
Db      29938 TTGCATCAAGTATGGGAACCAAAAGTTTCTGAGAACCTTCAAGAGTTGCCTTCAATGGAG 29997

Qy         20 ArgArgProGluThrIleSerGlu------ProLysThrTyrValAspLeuThrAsnGlu 37
              :::   |||   ::::::            |||  |||::::::|||||||||   :::
Db      29998 AAAATCCCACACTCTCTTGATAATCATCTTCCTGATACACACATAGATCTAACGAAAGAT 30057

Qy         38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
                    |||::::::                                    :::
Db      30058 TCAGCCACTGAGACC----------------------------------AAAAGCTTG 30081

Qy         58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn------Asn 75
              |||   :::     :::         |||:::|||   |||:::   :::        |||
Db      30082 GGGGAGCTAATGGAAGTAACAGTTTTAAATGTTGATCACTTGGAATGTTCTCAAACCAAC 30141

Qy         76 LeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||   :::   |||       ||||||   |||
Db      30142 TTAGATCAAGATGCAGAGATAACATGTAGTTCTTTA 30177
```

RESULT 56
AC109433/c

LOCUS	AC109433 233372 bp DNA linear HTG 08-OCT-2002
DEFINITION	Rattus norvegicus clone CH230-99D20, *** SEQUENCING IN PROGRESS ***, 6 unordered pieces.
ACCESSION	AC109433
VERSION	AC109433.4 GI:23270090
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 233372)
AUTHORS	Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 233372)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 233372)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Sep 23, 2002 this sequence version replaced gi:21738728.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GRGC
Center clone name: CH230-99D20
---------- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 186934 bases at least Q40
Consensus quality: 190979 bases at least Q30
Consensus quality: 193518 bases at least Q20
Estimated insert size: 213564; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 14442: contig of 14442 bp in length
* 14443 14542: gap of unknown length
* 14543 189203: contig of 174661 bp in length
* 189204 189303: gap of unknown length
* 189304 226103: contig of 36800 bp in length
* 226104 226203: gap of unknown length
* 226204 227571: contig of 1368 bp in length
* 227572 227671: gap of unknown length
* 227672 229227: contig of 1556 bp in length
* 229228 229327: gap of unknown length
* 229328 233372: contig of 4045 bp in length.

FEATURES Location/Qualifiers
source 1. .233372
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-99D20"
misc_feature 1. .1279
/note="wgs_contig"
gap 14443. .14542
/estimated_length=unknown
misc_feature 34647. .36034
/note="wgs_contig"
misc_feature 130566. .132005
/note="wgs_contig"
gap 189204. .189303
/estimated_length=unknown
gap 226104. .226203
/estimated_length=unknown
gap 227572. .227671
/estimated_length=unknown
gap 229228. .229327
/estimated_length=unknown
ORIGIN

Alignment Scores:

Pred. No.:	2.09e+03	Length:	233372
Score:	80.00	Matches:	28
Percent Similarity:	46.5%	Conservative:	12
Best Local Similarity:	32.6%	Mismatches:	32
Query Match:	17.9%	Indels:	14
DB:	14	Gaps:	3

US-10-757-745-2_COPY_54_140 (1-87) x AC109433 (1-233372)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20

```
Db     174882 ATGGAGAAACCAAGATATTCCATGACAGAACCAAATTTACACAATATCTTTCTACAAATC 174823

Qy         21 ArgProGluThr---------IleSerGluProLysThrTyrValAspLeu--------- 34
                 |||               :::   :::|||||||||
Db     174822 CATCCCCATAAAGGACAATAAATGGTAAAGCCCAAGACAAGGAGGCAAACTACACCCTAG 174763

Qy         35 ------------ThrAsnGluGluThrThr------------AspSerThrThrSerLys 46
                        ||||||      ||||||            |||      |||:::
Db     174762 AAAAAGCAAGAAACTAACCGTTTTACAACAAAACAAGGAGAAGACAAGCACACAAACATA 174703

Qy         47 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
              |||||||||           :::   ||||||   :::   :::||||||||||:::
Db     174702 ATCTCTCCTCCAAATACAAAGAAAACAGAAAACAAAAATTGCTATTCCTTAATATCTCTC 174643

Qy         67 AsnIleAspGlyLeuAsp 72
              ||||||:::|||||||:::
Db     174642 AACATCAATGGTCTCAAC 174625
```

RESULT 57
AC128484

```
LOCUS       AC128484               239241 bp    DNA     linear   HTG 21-SEP-2002
DEFINITION  Rattus norvegicus clone CH230-172L11, *** SEQUENCING IN PROGRESS
            ***, 2 unordered pieces.
ACCESSION   AC128484
VERSION     AC128484.2  GI:23265131
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 239241)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 239241)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 239241)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 21, 2002 this sequence version replaced gi:21909206.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequening reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/).  As a result, the
            sequence may extend beyond the ends of the clone and there may be
            contigs that consist entirely of whole genome shotgun sequence
            reads.  Both end sequences and whole genome shotgun sequence only
            contigs will be indicated in the feature table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GYHA
                Center clone name: CH230-172L11
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 213742 bases at least Q40
                Consensus quality: 216345 bases at least Q30
                Consensus quality: 217596 bases at least Q20
                Estimated insert size: 238142; sum-of-contigs estimation
                Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   238012: contig of 238012 bp in length
            *  238013   238112: gap of unknown length
            *  238113   239241: contig of 1129 bp in length.
FEATURES             Location/Qualifiers
     source          1. .239241
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-172L11"
     misc_feature    1. .1177
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    5831. .6663
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH292730"
```

```
     misc_feature    210665. .211873
                     /note="wgs_contig"
     misc_feature    232062. .232918
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH292731"
     misc_feature    235232. .236396
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    236613. .238012
                     /note="wgs_end_extension
                     clone_end:Sp6"
     gap             238013. .238112
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:              2.15e+03        Length:         239241
Score:                  80.00           Matches:        28
Percent Similarity:     46.5%           Conservative:   12
Best Local Similarity:  32.6%           Mismatches:     32
Query Match:            17.9%           Indels:         14
DB:                     14              Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x AC128484 (1-239241)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||:::      |||      ||||||   :::          |||:::
Db     180874 ATGGAGAAACCAAGATATTCCATGACAGAACCAAATTTACACAATATCTTTCTACAAATC 180933

Qy         21 ArgProGluThr---------IleSerGluProLysThrTyrValAspLeu--------- 34
                |||              :::   :::||||||||||
Db     180934 CATCCCCATAAAGGACAATAAATGGTAAAGCCCAAGACAAGGAGGCAAACTACACCCTAG 180993

Qy         35 ------------ThrAsnGluGluThrThr------------AspSerThrThrSerLys 46
                        ||||||      ||||||            |||      |||:::
Db     180994 AAAAAGCAAGAAACTAACCGTTTTACAACAAAACAAGGAGAAGACAAGCACACAAACATA 181053

Qy         47 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
              |||||||||        :::   ||||||   :::   :::||||||||||:::
Db     181054 ATCTCTCCTCCAAATACAAAGAAAACAGAAAACAAAAATTGCTATTCCTTAATATCTCTC 181113

Qy         67 AsnIleAspGlyLeuAsp 72
              ||||||:::|||||||:::
Db     181114 AACATCAATGGTCTCAAC 181131

RESULT 58
CP000099_47
WPCOMMENT
  Sequence split into 49 fragments  LOCUS CP000099 Accession CP000099
     Fragment Name      Begin     End
     CP000099_00            1    110000
     CP000099_01       100001    210000
     CP000099_02       200001    310000
     CP000099_03       300001    410000
     CP000099_04       400001    510000
     CP000099_05       500001    610000
     CP000099_06       600001    710000
     CP000099_07       700001    810000
     CP000099_08       800001    910000
     CP000099_09       900001   1010000
     CP000099_10      1000001   1110000
     CP000099_11      1100001   1210000
     CP000099_12      1200001   1310000
     CP000099_13      1300001   1410000
     CP000099_14      1400001   1510000
     CP000099_15      1500001   1610000
     CP000099_16      1600001   1710000
     CP000099_17      1700001   1810000
     CP000099_18      1800001   1910000
     CP000099_19      1900001   2010000
     CP000099_20      2000001   2110000
     CP000099_21      2100001   2210000
     CP000099_22      2200001   2310000
     CP000099_23      2300001   2410000
     CP000099_24      2400001   2510000
     CP000099_25      2500001   2610000
     CP000099_26      2600001   2710000
     CP000099_27      2700001   2810000
     CP000099_28      2800001   2910000
     CP000099_29      2900001   3010000
     CP000099_30      3000001   3110000
     CP000099_31      3100001   3210000
     CP000099_32      3200001   3310000
     CP000099_33      3300001   3410000
     CP000099_34      3400001   3510000
     CP000099_35      3500001   3610000
     CP000099_36      3600001   3710000
     CP000099_37      3700001   3810000
     CP000099_38      3800001   3910000
     CP000099_39      3900001   4010000
     CP000099_40      4000001   4110000
     CP000099_41      4100001   4210000
     CP000099_42      4200001   4310000
     CP000099_43      4300001   4410000
     CP000099_44      4400001   4510000
     CP000099_45      4500001   4610000
     CP000099_46      4600001   4710000
     CP000099_47      4700001   4810000
     CP000099_48      4800001   4837408
Continuation (48 of 49) of CP000099 from base 4700001 (CP000099 Methanosarcina barkeri s

Alignment Scores:
Pred. No.:              1.03e+03        Length:         110000
Score:                  79.50           Matches:        22
Percent Similarity:     51.9%           Conservative:   20
Best Local Similarity:  27.2%           Mismatches:     30
Query Match:            17.7%           Indels:         9
DB:                     1               Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x CP000099_47 (1-110000)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              ||||||:::::||||||      :::   :::|||   :::|||:::   ||||||
Db      78249 TTAAATACTTTTTTTGAA------ATTATAGATTCAGAAATTGAAAAAAATCCTGAACTT 78302

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||   |||:::          :::|||   :::
Db      78303 AGAGACTACCTTAAGCTTTACTTAGCGAAAAAAGCACATAAAACAATTGAAGAAATAAAT 78362

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 ::::::         ::::::|||   ||||||              :::
Db      78363 CAAGAACTT---------AAAGAAACACATCAAGAG-----------ATCCAAAAATTA 78401

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCys 84
              ::::::   ||||||||||||   :::|||||||||      ::::::      :::
Db      78402 TCCTTCAAAATTGATGGACTTCTAATCAATAATCTTGATAACAAGTCTGCTCCTCTCACA 78461

Qy         85 Ser 85
              |||
Db      78462 TCT 78464

RESULT 59
CLEGCGA/c
LOCUS       CLEGCGA                143171 bp    DNA     circular PLN 18-APR-2005
DEFINITION  Euglena gracilis chloroplast complete genome.
ACCESSION   X70810
VERSION     X70810.2  GI:15526340
KEYWORDS    16S ribosomal RNA; 16S rRNA gene; 23S ribosomal RNA; 23S rRNA gene;
            30S ribosomal protein S11; 30S ribosomal protein S12; 30S ribosomal
            protein S14; 30S ribosomal protein S18; 30S ribosomal protein S19;
            30S ribosomal protein S2; 30S ribosomal protein S3; 30S ribosomal
            protein S4; 30S ribosomal protein S7; 30S ribosomal protein S8; 50S
            ribosomal protein L14; 50S ribosomal protein L16; 50S ribosomal
            protein L2; 50S ribosomal protein L20; 50S ribosomal protein L22;
```

```
            50S ribosomal protein L23; 50S ribosomal protein L36; 50S ribosomal
            protein L5; 5S ribosomal RNA; 5S rRNA gene; alternative splicing;
            ATP synthase beta subunit; ATP synthase epsilon subunit; ATP
            synthase, CF0-subunit I; ATP synthase, CF0-subunit III; ATP
            synthase, CF0-subunit IV; ATP synthase, CF1-alpha subunit; atpA
            gene; atpB gene; atpE gene; atpF gene; atpH gene; atpI gene; ccsA
            gene; complete genome; cytochrome B559 alpha subunit; cytochrome
            B559 beta subunit; cytochrome b6; cytochrome b6-f complex, subunit
            V; elongation factor Ef-Tu; J-polypeptide; MAT1 gene; mat2 gene;
            petB gene; petG gene; psaA gene; psaB gene; psaC gene; psaJ gene;
            psaM gene; psbA gene; psbB gene; psbC gene; psbD gene; psbE gene;
            psbF gene; psbH gene; psbI gene; psbK gene; psbL gene; psbT gene;
            PSI Fe-S polypeptide; PSI M-polypeptide; PSI P700 apoprotein,
            subunit 1a; PSI P700 apoprotein, subunit 1b; PSI polypeptide; PSII
            4.8 kD polypeptide; PSII D1-polypeptide; PSII D2-polypeptide; PSII
            K-polypeptide; PSII L-polypeptide; PSII polypeptide (CP43);
            PSII(CP47) polypeptide; PSII, 10 kDa phosphoprotein; rbcL gene;
            repeat region; ribosomal protein L12; RNA polymerase subunit; roaA
            gene; rpl12 gene; rpl14 gene; rpl16 gene; rpl2 gene; rpl20 gene;
            rpl22 gene; rpl23 gene; rpl32 gene; rpl36 gene; rpl5 gene; rpoA
            gene; rpoB gene; rpoC1 gene; rpoC2 gene; rps11 gene; rps12 gene;
            rps14 gene; rps18 gene; rps19 gene; rps2 gene; rps3 gene; rps4
            gene; rps7 gene; rps8 gene; rps9 gene; rubisco large subunit;
            transfer RNA-Ala; transfer RNA-Arg; transfer RNA-Asn; transfer
            RNA-Asp; transfer RNA-Cys; transfer RNA-Gln; transfer RNA-Glu;
            transfer RNA-Gly; transfer RNA-His; transfer RNA-Ile; transfer
            RNA-Leu; transfer RNA-Lys; transfer RNA-Met; transfer RNA-Phe;
            transfer RNA-Pro; transfer RNA-Ser; transfer RNA-Thr; transfer
            RNA-Trp; transfer RNA-Tyr; transfer RNA-Val; tRNA-Ala gene;
            tRNA-Arg gene; tRNA-Asn gene; tRNA-Asp gene; tRNA-Cys gene;
            tRNA-Gln gene; tRNA-Glu gene; tRNA-Gly gene; tRNA-His gene;
            tRNA-Ile gene; tRNA-Leu gene; tRNA-Lys gene; tRNA-Met gene;
            tRNA-Phe gene; tRNA-Pro gene; tRNA-Ser gene; tRNA-Thr gene;
            tRNA-Trp gene; tRNA-Tyr gene; tRNA-Val gene; tufA gene; ycf12 gene;
            ycf4 gene; ycf9 gene.
SOURCE      chloroplast Euglena gracilis
  ORGANISM  Euglena gracilis
            Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
REFERENCE   1
  AUTHORS   Schlunegger,B. and Stutz,E.
  TITLE     The Euglena gracilis chloroplast genome: structural features of a
            DNA region possibly carrying the single origin of DNA replication
  JOURNAL   Curr. Genet. 8, 629-634 (1984)
REFERENCE   2
  AUTHORS   Hallick,R.B., Hong,L., Drager,R.G., Favreau,M.R., Monfort,A.,
            Orsat,B., Spielmann,A. and Stutz,E.
  TITLE     Complete sequence of Euglena gracilis chloroplast DNA
  JOURNAL   Nucleic Acids Res. 21 (15), 3537-3544 (1993)
   PUBMED   8346031
REFERENCE   3
  AUTHORS   Stevenson,J.K. and Hallick,R.B.
  TITLE     The psaA operon pre-mRNA of the Euglena gracilis chloroplast is
            processed into photosystem I and II mRNAs that accumulate
            differentially depending on the conditions of cell growth
  JOURNAL   Plant J. 5 (2), 247-260 (1994)
   PUBMED   8148879
REFERENCE   4
  AUTHORS   Hong,L. and Hallick,R.B.
  TITLE     Gene structure and expression of a novel Euglena gracilis
            chloroplast operon encoding cytochrome b6 and the beta and epsilon
            subunits of the H(+)-ATP synthase complex
  JOURNAL   Curr. Genet. 25 (3), 270-281 (1994)
   PUBMED   7923415
REFERENCE   5
  AUTHORS   Orsat,B., Spielmann,A., Marc-Martin,S., Lemberger,T. and Stutz,E.
  TITLE     Analysis of the 22 kbp long psbD-psbC gene cluster of Euglena
            gracilis chloroplast DNA: evidence for overlapping transcription
            units undergoing differential processing
  JOURNAL   Biochim. Biophys. Acta 1218 (1), 75-81 (1994)
   PUBMED   8193167
REFERENCE   6
  AUTHORS   Hong,L. and Hallick,R.B.
  TITLE     A group III intron is formed from domains of two individual group
            II introns
  JOURNAL   Genes Dev. 8 (13), 1589-1599 (1994)
   PUBMED   7958842
REFERENCE   7
  AUTHORS   Jenkins,K.P., Hong,L. and Hallick,R.B.
  TITLE     Alternative splicing of the Euglena gracilis chloroplast roaA
            transcript
  JOURNAL   RNA 1 (6), 624-633 (1995)
   PUBMED   7489521
REFERENCE   8
  AUTHORS   Hallick,R.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-1993) R.B. Hallick, University of Arizona, Dept.
            of Biochemistry, Bioscience West 524, Tucson, Arizona 85721, USA
  REMARK    Revised by [13]
REFERENCE   9  (bases 1 to 143171)
  AUTHORS   Hallick,R.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-AUG-2000) R.B. Hallick, University of Arizona, Dept.
            of Biochemistry, Bioscience West 524, Tucson, Arizona 85721, USA
COMMENT     On Sep 7, 2001 this sequence version replaced gi:415327.
            Related sequences X03851, X65484, X04509, X00929, X00735, X15903,
            X17191, X12890, X13310, X05005, Z11874, M12109, M22010, J01414,
            K02392, K02483, V00159.
FEATURES             Location/Qualifiers
     source          1. .143171
                     /organism="Euglena gracilis"
                     /organelle="plastid:chloroplast"
                     /mol_type="genomic DNA"
                     /strain="Z"
                     /db_xref="taxon:3039"
     intron          complement(join(30467. .30512,30309. .30355))
                     /note="external group III intron of twintron"
                     /number=2
     intron          complement(join(89059. .89399,88643. .88656))
                     /note="external group II intron of twintron
                     number 1B"
     intron          complement(join(94551. .94761,94132. .94146,93853. .94025))
                     /note="external group II intron of twintron
                     number 1B"
     intron          complement(join(96895. .96911,96209. .96293))
                     /note="group III
                     alternate to internal intron C splicing, 102 nt variant
                     number 1D
                     intron 1D has alternate 5'-splice sites"
     intron          complement(join(96895. .96913,96209. .96293))
                     /note="group III
                     alternate to internal intron C splicing, 104 nt variant
                     number 1D
                     intron 1D has alternate 5'-splice sites"
     intron          complement(join(96895. .97009,96254. .96293,95658. .95861))
                     /note="external group II intron of twintron
                     number 1B"
     intron          complement(join(106714. .106883,106606. .106617))
                     /note="external group III intron of twintron"
                     /number=11
     intron          complement(join(106724. .106883,106606. .106617))
                     /note="external group III intron of twintron, alternate
                     splicing variant"
                     /number=11
     intron          complement(join(109232. .109237,109115. .109129))
                     /note="external group III intron of twintron"
                     /number=3
     intron          complement(join(109820. .109915,109273. .109706))
                     /note="external group III intron of twintron"
                     /number=1
     gene            complement(89. .137)
                     /gene="tRNA-Trp"
     tRNA            complement(89. .137)
                     /gene="tRNA-Trp"
                     /product="tRNA-Trp"
                     /pseudo
     repeat_unit     180. .275
```

```
                /rpt_type=INVERTED
     gene       complement(2011. .2089)
                /gene="tRNA-Leu"
     tRNA       complement(2011. .2089)
                /gene="tRNA-Leu"
                /product="tRNA-Leu"
                /note="codon recognized: UUG"
                /anticodon=(pos:2054. .2056,aa:Leu)
     gene       complement(<2171. .3549)
                /gene="ccsA"
     CDS        complement(join(2171. .3152,3485. .3549))
                /gene="ccsA"
                /codon_start=1
                /transl_table=11
                /product="CcsA protein"
                /protein_id="CAA50075.1"
                /db_xref="GI:295952"
                /db_xref="GOA:P31205"
                /db_xref="InterPro:IPR000523"
                /db_xref="InterPro:IPR003593"
                /db_xref="UniProt/Swiss-Prot:P31205"
                /translation="MNKKTNERPVFPPTSIVGQEEMKLSLILNVIDPKIGGVMIMGDR
                GTGKSTIVRALVDLLPPIDVIENDPYNSDPYDTELMSDDVLEKIKKNEKVSIIQVKTP
                MVDLPLGGTEDRVCGTIDIEKAISEGKKAFEPGLLAQANRGILYVDEVNLLDDHLVDV
                LLDSAASGWNTVEREGVSICHPARFILVGSGNPEEGELRPQLLDRFGMHAQIKTLKEP
                ALRVKIVQQRELFEKSPKEFKEKYKEEQNKLMEKIINARKKLKNIIIKYELLEKISQI
                CSELNVDGLRGDMVTSRAAKALVAFEDRTEVTPKDIFTVITLCLRHRLRKDPLESIDS
                GYKVQETFKKVFNY"
     exon       complement(<2171. .3152)
                /gene="ccsA"
                /number=2
     intron     complement(3153. .3484)
                /gene="ccsA"
                /note="group II"
                /number=1
     exon       complement(<3485. .3549)
                /gene="ccsA"
                /number=1
     gene       join(<3688. .3930,3931. .5028,5029. .5061,5062. .5425,
                5426. .5513,5514. .6118,6119. .6190,6191. .6841,6842. .6882,
                6883. .7380,7381. .7495,7496. .8101,8102. .8164,8165. .8744,
                8745. .8903,8904. .12561,12562. .12677,12678. .13050,
                13051. .13176,13720. .>13722)

Alignment Scores:
Pred. No.:                1.39e+03        Length:          143171
Score:                    79.50           Matches:         26
Percent Similarity:       44.9%           Conservative:    18
Best Local Similarity:    26.5%           Mismatches:      43
Query Match:              17.7%           Indels:          11
DB:                       15              Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x CLEGCGA (1-143171)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              :::|||:::|||      :::||||||   |||   :::          |||::::::
Db      84293 ATTGAGAAAGCAAGGGGAAATTATTTTCTACCTAAAATTTTCCTTGAAGTTTTAAAAAAA 84234

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
                ||||||   |||   :::   |||            |||      :::
Db      84233 GGTCCGGAAGAGATTGATAAAGAAAAAAGAAAAACTGTTTTAGCAAAAAAAATCGATAAT 84174

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu----------- 56
              |||            ||||||         ::::::   |||   |||
Db      84173 GACGAAGAAATAGAAAAAATAGAAATAAAAAAAGAAAAACAATTTGAAATTAAGAGCTTT 84114

Qy         57 ---AsnGlySerMetPheSerLeuIleThrTrpAsnIle----------AspGly 70
                 |||   |||      ||||||                        |||
Db      84113 TTCGAGGATTCAGAATTTTTAGCAATAACCGAAAAATATACAAAAAAGTGTGAGGATTTA 84054

Qy         71 LeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVal---CysSerTyrLeu 87
              :::::::::||||||      ::::::   |||   :::   |||:::|||:::
Db      84053 ATAAATATTAACAATAAAGGGAAAAAAATAAGAGATTTATTTTGCAATTATATT 84000
```

```
RESULT 60
AC152575
LOCUS       AC152575               214994 bp    DNA     linear   HTG 01-JUL-2005
DEFINITION  Bos taurus clone CH240-14A22, *** SEQUENCING IN PROGRESS ***, 26
            unordered pieces.
ACCESSION   AC152575
VERSION     AC152575.3  GI:68226842
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 214994)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 214994)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2004) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 214994)
```

```
AUTHORS   Cow Genome Sequencing Consortium.
TITLE     Direct Submission
JOURNAL   Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Jun 26, 2005 this sequence version replaced gi:58037717.
          The sequence in this assembly is a combination of BAC based reads
          and whole genome shotgun sequencing reads assembled using Atlas
          (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
          in the feature table below represents a scaffold in the Atlas
          assembly (a 'contig-scaffold'). Within each contig-scaffold,
          individual sequence contigs are ordered and oriented, and separated
          by sized gaps filled with Ns to the estimated size. The sequence
          may extend beyond the ends of the clone and there may be sequence
          contigs within a contig-scaffold that consist entirely of whole
          genome shotgun sequence reads. Both end sequences and whole genome
          shotgun sequence only contigs will be indicated in the feature
          table.
          -------------- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
          -------------- Project Information
              Center project name: FAUB
              Center clone name: CH240-14A22
          -------------- Summary Statistics
              Assembly program: Atlas 3.0;
              Consensus quality: 200225 bases at least Q40
              Consensus quality: 202716 bases at least Q30
              Consensus quality: 205063 bases at least Q20
              Estimated insert size: 206058; sum-of-contigs estimation
              Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
          --------------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 26 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *        1     2265: contig of 2265 bp in length
          *     2266     2424: gap of 159 bp
          *     2425     4974: contig of 2550 bp in length
          *     4975     5024: gap of 50 bp
          *     5025    17462: contig of 12438 bp in length
          *    17463    17512: gap of 50 bp
          *    17513    32196: contig of 14684 bp in length
          *    32197    32296: gap of unknown length
          *    32297    38694: contig of 6398 bp in length
          *    38695    38744: gap of 50 bp
          *    38745    47441: contig of 8697 bp in length
          *    47442    47491: gap of 50 bp
          *    47492    49154: contig of 1663 bp in length
          *    49155    54827: gap of 5673 bp
          *    54828    72490: contig of 17663 bp in length
          *    72491    72540: gap of 50 bp
          *    72541    79037: contig of 6497 bp in length
          *    79038    79087: gap of 50 bp
          *    79088    90993: contig of 11906 bp in length
          *    90994    91043: gap of 50 bp
          *    91044    98880: contig of 7837 bp in length
          *    98881    98930: gap of 50 bp
          *    98931   112165: contig of 13235 bp in length
          *   112166   112215: gap of 50 bp
          *   112216   122234: contig of 10019 bp in length
          *   122235   122598: gap of 364 bp
          *   122599   124664: contig of 2066 bp in length
          *   124665   124714: gap of 50 bp
          *   124715   128852: contig of 4138 bp in length
          *   128853   128902: gap of 50 bp
          *   128903   136723: contig of 7821 bp in length
          *   136724   136823: gap of unknown length
          *   136824   147938: contig of 11115 bp in length
          *   147939   147988: gap of 50 bp
          *   147989   168354: contig of 20366 bp in length
          *   168355   168404: gap of 50 bp
          *   168405   204969: contig of 36565 bp in length
          *   204970   205069: gap of unknown length
          *   205070   206072: contig of 1003 bp in length
          *   206073   206172: gap of unknown length
          *   206173   207738: contig of 1566 bp in length
          *   207739   207838: gap of unknown length
          *   207839   209024: contig of 1186 bp in length
          *   209025   209124: gap of unknown length
          *   209125   210197: contig of 1073 bp in length
          *   210198   210297: gap of unknown length
          *   210298   211403: contig of 1106 bp in length
          *   211404   211503: gap of unknown length
          *   211504   213223: contig of 1720 bp in length
          *   213224   213323: gap of unknown length
          *   213324   214994: contig of 1671 bp in length.
FEATURES             Location/Qualifiers
     source          1. .214994
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9913"
                     /clone="CH240-14A22"
     gap             2266. .2424
                     /estimated_length=159
     gap             4975. .5024
                     /estimated_length=50
     gap             17463. .17512
                     /estimated_length=50
     gap             32197. .32296
                     /estimated_length=unknown
     gap             38695. .38744
                     /estimated_length=50
     gap             47442. .47491
                     /estimated_length=50
     gap             49155. .54827
                     /estimated_length=5673
     gap             72491. .72540
                     /estimated_length=50
     gap             79038. .79087
                     /estimated_length=50
     gap             90994. .91043
                     /estimated_length=50
     gap             98881. .98930
                     /estimated_length=50
     gap             112166. .112215
                     /estimated_length=50
     gap             122235. .122598
                     /estimated_length=364
     gap             124665. .124714
                     /estimated_length=50
     gap             128853. .128902

Alignment Scores:
Pred. No.:              2.19e+03        Length:         214994
Score:                  79.50           Matches:        19
Percent Similarity:     52.8%           Conservative:   9
Best Local Similarity:  35.8%           Mismatches:     24
Query Match:            17.7%           Indels:         1
DB:                     14              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x AC152575 (1-214994)

Qy          6 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle 25
              ||||||:::      |||         ||||||      :::   :::|||:::|||
Db     179072 AATAGCCACCCAAGTCCAAGAAACCCAGAGAGTCCCAAACAGGATAAACCCAAGACA--- 179128

Qy         26 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 45
```

```
                    |||||||||:::::::::   |||||||||:::   |||      |||
Db      179129 AAACACCCCAAGACACATATTAATCAAACTAACGAAGATCAAACACAAAGAACAAATATT 179188

Qy          46 LysIleSerProSerGluAspThrGlnGlnGluAsnGly 58
               |||   |||     :::   |||          |||
Db      179189 AAAAGCAGCAAGGGAAAAACAACATATAACACACAAGGG 179227

RESULT 61
AY228365/c
LOCUS       AY228365               479 bp    mRNA    linear   PRI 11-MAR-2004
DEFINITION  Homo sapiens SLC36A3 mRNA, partial sequence.
ACCESSION   AY228365
VERSION     AY228365.1  GI:29501539
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 479)
  AUTHORS   Bermingham,J.R. Jr. and Pennington,J.
  TITLE     Organization and expression of the SLC36 cluster of amino acid
            transporter genes
  JOURNAL   Mamm. Genome 15 (2), 114-125 (2004)
REFERENCE   2  (bases 1 to 479)
  AUTHORS   Bermingham,J.R. Jr. and Pennington,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2003) McLaughlin Research Institute, 1520 23rd
            Street South, Great Falls, MT 59405, USA
FEATURES             Location/Qualifiers
     source          1. .479
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q33"
                     /clone="IMAGE: 1837427"
     gene            <1. .479
                     /gene="SLC36A3"
                     /note="missing stop codon"
ORIGIN

Alignment Scores:
Pred. No.:              2.63          Length:        479
Score:                  79.00         Matches:       23
Percent Similarity:     43.3%         Conservative:  6
Best Local Similarity:  34.3%         Mismatches:    36
Query Match:            17.6%         Indels:        2
DB:                     8             Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x AY228365 (1-479)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               |||   ||||||                  |||      |||           |||
Db         275 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 216

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               |||         :::::::|||      |||  |||      |||     |||  |||
Db         215 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 162

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               |||   :::|||:::           |||:::     :::      ||||||   |||
Db         161 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 102

Qy          61 PheSerLeuIleThrTrpAsn 67
                  |||      |||||||||
Db         101 ATCTCAGCTGGGACGTGGAAC 81

RESULT 62
CQ728084/c
LOCUS       CQ728084               1006 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 14018 from Patent WO02068579.
ACCESSION   CQ728084
VERSION     CQ728084.1  GI:42295966
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 14018 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1006
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              6.07          Length:        1006
Score:                  79.00         Matches:       23
Percent Similarity:     43.3%         Conservative:  6
Best Local Similarity:  34.3%         Mismatches:    36
Query Match:            17.6%         Indels:        2
DB:                     6             Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x CQ728084 (1-1006)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               |||   ||||||                  |||      |||           |||
Db         836 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 777

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               |||         :::::::|||      |||  |||      |||     |||  |||
Db         776 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 723

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               |||   :::|||:::           |||:::     :::      ||||||   |||
Db         722 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 663

Qy          61 PheSerLeuIleThrTrpAsn 67
                  |||      |||||||||
Db         662 ATCTCAGCTGGGACGTGGAAC 642

RESULT 63
CQ728085/c
LOCUS       CQ728085               1093 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 14019 from Patent WO02068579.
ACCESSION   CQ728085
VERSION     CQ728085.1  GI:42295969
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 14019 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1093
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                6.66          Length:          1093
Score:                    79.00         Matches:         23
Percent Similarity:       43.3%         Conservative:    6
Best Local Similarity:    34.3%         Mismatches:      36
Query Match:              17.6%         Indels:          2
DB:                       6             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x CQ728085 (1-1093)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                          |||      |||          |||
Db        923 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 864

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||         ::::::|||      |||  |||      |||      |||  |||
Db        863 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 810

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::              |||:::     :::     ||||||   |||
Db        809 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 750

Qy         61 PheSerLeuIleThrTrpAsn 67
              |||      |||||||||
Db        749 ATCTCAGCTGGGACGTGGAAC 729

RESULT 64
AX571858/c
LOCUS       AX571858               1473 bp    DNA     linear   PAT 29-NOV-2002
DEFINITION  Sequence 1 from Patent WO0248363.
ACCESSION   AX571858
VERSION     AX571858.1  GI:26003992
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Webster,M., di Francesco,V. and Beasley,E.M.
  TITLE     Human transporter proteins, nucleic acid molecules encoding them,
            and uses thereof
  JOURNAL   Patent: WO 0248363-A 1 20-JUN-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1473
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                9.32          Length:          1473
Score:                    79.00         Matches:         23
Percent Similarity:       43.3%         Conservative:    6
Best Local Similarity:    34.3%         Mismatches:      36
Query Match:              17.6%         Indels:          2
DB:                       6             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x AX571858 (1-1473)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                          |||      |||          |||
Db       1289 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1230

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||         ::::::|||      |||  |||      |||      |||  |||
Db       1229 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1176

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::              |||:::     :::     ||||||   |||
Db       1175 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1116

Qy         61 PheSerLeuIleThrTrpAsn 67
              |||      |||||||||
Db       1115 ATCTCAGCTGGGACGTGGAAC 1095

RESULT 65
BC101094/c
LOCUS       BC101094               1597 bp    mRNA    linear   PRI 02-AUG-2005
DEFINITION  Homo sapiens solute carrier family 36 (proton/amino acid
            symporter), member 3, mRNA (cDNA clone MGC:119640 IMAGE:40009832),
            complete cds.
ACCESSION   BC101094
VERSION     BC101094.1  GI:71682770
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1597)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED    12477932
REFERENCE   2  (bases 1 to 1597)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Baylor Human Genome Sequencing Center
            cDNA Library Preparation: Baylor Human Genome Sequencing Center
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAM Plate: 6 Row: j Column: 11.
FEATURES             Location/Qualifiers
     source          1. .1597
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:119640 IMAGE:40009832"
                     /tissue_type="PCR rescued clones"
                     /clone_lib="NIH_MGC_283"
                     /note="Vector: PCR-Blunt II-TOPO with reversed insert;
```

```
                     Clone identification sequence tag: CATTCATA sequenced from
                     the reverse primer"
     gene            1. .1597
                     /gene="SLC36A3"
                     /note="synonyms: TRAMD2, tramdorin2, PAT3"
                     /db_xref="GeneID:285641"
                     /db_xref="MIM:608332"
     CDS             143. .1555
                     /gene="SLC36A3"
                     /codon_start=1
                     /product="solute carrier family 36 (proton/amino acid
                     symporter), member 3"
                     /protein_id="AAI01095.1"
                     /db_xref="GI:71682771"
                     /db_xref="GeneID:285641"
                     /db_xref="MIM:608332"
                     /translation="MSLLGRDYNSELNSLDNGPQSPSESSSSITSENVHPAGEAGLSM
                     MQTLIHLLKCNIGTGLLGLPLAIKNAGLLVGPVSLLAIGVLTVHCMVILLNCAQHLSQ
                     RLQKTFVNYGEATMYGLETCPNTWLRAHAVWGRYTVSFLLVITQLGFCSVYFMFMADN
                     LQQMVEKAHVTSNICQPREILTLTPILDIRFYMLIILPFLILLVFIQNLKVLSVFSTL
                     ANITTLGSMALIFEYIMEGIPYPSNLPLMANWKTFLLFFGTAIFTFEGVGMVLPLKNQ
                     MKHPQQFSFVLYLGMSIVIILYILLGTLGYMKFGSDTRASITLNLPNCWLYQSVKLMY
                     SIGIFFTYALQFHVPAEIIIPPAISQVSESWALFVDLSVRSALVCLTCVSAILIPRLD
                     LVISLVGSVSSSALALIIPALLEIVIFYSEDMSCVTIAKDIMISIVGLLGCIFGTYQA
                     LYELPQPISHSMANSTGVHA"
ORIGIN

Alignment Scores:
Pred. No.:                10.2          Length:          1597
Score:                    79.00         Matches:         23
Percent Similarity:       43.3%         Conservative:    6
Best Local Similarity:    34.3%         Mismatches:      36
Query Match:              17.6%         Indels:          2
DB:                       8             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x BC101094 (1-1597)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                  |||      |||              |||
Db       1371 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1312

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||       :::::::|||      |||  |||      |||      |||  |||
Db       1311 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1258

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::             |||:::     :::      ||||||   |||
Db       1257 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1198

Qy         61 PheSerLeuIleThrTrpAsn 67
              |||      |||||||||
Db       1197 ATCTCAGCTGGGACGTGGAAC 1177

RESULT 66
BC101095/c
LOCUS       BC101095                1597 bp    mRNA    linear   PRI 12-AUG-2005
DEFINITION  Homo sapiens solute carrier family 36 (proton/amino acid
            symporter), member 3, mRNA (cDNA clone MGC:119641 IMAGE:40009834),
            complete cds.
ACCESSION   BC101095
VERSION     BC101095.2  GI:72533329
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1597)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1597)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Aug 12, 2005 this sequence version replaced gi:71680390.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Baylor Human Genome Sequencing Center
            cDNA Library Preparation: Baylor Human Genome Sequencing Center
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAM Plate: 6 Row: j Column: 12.
FEATURES              Location/Qualifiers
     source           1. .1597
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="MGC:119641 IMAGE:40009834"
                      /tissue_type="PCR rescued clones"
                      /clone_lib="NIH_MGC_283"
                      /note="Vector: pCR-Blunt II-TOPO with reversed insert;
                      Clone identification sequence tag: GATTTACC sequenced from
                      the forward primer"
     gene             1. .1597
                      /gene="SLC36A3"
                      /note="synonyms: TRAMD2, tramdorin2, PAT3"
                      /db_xref="GeneID:285641"
                      /db_xref="MIM:608332"
     CDS              143. .1555
                      /gene="SLC36A3"
                      /codon_start=1
                      /product="solute carrier family 36 (proton/amino acid
                      symporter), member 3"
                      /protein_id="AAI01096.1"
                      /db_xref="GI:72533330"
                      /db_xref="GeneID:285641"
                      /db_xref="MIM:608332"
                      /translation="MSLLGRDYNSELNSLDNGPQSPSESSSSITSENVHPAGEAGLSM
                      MQTLIHLLKCNIGTGLLGLPLAIKNAGLLVGPVSLLAIGVLTVHCMVILLNCAQHLSQ
                      RLQKTFVNYGEATMYGLETCPNTWLRAHAVWGRYTVSFLLVITQLGFCSVYFMFMADN
                      LQQMVEKAHVTSNICQPREILTLTPILDIRFYMLIILPFLILLVFIQNLKVLSVFSTL
                      ANITTLGSMALIFEYIMEGIPYPSNLPLMANWKTFLLFFGTAIFTFEGVGMVLPLKNQ
                      MKHPQQFSFVLYLGMSIVIILYILLGTLGYMKFGSDTQASITLNLPNCWLYQSVKLMY
                      SIGIFFTYALQFHVPAEIIIPFAISQVSESWALFVDLSVRSALVCLTCVSAILIPRLD
```

LVISLVGSVSSSALALIIPALLEIVIFYSEDMSCVTIAKDIMISIVGLLGCIFGTYQA
LYELPQPISHSMANSTGVHA"

ORIGIN

Alignment Scores:

Pred. No.:	10.2	Length:	1597
Score:	79.00	Matches:	23
Percent Similarity:	43.3%	Conservative:	6
Best Local Similarity:	34.3%	Mismatches:	36
Query Match:	17.6%	Indels:	2
DB:	8	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x BC101095 (1-1597)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                 |||      |||              |||
Db       1371 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1312

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||          ::::::|||      |||  |||        |||     |||  |||
Db       1311 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1258

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::            |||:::     :::      ||||||   |||
Db       1257 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1198

Qy         61 PheSerLeuIleThrTrpAsn 67
              |||      |||||||||
Db       1197 ATCTCAGCTGGGACGTGGAAC 1177
```

RESULT 67
BC101092/c

LOCUS	BC101092 1720 bp mRNA linear PRI 02-AUG-2005
DEFINITION	Homo sapiens cDNA clone MGC:119638 IMAGE:40009824, complete cds.
ACCESSION	BC101092
VERSION	BC101092.1 GI:71680387
KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1720)
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
CONSRTM	Mammalian Gene Collection Program Team
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 1720)
AUTHORS	.
CONSRTM	NIH MGC Project
TITLE	Direct Submission
JOURNAL	Submitted (01-AUG-2005) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: Baylor Human Genome Sequencing Center cDNA Library Preparation: Baylor Human Genome Sequencing Center cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers, R. M. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAM Plate: 6 Row: i Column: 11.

FEATURES Location/Qualifiers

source	1. .1720 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:119638 IMAGE:40009824" /tissue_type="PCR rescued clones" /clone_lib="NIH_MGC_283" /note="Vector: pCR-Blunt II-TOPO with reversed insert; Clone identification sequence tag: GTAACCCC sequenced from the reverse primer"
CDS	143. .1678 /codon_start=1 /product="Unknown (protein for MGC:119638)" /protein_id="AAI01093.1" /db_xref="GI:71680388" /translation="MSLLGRDYNSELNSLDNGPQSPSESSSSITSENVHPAGEAGLSM MQTLIHLLKCNIGTGLLGLPLAIKNAGLLVGPVSLLAIGVLTVHCMVILLNCAQHLSQ RLQKTFVNYGEATMYGLETCPNTWLRAHAVWGRWNLALSPRLECSGKISAHCNPHLQG SSNSPAQASRVAGIYRYTVSFLLVITQLGFCSVYFMFMADNLQQMVEKAHVTSNICQP REILTLTPILDIRFYMLIILPFLILLVFIQNLKVLSVFSTLANITTLGSMALIFEYIM EGIPYPSNLPLMANWKTFLLFFGTAIFTFEGVGMVLPLKNQMKHPQQFSFVLYLGMSI VIILYILLGTLGYMKFGSDTQASITLNLPNCWLYQSVKLMYSIGIFFTYALQFHVPAE IIIPFAISQVSESWALFVDLSVRSALVCLTCVSAILIPRLDLVISLVGSVSSSALALI IPALLEIVIFYSEDMSCVTIAKDIMISIVGLLGCIFGTYQALYELPQPISHSMANSTG VHA"

ORIGIN

Alignment Scores:

Pred. No.:	11.1	Length:	1720
Score:	79.00	Matches:	23
Percent Similarity:	43.3%	Conservative:	6
Best Local Similarity:	34.3%	Mismatches:	36
Query Match:	17.6%	Indels:	2
DB:	8	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x BC101092 (1-1720)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                 |||      |||              |||
Db       1494 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1435

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||          ::::::|||      |||  |||        |||     |||  |||
Db       1434 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1381

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::            |||:::     :::      ||||||   |||
Db       1380 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1321

Qy         61 PheSerLeuIleThrTrpAsn 67
              |||      |||||||||
Db       1320 ATCTCAGCTGGGACGTGGAAC 1300
```

RESULT 68
BC101093/c

LOCUS	BC101093 1720 bp mRNA linear PRI 02-AUG-2005
DEFINITION	Homo sapiens cDNA clone MGC:119639 IMAGE:40009826, complete cds.

```
ACCESSION   BC101093
VERSION     BC101093.1  GI:71681850
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1720)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1720)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Baylor Human Genome Sequencing Center
            cDNA Library Preparation: Baylor Human Genome Sequencing Center
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAM Plate: 6 Row: i Column: 12.
FEATURES             Location/Qualifiers
     source          1. .1720
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:119639 IMAGE:40009826"
                     /tissue_type="PCR rescued clones"
                     /clone_lib="NIH_MGC_283"
                     /note="Vector: pCR-Blunt II-TOPO with reversed insert;
                     Clone identification sequence tag: CGAACATC sequenced from
                     the reverse primer"
     CDS             143. .1678
                     /codon_start=1
                     /product="Unknown (protein for MGC:119639)"
                     /protein_id="AAI01094.1"
                     /db_xref="GI:71681851"
                     /translation="MSLLGRDYNSELNSLDNGPQSPSESSSSITSENVHPAGEAGLSM
                     MQTLIHLLKCNIGTGLLGLPLAIKNAGLLVGPVSLLAIGVLTVHCMVILLNCAQHLSQ
                     RLQKTFVNYGEATMYGLETCPNTWLRAHAVWGRWNLALSPRLECSGKISAHCNPHLQG
                     SSNSPAQASRVAGIYRYTVSFLLVITQLGFCSVYFMFMADNLQQMVEKAHVTSNICQP
                     REILTLTPILDIRFYMLIILPFLILLVFIQNLKVLSVFSTLANITTLGSMALIFEYIM
                     EGIPYPSNLPLMANWKTFLLFFGTAIFTFEGVGMVLPLKNQMKHPQQFSFVLYLGMSI
                     VIILYILLGTLGYMKFGSDTQASITLNLPNCWLYQSVKLMYSIGIFFTYALQFHVPAE
                     IIIPFAISQVSESWALFVDLSVRSALVCLTCVSAILIPRLDLVISLVGSVSSSALALI
                     IPALLEIVIFYSEDMSCVTIAKDIMISIVGLLGCIFGTYQALYELPQPISHSMANSTG
                     VHA"
ORIGIN

Alignment Scores:
Pred. No.:                11.1          Length:          1720
Score:                    79.00         Matches:         23
Percent Similarity:       43.3%         Conservative:    6
Best Local Similarity:    34.3%         Mismatches:      36
Query Match:              17.6%         Indels:          2
DB:                       8             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x BC101093 (1-1720)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                    |||       |||             |||
Db       1494 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1435

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||          ::::::|||     |||  |||     |||     |||  |||
Db       1434 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1381

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::           |||:::     :::      ||||||   |||
Db       1380 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1321

Qy         61 PheSerLeuIleThrTrpAsn 67
                 |||      |||||||||
Db       1320 ATCTCAGCTGGGACGTGGAAC 1300

RESULT 69
AY162215/c
LOCUS       AY162215               2268 bp    mRNA    linear   PRI 17-JUN-2003
DEFINITION  Homo sapiens proton/amino acid transporter 3 (SLC36A3) mRNA,
            complete cds.
ACCESSION   AY162215
VERSION     AY162215.1  GI:31871294
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2268)
  AUTHORS   Boll,M., Foltz,M., Rubio-Aliaga,I. and Daniel,H.
  TITLE     A cluster of proton/amino acid transporter genes in the human and
            mouse genomes small star, filled
  JOURNAL   Genomics 82 (1), 47-56 (2003)
   PUBMED   12809675
REFERENCE   2  (bases 1 to 2268)
  AUTHORS   Boll,M., Foltz,M., Rubio-Aliaga,I., Kolmeder,D. and Daniel,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-OCT-2002) Molecular Nutrition Unit, Technical
            University of Munich, Hochfeldweg 2, Freising 85350, Germany
FEATURES             Location/Qualifiers
     source          1. .2268
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q33"
                     /tissue_type="testis"
     gene            1. .2268
                     /gene="SLC36A3"
     CDS             238. .1650
                     /gene="SLC36A3"
                     /note="solute carrier family 36 member 3; PAT3"
                     /codon_start=1
```

```
                     /product="proton/amino acid transporter 3"
                     /protein_id="AAO11789.1"
                     /db_xref="GI:31871295"
                     /translation="MSLLGRDYNSELNSLDNGPQSPSESSSSITSENVHPAGEAGLSM
                     MQTLIHLLKCNIGTGLLGLPLAIKNAGLLVGPVSLLAIGVLTVHCMVILLNCAQHLSQ
                     RLQKTFVNYGEATMYGLETCPNTWLRAHAVWGRYTVSFLLVITQLGFCSVYFMFMADN
                     LQQMVEKAHVTSNICQPREILTLTPILDIRFYMLIILPFLILLVFIQNLKVLSVFSTL
                     ANITTTLGSMALIFEYIMEGIPYPSNLPLMANWKTPLLFFGTAIFTFEGVGMVLPLKNQ
                     MKHPQQFSFVLYLGMSIVIILYILLGTLGYMKFGSDTQASITLNLPNCWLYQSVKLMY
                     SIGIFFTYALQFHVPAEIIIPFAISQVSESWALFVDLSVRSALVCLTCVSAILIPRLD
                     LVISLVGSVSSSALALIIPALLEIVIFYSEDMSCVTIAKDIMISIVGLLGCIFGTYQA
                     LYELPQPISHSMANSTGVHA"
ORIGIN

Alignment Scores:
Pred. No.:                15.1          Length:          2268
Score:                    79.00         Matches:         23
Percent Similarity:       43.3%         Conservative:    6
Best Local Similarity:    34.3%         Mismatches:      36
Query Match:              17.6%         Indels:          2
DB:                       8             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x AY162215 (1-2268)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                          |||       |||           |||
Db       1466 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1407

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||         ::::::|||     |||  |||     |||     |||  |||
Db       1406 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1353

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::          |||:::     :::      ||||||   |||
Db       1352 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1293

Qy         61 PheSerLeuIleThrTrpAsn 67
                 |||       |||||||||
Db       1292 ATCTCAGCTGGGACGTGGAAC 1272

RESULT 70
CQ850796/c
LOCUS       CQ850796                2456 bp    DNA     linear   PAT 23-AUG-2004
DEFINITION  Sequence 1265 from Patent EP1447413.
ACCESSION   CQ850796
VERSION     CQ850796.1  GI:51509008
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Yamamoto,J., Nishikawa,T., Isono,Y., Sugiyama,T.,
            Otsuki,T., Wakamatsu,A., Ishii,S., Nagai,K. and Irie,R.
  TITLE     Full-length human cDNA
  JOURNAL   Patent: EP 1447413-A 1265 18-AUG-2004;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .2456
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                16.6          Length:          2456
Score:                    79.00         Matches:         23
Percent Similarity:       43.3%         Conservative:    6
Best Local Similarity:    34.3%         Mismatches:      36
Query Match:              17.6%         Indels:          2
DB:                       6             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x CQ850796 (1-2456)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                          |||       |||           |||
Db       1641 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1582

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||         ::::::|||     |||  |||     |||     |||  |||
Db       1581 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1528

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::          |||:::     :::      ||||||   |||
Db       1527 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1468

Qy         61 PheSerLeuIleThrTrpAsn 67
                 |||       |||||||||
Db       1467 ATCTCAGCTGGGACGTGGAAC 1447

RESULT 71
AK127978/c
LOCUS       AK127978                2456 bp    mRNA    linear   PRI 19-FEB-2004
DEFINITION  Homo sapiens cDNA FLJ46093 fis, clone TESTI2018335.
ACCESSION   AK127978
VERSION     AK127978.1  GI:34535124
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kanehori,K., Ishibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S.,
            Tanai,H., Watanabe,S., Ishida,S., Ono,Y., Hotuta,T., Watanabe,M.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Takahashi-Fujii,A., Oshima,A., Sugiyama,A.,
            Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
            and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2456)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: Reverse Proteomics Research Institute, HRI and
            RAB.
FEATURES             Location/Qualifiers
     source          1. .2456
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TESTI2018335"
                     /tissue_type="testis"
                     /clone_lib="TESTI2"
                     /note="cloning vector: pME18SFL3"
     CDS             413. .1825
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC87215.1"
                     /db_xref="GI:34535125"
                     /translation="MSLLGRDYNSELNSLDNGPQSPSESSSSITSENVHPAGEAGLSM
                     MQTLIHLLKCNIGTGLLGLPLAIKNAGLLVGPVSLLAIGVLTVHCMVILLNCAQHLSQ
                     RLQKTFVNYGEATMYGLETCPNTWLRAHAVWGRYTVSFLLVITQLGFCSVYFMFMADN
```

```
                    LQQMVEEAHVTSNICQPREILTLTPILDIRFYMLIILPFLILLVFIQNLKVLSVFSTL
                    ANITTLGSMALIFEYIMEGIPYPSNLPLMANWKTFLLFFGTAIFTFEGVGMVLPLKNQ
                    MKHPQQFSFVLYLGMSIVIILYILLGTLGYMKFGSDTQASITLNLPNCWLYQSVKLMY
                    SIGIFFTYALQFHVPAEIIIPFAISQVSESWALFVDLSVRSALVCLTCVSAILIPRLD
                    LVISLVGSVSSSALALIIPALLEIVIFYSEDMSCVTIAKDIMISIVGLLGCIFGTYQA
                    LYELPQPISHSMANSTGVHA"
ORIGIN

Alignment Scores:
Pred. No.:              16.6          Length:         2456
Score:                  79.00         Matches:        23
Percent Similarity:     43.3%         Conservative:   6
Best Local Similarity:  34.3%         Mismatches:     36
Query Match:            17.6%         Indels:         2
DB:                     8             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x AK127978 (1-2456)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                        |||      |||          |||
Db       1641 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1582

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||       ::::::|||      |||  |||      |||      |||  |||
Db       1581 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1528

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::             |||:::     ::: .     ||||||   |||
Db       1527 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1468

Qy         61 PheSerLeuIleThrTrpAsn 67
                 |||      |||||||||
Db       1467 ATCTCAGCTGGGACGTGGAAC 1447

RESULT 72
CQ845983/c
LOCUS       CQ845983               3558 bp    DNA     linear   PAT 02-AUG-2004
DEFINITION  Sequence 4630 from Patent EP1440981.
ACCESSION   CQ845983
VERSION     CQ845983.1  GI:50895582
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J., Isono,Y., Nagai,K. and Irie,R.
  TITLE     Full-length human cdna
  JOURNAL   Patent: EP 1440981-A 4630 28-JUL-2004;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .3558
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              25.1          Length:         3558
Score:                  79.00         Matches:        23
Percent Similarity:     43.3%         Conservative:   6
Best Local Similarity:  34.3%         Mismatches:     36
Query Match:            17.6%         Indels:         2
DB:                     6             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x CQ845983 (1-3558)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                        |||      |||          |||
Db       3129 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 3070

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||       ::::::|||      |||  |||      |||      |||  |||
Db       3069 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 3016

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::             |||:::     ::: .     ||||||   |||
Db       3015 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 2956

Qy         61 PheSerLeuIleThrTrpAsn 67
                 |||      |||||||||
Db       2955 ATCTCAGCTGGGACGTGGAAC 2935

RESULT 73
AK131483/c
LOCUS       AK131483               3558 bp    mRNA    linear   PRI 07-MAY-2004
DEFINITION  Homo sapiens cDNA FLJ16658 fis, clone TESTI4041086.
ACCESSION   AK131483
VERSION     AK131483.1  GI:47077480
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
            Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
            and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3558)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2004) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .3558
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TESTI4041086"
                     /tissue_type="testis"
                     /clone_lib="TESTI4"
                     /note="cloning vector: pME18SFL3"
     CDS             1946. .3313
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAD18628.1"
                     /db_xref="GI:47077481"
                     /translation="MGVIISGSLPPFSSPLQVGPVSLLAIGVLTVHCMVILLNCAQHL
                     SQRLQKTFVNYGEATMYGLETCPNTWLRAHAVWGRWNLALSPRLECSGKISAHCNPHL
                     QGSSNSPAQASRVAGIYRYTVSFLLVITQLGFCSVYFMFMADNLQQMVEKAHVTSNIC
                     QPREILTLTPILDIRFYMLIILPFLILLVFIQNLKVLSVFSTLANITTLGSMALIFEY
                     IMEGIPYPSNLPLMANWKTFLLFFGTAIFTFEGVGMVLPLKNQMKHPQQFSFVLYLGM
                     SIVIILYILLGTLGYMKFGSDTQASITLNLPNCWLYQSVKLMYSIGIFFTYALQFHVP
                     AEIIIPFAISQVSESWALFVDLSVRSALVCLTCVSAILIPRLDLVISLVGSVSSSALA
                     LIIPALLEIVIFYSEDMSCVTIAKDIMISIVGLLGCIFGTYQALYELPQPISHSMANS
                     TGVHA"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                25.1          Length:          3558
Score:                    79.00         Matches:         23
Percent Similarity:       43.3%         Conservative:    6
Best Local Similarity:    34.3%         Mismatches:      36
Query Match:              17.6%         Indels:          2
DB:                       8             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x AK131483 (1-3558)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                  |||      |||             |||
Db       3129 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 3070

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||         ::::::|||      |||   |||      |||      |||   |||
Db       3069 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 3016

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::           |||:::     :::      ||||||   |||
Db       3015 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 2956

Qy         61 PheSerLeuIleThrTrpAsn 67
                 |||      |||||||||
Db       2955 ATCTCAGCTGGGACGTGGAAC 2935

RESULT 74
HUMIGFBP5X
LOCUS       HUMIGFBP5X              3672 bp    mRNA    linear   PRI 23-AUG-2004
DEFINITION  Homo sapiens IGFBP5 mRNA, 3' UTR.
ACCESSION   L27560
VERSION     L27560.1  GI:452059
KEYWORDS    insulin-like growth factor binding protein 5.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kiefer,M.C., Ioh,R.S., Bauer,D.M. and Zapf,J.
  TITLE     Molecular cloning of a new human insulin-like growth factor binding
            protein
  JOURNAL   Biochem. Biophys. Res. Commun. 176 (1), 219-225 (1991)
   PUBMED   1850258
REFERENCE   2  (bases 1 to 3672)
  AUTHORS   Allander,S.V., Larsson,C., Ehrenborg,E., Suwanichkul,A., Weber,G.,
            Morris,S.L., Bajalica,S., Kiefer,M.C., Luthman,H. and Powell,D.R.
  TITLE     Characterization of the chromosomal gene and promoter for human
            insulin-like growth factor binding protein-5
  JOURNAL   J. Biol. Chem. 269 (14), 10891-10898 (1994)
   PUBMED   7511611
FEATURES             Location/Qualifiers
     source          1. .3672
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="2q33-34"
     gene            <1. .>3672
                     /gene="IGFBP5"
     3'UTR           <1. .>3672
                     /gene="IGFBP5"
ORIGIN

Alignment Scores:
Pred. No.:                34.6          Length:          3672
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       8             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x HUMIGFBP5X (1-3672)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::            ||||||:::    |||      |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||          |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||          |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 75
CQ787212/c
LOCUS       CQ787212                4001 bp    DNA     linear   PAT 24-MAR-2004
DEFINITION  Sequence 62 from Patent WO2004020662.
ACCESSION   CQ787212
VERSION     CQ787212.1  GI:45722195
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maier,S.
  TITLE     Method and nucleic acids for the analysis of breast cell
            proliferative disorders
  JOURNAL   Patent: WO 2004020662-A 62 11-MAR-2004;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .4001
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                38.1          Length:          4001
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       6             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x CQ787212 (1-4001)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::            ||||||:::    |||      |||
Db        625 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 566

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||          |||||||||
Db        565 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 506

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||          |||   ||||||:::|||:::
Db        505 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 446

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db        445 TGTCTGGACTTTGAGCAGGCTT 424


Search completed: March 10, 2006, 22:50:12
Job time : 3369.57 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 15:40:34 ; Search time 9695.3 Seconds
(without alignments)
1746.920 Million cell updates/sec

Title: US-10-757-745-2
Perfect score: 1902
Sequence: 1 MELGSCLEGGREAAEEEGEP.....DCGRFPSDHWGLLCNLDIIL 362

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170329_4058/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_12397_@runat_09032006_170329_4058 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1902	100.0	1089	11	DQ049205	DQ049205 Homo sapi
2	1902	100.0	1168	4	CR601303	CR601303 full-leng
3	1902	100.0	1894	4	CR597293	CR597293 full-leng
4	1902	100.0	1909	4	CR595644	CR595644 full-leng
5	1867	98.2	1089	11	DQ049206	DQ049206 Pan trogl
6	1724	90.6	1620	4	CR592636	CR592636 full-leng
7	1653	86.9	1081	5	BX444691	BX444691 BX444691
8	1616	85.0	1743	4	CR602029	CR602029 full-leng
9	1603	84.3	981	5	BX338160	BX338160 BX338160
10	1572	82.6	1067	5	BX337905	BX337905 BX337905
11	1570	82.5	972	3	BM468826	BM468826 AGENCOURT
12	1545	81.2	1018	5	BX337141	BX337141 BX337141
13	1523	80.1	904	5	BX422491	BX422491 BX422491
14	1497	78.7	883	5	BU179107	BU179107 AGENCOURT
c 15	1483	78.0	952	5	BX433489	BX433489 BX433489
16	1476	77.6	1046	3	BM926092	BM926092 AGENCOURT
17	1468	77.2	942	5	BX358707	BX358707 BX358707
18	1460	76.8	965	7	CN803230	CN803230 ILLUMIGEN
19	1452	76.3	1192	3	BM553049	BM553049 AGENCOURT
20	1447	76.1	1090	7	CN641671	CN641671 ILLUMIGEN
21	1433.5	75.4	948	1	AL555333	AL555333 AL555333
22	1413	74.3	1103	3	BM555041	BM555041 AGENCOURT
23	1390	73.1	846	7	CR765451	CR765451 DKFZp469E
24	1389	73.0	870	2	BG740396	BG740396 602634171
25	1381	72.6	1087	7	CN642527	CN642527 ILLUMIGEN
26	1367	71.9	810	8	CX166335	CX166335 HESC2_39
27	1366	71.8	922	7	CN801806	CN801806 ILLUMIGEN
28	1341.5	70.5	870	5	BU159911	BU159911 AGENCOURT
29	1329.5	69.9	1035	3	BM554324	BM554324 AGENCOURT
30	1327	69.8	906	6	CD251503	CD251503 AGENCOURT
31	1299.5	68.3	838	3	BI908925	BI908925 603067028
32	1299.5	68.3	932	2	BG740339	BG740339 602635289
33	1293	68.0	783	5	BX365835	BX365835 BX365835
34	1290	67.8	780	5	BX374579	BX374579 BX374579
35	1272.5	66.9	858	3	BI553412	BI553412 603193358
36	1240	65.2	883	5	BU169945	BU169945 AGENCOURT
37	1238	65.1	767	2	BI258848	BI258848 602969663
38	1235	64.9	937	2	BG541819	BG541819 602569712
39	1234	64.9	729	7	CN298922	CN298922 170006000
40	1234	64.9	729	7	CN298924	CN298924 170004244
41	1197	62.9	711	7	CN298923	CN298923 170006000
42	1190.5	62.6	731	8	CX760857	CX760857 AGENCOURT
43	1190	62.6	749	7	CR767800	CR767800 DKFZp469C
44	1188	62.5	710	5	BU661472	BU661472 cl72d11.z
45	1187	62.4	704	3	BI334820	BI334820 602998939
46	1187	62.4	757	3	BI760756	BI760756 603044763
47	1182	62.1	752	1	AU139147	AU139147 AU139147
48	1180	62.0	701	7	CR767236	CR767236 DKFZp469E
49	1162.5	61.1	816	1	AL516230	AL516230 AL516230
50	1159	60.9	1289	4	AK009089	AK009089 Mus muscu
51	1155	60.7	739	7	CR791801	CR791801 DKFZp468J
52	1153	60.6	774	3	BI754101	BI754101 603027659
53	1133	59.6	657	7	CN298920	CN298920 170006001
54	1129	59.4	678	2	BG719113	BG719113 602699045
55	1118.5	58.8	853	2	BE784416	BE784416 601473891
56	1116	58.7	798	8	DN122208	DN122208 1122397 M
57	1112.5	58.5	927	2	BG119064	BG119064 602347589
58	1102	57.9	770	2	BG498689	BG498689 602544301
59	1098	57.7	908	2	BG391213	BG391213 602417244
60	1095	57.6	634	6	CB157906	CB157906 K-EST0217
61	1095	57.6	650	2	BG431303	BG431303 602499879
62	1085	57.0	680	7	CK952220	CK952220 4091710 B
63	1082	56.9	786	2	BG719977	BG719977 602691335
64	1055.5	55.5	995	2	BE892886	BE892886 601435730
65	1050.5	55.2	799	8	CX756424	CX756424 AGENCOURT
66	1045	54.9	621	6	CB069952	CB069952 is31e10.y
67	1044.5	54.9	915	3	BI915865	BI915865 603184527
68	1044	54.9	634	7	CV023369	CV023369 288 Full
69	1043	54.8	909	5	BU928818	BU928818 AGENCOURT
70	1036	54.5	1110	3	BM476771	BM476771 AGENCOURT
71	1030	54.2	602	1	AL703449	AL703449 DKFZp686G
72	1018.5	53.5	793	2	BG533717	BG533717 602562373
73	1018.5	53.5	959	2	BI161201	BI161201 602865659
74	1017.5	53.5	652	7	CR753214	CR753214 DKFZp469I
75	1017	53.5	581	3	BP262101	BP262101 BP262101

ALIGNMENTS

RESULT 1

```
DQ049205
LOCUS       DQ049205               1089 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Homo sapiens TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ049205
VERSION     DQ049205.1  GI:66902404
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1089)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 1089)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.  Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1. .1089
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
     gene            <1. .>1089
                     /gene="TTRAP"
                     /locus_tag="HC17203"
ORIGIN

Alignment Scores:
Pred. No.:              1.84e-212       Length:         1089
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2 (1-362) x DQ049205 (1-1089)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1080

Qy        361 IleLeu 362
              ||||||
Db       1081 ATATTG 1086

RESULT 2
CR601303
LOCUS       CR601303               1168 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DN005YN02 of Adult brain of Homo sapiens
            (human).
ACCESSION   CR601303
VERSION     CR601303.1  GI:50482110
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1168)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
```

```
            Faraday Avenue
REFERENCE   2  (bases 1 to 1168)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1168
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="Adult brain"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Alignment Scores:
Pred. No.:                2.06e-212      Length:          1168
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       4              Gaps:            0

US-10-757-745-2 (1-362) x CR601303 (1-1168)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
              ||||||
Db       1106 ATATTG 1111

RESULT 3
CR597293
LOCUS       CR597293               1894 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DK007YK05 of HeLa cells Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR597293
VERSION     CR597293.1  GI:50478100
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1894)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1894)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK007YK05"
                     /tissue_type="HeLa cells Cot 25-normalized"
```

```
                     /plasmid="pCMVSPORT_6"
ORIGIN

Alignment Scores:
Pred. No.:                4.48e-212       Length:          1894
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       4               Gaps:            0

US-10-757-745-2 (1-362) x CR597293 (1-1894)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 78

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 138

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 198

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 258

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 318

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 378

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 438

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 498

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 558

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 618

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 678

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 738

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 798

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 858

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 918

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 978

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1038

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1039 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1098

Qy        361 IleLeu 362
              ||||||
Db       1099 ATATTG 1104

RESULT 4
CR595644
LOCUS       CR595644               1909 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DI052YN13 of Placenta Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR595644
VERSION     CR595644.1  GI:50476451
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1909)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1909)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI052YN13"
                     /tissue_type="Placenta Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Alignment Scores:
Pred. No.:                4.54e-212       Length:          1909
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       4               Gaps:            0

US-10-757-745-2 (1-362) x CR595644 (1-1909)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 84

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         85 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 144

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60

Db        145 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 204

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db        205 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 264

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        265 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 324

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        325 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 384

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        385 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 444

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        445 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 504

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db        505 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 564

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db        565 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 624

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220

Db        625 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 684

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240

Db        685 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 744

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260

Db        745 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 804

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280

Db        805 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 864

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        865 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 924

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        925 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 984

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340

Db        985 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1044

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360

Db       1045 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1104

Qy        361 IleLeu 362

Db       1105 ATATTG 1110

RESULT 5
DQ049206
LOCUS       DQ049206                1089 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Pan troglodytes TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ049206
VERSION     DQ049206.1  GI:66902405
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 1089)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 1089)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.  Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1. .1089
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>1089
                     /gene="TTRAP"
                     /locus_tag="HC17203"
ORIGIN

Alignment Scores:
Pred. No.:              2.43e-208       Length:         1089
Score:                  1867.00         Matches:        355
Percent Similarity:     98.3%           Conservative:   1
Best Local Similarity:  98.1%           Mismatches:     6
Query Match:            98.2%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2 (1-362) x DQ049206 (1-1089)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20

Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCC 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40

Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTNNNNNGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60

Db        121 GCTGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db        181 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        241 CCCANNNNNNNTGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 660

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              |||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGTGGAACGAATGAATCAG 720

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAATCAGCTACAGTTATATTTGCA 780

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAGCTTCGTTTTGATCGAATATTTTTC 960

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db       1021 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATGTA 1080

Qy        361 IleLeu 362
              ||||||
Db       1081 ATATTG 1086
```

RESULT 6
CR592636

```
LOCUS       CR592636                1620 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DM007YE11 of Fetal liver of Homo sapiens
            (human).
ACCESSION   CR592636
VERSION     CR592636.1  GI:50473443
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1620)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1620)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1..1620
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DM007YE11"
                     /tissue_type="Fetal liver"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Alignment Scores:
Pred. No.:               3.17e-191      Length:          1620
Score:                   1724.00        Matches:         329
Percent Similarity:      100.0%         Conservative:    0
Best Local Similarity:   100.0%         Mismatches:      0
Query Match:             90.6%          Indels:          0
DB:                      4              Gaps:            0

US-10-757-745-2 (1-362) x CR592636 (1-1620)

Qy         20 ProGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAsp 39
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGAT 60

Qy         40 AlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsn 59
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAAC 120

Qy         60 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCT 180

Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAA 240

Qy        100 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGG 300

Qy        120 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTAC 360

Qy        140 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSer 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGC 420

Qy        160 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTC 480

Qy        180 ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTT 540

Qy        200 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAG 600

Qy        220 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAAT 660
```

```
Qy        240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTT 720

Qy        260 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsn 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAAC 780

Qy        280 IleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThr 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACA 840

Qy        300 GlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTT 900

Qy        320 PheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeu 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTA 960

Qy        340 GluLysLeuAspCysGlyArgPhePro 348
              ||||||||||||||||||||||||||
Db        961 GAAAAACTGGACTGTGGTAGATTTCCT 987
```

```
RESULT 7
BX444691
LOCUS       BX444691               1081 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX444691 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN005YN02 5-PRIME, mRNA sequence.
ACCESSION   BX444691
VERSION     BX444691.2  GI:47009162
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1081)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30780264.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DN005DG01QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .1081
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

Alignment Scores:
Pred. No.:          3.79e-183      Length:       1081
Score:              1653.00        Matches:      340
```

```
Percent Similarity:     96.3%          Conservative:  1
Best Local Similarity:  96.0%          Mismatches:    12
Query Match:            86.9%          Indels:        7
DB:                     5              Gaps:          0

US-10-757-745-2 (1-362) x BX444691 (1-1081)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGly-ArgGluAlaAlaGluGluGluGlyGluPr 20
              |||||||||||||||||||||  |||  |||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGMGGGCGCGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCC 85

Qy         20 oGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAl 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 TGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGC 145

Qy         40 aAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSe 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 CGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTC 205

Qy         60 rTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGl 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGA 265

Qy         80 uProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIl 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 GCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAAT 325

Qy        100 eSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAs 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 CAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAA 385

Qy        120 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 TATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTT 445

Qy        140 uAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTy 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 AGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTA 505

Qy        160 rLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheTh 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCAC 565

Qy        180 rAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 AGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCC 625

Qy        200 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCT 685

Qy        220 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCA 745

Qy        240 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 260
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db        746 GTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGC 805

Qy        260 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||
Db        806 AGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCMAC-AACAT 864

Qy        280 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGl 300
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        865 TGTGGATGTCTGGGAGTTTTTGGGCAA-CCTAAACATTGCCAGTATACATGGGATACACA 923

Qy        300 nMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePh 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 CATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTT 983

Qy        320 eArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGl 340
              ||||  |||||||||||||||||||||||||||  |||||||||  |||||||||
```

```
Db        984 CAGARCAGCAGCAGAAGAGGGACACATTATYCCCGA-AGTTTGGACCYTCTTGGATTAGA 1042

Qy        340 uLysLeuAspCysGlyArgPheProSerAspHisTrpGly 353
              ||||||||||||   ||||||   ||||||   ||||||
Db       1043 -AAACTGGACTGTWGTAGATTYCT-AGTGATAC-TGGGGK 1079

RESULT 8
CR602029
LOCUS       CR602029               1743 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DI042YL19 of Placenta Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR602029
VERSION     CR602029.1  GI:50482836
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1743)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1743)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1743
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI042YL19"
                     /tissue_type="Placenta Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Alignment Scores:
Pred. No.:              1.86e-178       Length:         1743
Score:                  1616.00         Matches:        307
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  99.7%           Mismatches:     0
Query Match:            85.0%           Indels:         0
DB:                     4               Gaps:           0

US-10-757-745-2 (1-362) x CR602029 (1-1743)

Qy         55 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 74
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 91

Qy         75 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 CCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 151

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 TCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTC 211

Qy        115 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 271

Qy        135 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATT 331

Qy        155 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 CCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCAT 391

Qy        175 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAA 451

Qy        195 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAC 511

Qy        215 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 GTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCT 571

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 GCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCA 631

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        632 GCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGT 691

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        692 GGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAG 751

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        752 TATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGT 811

Qy        315 PheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeu 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        812 TTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 871

Qy        335 AspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeu 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        872 GACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTT 931

Qy        355 LeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||||||
Db        932 CTGTGCAACTTAGATATAATATTG 955

RESULT 9
BX338160
LOCUS       BX338160               981 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX338160 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI056YC23 5-PRIME, mRNA sequence.
ACCESSION   BX338160
VERSION     BX338160.2  GI:46283046
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 981)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30345671.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
```

```
          was normalized. Library was constructed by Life Technologies, a
          division of Invitrogen. This sequence belongs to sequence cluster
          3474.r
          For more information about this cluster, see
          http://www.genoscope.cns.fr/cdna?s=CS0DI056AB12QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .981
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI056YC23"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:                2.5e-177      Length:          981
Score:                    1603.00       Matches:         317
Percent Similarity:       99.1%         Conservative:    1
Best Local Similarity:    98.8%         Mismatches:      2
Query Match:              84.3%         Indels:          3
DB:                       5             Gaps:            0

US-10-757-745-2 (1-362) x BX338160 (1-981)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         13 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 72

Qy         21 GluVal-LysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAl 40
              |||||| |||:::|||||||||||||||||||||||||||||||||||||||||||||||
Db         73 GAGGTGAAAACAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGC 132

Qy         40 aAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSe 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 CGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTC 192

Qy         60 rTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGl 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 CTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGA 252

Qy         80 uProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIl 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAAT 312

Qy        100 eSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAs 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 CAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAA 372

Qy        120 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 TATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTT 432

Qy        140 uAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTy 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 AGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTA 492

Qy        160 rLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheTh 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 CCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCAC 552

Qy        180 rAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 AGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCC 612

Qy        200 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 AAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCT 672

Qy        220 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCA 732

Qy        240 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 GTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGC 792

Qy        260 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 AGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACAT 852

Qy        280 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGl 300
              |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        853 TGTGGATGTCTGGGAGTTTTTGGGCAA-CCTAAACATTGCCAGTATACATGGGATACACA 911

Qy        300 nMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePh 320
              |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        912 AATGAACTCTAATCTTGGAATAMT-GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTT 970

Qy        320 e 320
              |
Db        971 C 971

RESULT 10
BX337905
LOCUS       BX337905               1067 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX337905 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI052YN13 5-PRIME, mRNA sequence.
ACCESSION   BX337905
VERSION     BX337905.2  GI:46273926
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1067)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30339657.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI052CG07QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .1067
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI052YN13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:                1.28e-173     Length:          1067
Score:                    1572.00       Matches:         313
Percent Similarity:       96.0%         Conservative:    2
Best Local Similarity:    95.4%         Mismatches:      11
```

```
Query Match:            82.6%          Indels:        4
DB:                     5              Gaps:          1

US-10-757-745-2 (1-362) x BX337905 (1-1067)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         25 ATGGAGTTGGGGAGTTGCCTGGWGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 84

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 144

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 204

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 264

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 324

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 384

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 444

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 504

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 564

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 624

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 684

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 744

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db        745 TTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 804

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCAAC-AACATT 863

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 GTGGATGTCTGGGAGTTTTTKGGCAAACCTAAACATTGCCAGTATACATGGGRTACACAA 923

Qy        301 MetAsnSer-AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePh 320
              |||||||||| |||||||||   ||||||   |||   ||||||||||||      |||||
Db        924 ATGAACTCTHAATCTTGGASTAACTGCTGC-TGTTAACTTCGTTTTGAT---CRATTTTT 979

Qy        320 eArgAlaAlaAlaGluGluGly 327
              |  :::   ||||||   |||
Db        980 TCHGAGCTGCGCAGARGGGGGC 1001
```

```
RESULT 11
BM468826
LOCUS       BM468826                 972 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6445782 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5587154
            5', mRNA sequence.
ACCESSION   BM468826
VERSION     BM468826.1  GI:18517868
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 972)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12356  row: d  column: 03
            High quality sequence stop: 707.
FEATURES             Location/Qualifiers
     source          1. .972
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5587154"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_92"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.9e-173        Length:         972
Score:                  1570.00         Matches:        311
Percent Similarity:     96.6%           Conservative:   1
Best Local Similarity:  96.3%           Mismatches:     5
Query Match:            82.5%           Indels:         6
DB:                     3               Gaps:           0

US-10-757-745-2 (1-362) x BM468826 (1-972)

Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 AGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGG 63

Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAA 123

Qy         77 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 96
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 ACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACC 183

Qy         97 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTC 243

Qy        117 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVal 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTG 303
```

```
Qy        137 CysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProPro 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 TGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCA 363

Qy        157 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAA 423

Qy        177 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATT 483

Qy        197 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 ATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCA 543

Qy        217 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAA 603

Qy        237 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 CGAATGAATCAGTTAAAAATGGTTTTGAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACA 663

Qy        257 ValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeu 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 GTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTA 723

Qy        277 ProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 CCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACA 783

Qy        297 TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys-LysLeuArgPheAs 316
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        784 TGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAAACTTCGTTTTGA 843

Qy        316 pArgIlePhePheArgAlaAlaAlaGlu-GluGlyHisIleIleProArgSerLeu-Asp 335
              |||||||||||||||||||||||||||| |||||||||||||||   ||| ::::||| |||
Db        844 TCGAATATTTTTCAGAGCAGCAGCAGAAAGAGGGACACATTATTTCCCGGAATTTGGGAC 903

Qy        336 LeuLeuGlyLeuGlu-LysLeuAspCysGly--ArgPheProSerAspHisTrpGly 353
              |||||||||||||||  |||||||||||||  ||||||  ||||||||||||||||||
Db        904 CCTTCTGGATTAGAAAAAACTGGACTGTGGGTAGAATTTCCTAAGGATCACTGGGGG 960

RESULT 12
BX337141
LOCUS       BX337141               1018 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX337141 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI036YJ02 5-PRIME, mRNA sequence.
ACCESSION   BX337141
VERSION     BX337141.2  GI:46271144
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1018)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30337585.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI036DE01QP1&c=3474.r.
FEATURES              Location/Qualifiers
     source           1. .1018
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DI036YJ02"
                      /tissue_type="PLACENTA COT 25-NORMALIZED"
                      /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:                1.8e-170      Length:          1018
Score:                    1545.00       Matches:         297
Percent Similarity:       94.9%         Conservative:    2
Best Local Similarity:    94.3%         Mismatches:      16
Query Match:              81.2%         Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2 (1-362) x BX337141 (1-1018)

Qy          5 SerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluProGluValLysLys 24
              |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AGTTGCCTGGMGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAG 62

Qy         25 ArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaValAlaGln 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 CGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAG 122

Qy         45 CysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluPro 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCT 182

Qy         65 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTAT 242

Qy         85 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 302

Qy        105 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 362

Qy        125 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGC 422

Qy        145 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGA 482

Qy        165 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTG 542

Qy        185 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 AAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATG 602

Qy        205 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 ATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACA 662

Qy        225 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 244
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TCCCATTTGGAGAGCAMCAGAGGGCATGCTGCGGAAMGAATGAATCAGTTAAAAATGGTT 722
```

```
Qy        245 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 264
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAAT 782

Qy        265 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 CTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTAMAMAAAAAMATTGTGGATGTATGG 842

Qy        285 GluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304
              |||  |||  ||||||  ||||||||||||  |||||||||  |||
Db        843 GAAGTTTTTGGGMAAMCTAAACATTGMMAGTATACATGGGATAMACAAATGAACTMTAAT 902

Qy        305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe 319
              |||||||||  |||  |||  :::||||||||||||||||||
Db        903 CTTGGRATAAMTGCTGMTTGTKAAATTCGTTTTGATMGAATATTT 947

RESULT 13
BX422491
LOCUS       BX422491                904 bp    mRNA    linear   EST 03-MAY-2004
DEFINITION  BX422491 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
            CS0DM007YE11 5-PRIME, mRNA sequence.
ACCESSION   BX422491
VERSION     BX422491.2  GI:46955237
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30766188.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DM007AC06QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .904
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DM007YE11"
                     /tissue_type="FETAL LIVER"
                     /dev_stage="fetal".
                     /clone_lib="Homo sapiens FETAL LIVER"
                     /note="Organ: liver; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

Alignment Scores:
Pred. No.:              5.78e-168       Length:         904
Score:                  1523.00         Matches:        292
Percent Similarity:     97.7%           Conservative:   2
Best Local Similarity:  97.0%           Mismatches:     7
Query Match:            80.1%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2 (1-362) x BX422491 (1-904)
```

```
Qy         20 ProGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAsp 39
              |||   :::|||||||||||||||||||||||||||||||||||||||||||||  |||
Db          1 CCCGGGATGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCYGCGAT 60

Qy         40 AlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsn 59
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAAC 120

Qy         60 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCT 180

Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCCCAARACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAA 240

Qy        100 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 119
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        241 ATCAGCCCATCTGAAGATACTCAGCAARAAAAWGGCAGCATGTTCTCTCTCATTACCTGG 300

Qy        120 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 139
              ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        301 AATATTGATGGATTAGATCTAAACAAWCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTAC 360

Qy        140 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSer 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGC 420

Qy        160 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTC 480

Qy        180 ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTT 540

Qy        200 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAG 600

Qy        220 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAAT 660

Qy        240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 259
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db        661 CAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTT 720

Qy        260 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsn 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCARATGTGGTGGTTTACCCAACAAC 780

Qy        280 IleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThr 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACA 840

Qy        300 GlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db        841 CAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAAWATTT 900

Qy        320 Phe 320
              |||
Db        901 TTC 903

RESULT 14
BU179107
LOCUS       BU179107                883 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7984768 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6174958
            5', mRNA sequence.
ACCESSION   BU179107
VERSION     BU179107.1  GI:22693091
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13548  row: o  column: 23
            High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6174958"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

Alignment Scores:
Pred. No.:                6.41e-165      Length:          883
Score:                    1497.00        Matches:         290
Percent Similarity:       99.0%          Conservative:    2
Best Local Similarity:    98.3%          Mismatches:      1
Query Match:              78.7%          Indels:          2
DB:                       5              Gaps:            0

US-10-757-745-2 (1-362) x BU179107 (1-883)

Qy         49 GluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGlu 68
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 CAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAG 62

Qy         69 SerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThr 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACC 122

Qy         89 AsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGln 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAA 182

Qy        109 GluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsn 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAAT 242

Qy        129 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATA 302

Qy        149 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTAT 362

Qy        169 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 422

Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        423 GTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTT 482

Qy        209 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAG 542

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 AGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATG 602

Qy        249 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGA 662

Qy        269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 722

Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 782

Qy        309 AlaAlaCysLysLeuArg-PheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHi 328
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        783 GCTGCTTGTAAACTTCGTTTTTGATCGAATATTTTTCCGAGCAGCAGCAGAAGAGGGACA 842

Qy        328 sIleIle-ProArgSerLeuAspLeuLeuGlyLeuGluLys 341
              |   ||| ||||||||||||||||||||||||:::|||
Db        843 CTTTATTCCCCCGAAGTTTGGACCTTCTTGGATTAAAAAAA 883

RESULT 15
BX433489/c
LOCUS       BX433489                952 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX433489 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN005YN02 3-PRIME, mRNA sequence.
ACCESSION   BX433489
VERSION     BX433489.2  GI:47002503
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 952)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30775195.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAI056ZD12_CS05372_1&c=3474.r

FEATURES             Location/Qualifiers
     source          1. .952
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
```

cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Alignment Scores:

Pred. No.:	3.22e-163	Length:	952
Score:	1483.00	Matches:	289
Percent Similarity:	97.7%	Conservative:	2
Best Local Similarity:	97.0%	Mismatches:	6
Query Match:	78.0%	Indels:	2
DB:	5	Gaps:	0

US-10-757-745-2 (1-362) x BX433489 (1-952)

```
Qy         66 ValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrVal 85
                ||||||||||||||||||:::      |||||||  ||||||||||||||||||||||
Db        948 GTGGAGGAGAGCGCCTTGGACGCCGAACCTGAACCCATCTCTGAGCCCAAGACCTATGT- 890

Qy         86 AspLeuThrAsnGluGluThrThr-AspSerThrThrSerLysIleSerProSerGluAs 105
                |||||||||||||||||||||||    |||||||  ||||||||||||||||||||||
Db        889 GACCTAACCAATGAAGAAACAACTGNATTCCACACTTTCTAAAATCAGCCCATCTGAAGA 830

Qy        105 pThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAs 125
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 TACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGA 770

Qy        125 pLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPr 145
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 TCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCC 710

Qy        145 oAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSe 165
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 AGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATC 650

Qy        165 rSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLy 185
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 AAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAA 590

Qy        185 sLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMe 205
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 GAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGAT 530

Qy        205 tArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSe 225
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 GAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATC 470

Qy        225 rHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLe 245
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 CCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTT 410

Qy        245 uLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLe 265
                ||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db        409 AAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCT 350

Qy        265 uArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGl 285
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 AAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGA 290

Qy        285 uPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLe 305
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 GTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCT 230

Qy        305 uGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGl 325
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 TGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGA 170

Qy        325 uGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGl 345
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 AGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGG 110

Qy        345 yArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 TAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 58
```

RESULT 16
BM926092

LOCUS	BM926092 1046 bp mRNA linear EST 12-MAR-2002
DEFINITION	AGENCOURT_6649780 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764440 5', mRNA sequence.
ACCESSION	BM926092
VERSION	BM926092.1 GI:19376471
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1046)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12817 row: o column: 01 High quality sequence start: 4 High quality sequence stop: 632.

FEATURES	Location/Qualifiers
source	1. .1046 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5764440" /lab_host="DH10B" /clone_lib="NIH_MGC_114" /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	2.5e-162	Length:	1046
Score:	1476.00	Matches:	287
Percent Similarity:	96.7%	Conservative:	4
Best Local Similarity:	95.3%	Mismatches:	8
Query Match:	77.6%	Indels:	2
DB:	3	Gaps:	0

US-10-757-745-2 (1-362) x BM926092 (1-1046)

```
Qy         55 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 74
                :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 95

Qy         75 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 94
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 CCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 155

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 TCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTC 215

Qy        115 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        216 TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 275

Qy        135 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 154

Db        276 GGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATT 335

Qy        155 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 174

Db        336 CCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCAT 395

Qy        175 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194

Db        396 GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAA 455

Qy        195 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 214

Db        456 GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAT 515

Qy        215 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 234

Db        516 GTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCT 575

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254

Db        576 GCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCA 635

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274

Db        636 GCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGT 695

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294

Db        696 GGTTTACCCAACAACATTGTGGATGTCTGGGGAGTTTTGGGCAAACCTAAACATTGCCAG 755

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314

Db        756 TATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGT 815

Qy        315 PheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeu 334

Db        816 TTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGGTTG 875

Qy        335 Asp-LeuLeuGlyLeuGluLysLeuAspCys-GlyArgPheProSerAspHisTrpGly 353

Db        876 GACCCTTCTTGGATAANAAAAACTGGACTGGGGGAAAATTTCCTAATGGATCCTGGGGG 934
```

```
RESULT 17
BX358707
LOCUS       BX358707                942 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX358707 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI042YL19 5-PRIME, mRNA sequence.
ACCESSION   BX358707
VERSION     BX358707.2  GI:46306560
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 942)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 5, 2003 this sequence version replaced gi:30378261.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI042CF10QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .942
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI042YL19"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

```
Alignment Scores:
Pred. No.:                1.85e-161       Length:          942
Score:                    1468.00         Matches:         289
Percent Similarity:       96.7%           Conservative:    1
Best Local Similarity:    96.3%           Mismatches:      9
Query Match:              77.2%           Indels:          3
DB:                       5               Gaps:            1

US-10-757-745-2 (1-362) x BX358707 (1-942)

Qy         55 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 74

Db         32 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 91

Qy         75 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 94

Db         92 CCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 151

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114

Db        152 TCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTC 211

Qy        115 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134

Db        212 TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 271

Qy        135 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 154

Db        272 GGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATT 331

Qy        155 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 174

Db        332 CCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCAT 391

Qy        175 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194

Db        392 GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAA 451

Qy        195 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 214

Db        452 GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAT 511

Qy        215 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 234

Db        512 GTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCT 571

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254

Db        572 GCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCA 631

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274

Db        632 GCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGT 691

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294
```

```
Db        692 GGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAG 751

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        752 TATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGT 811

Qy        315 PheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeu 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        812 TTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 871

Qy        335 AspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeu 354
              |||   ||||||   ||||||      ||||||                ||||||||| 
Db        872 GACHT-CTTGGAWTAGAAAAAWG-GGMTGTGGAAATTYC---YARTGATMATGGGGTCTC 926

RESULT 18
CN803230
LOCUS       CN803230                965 bp    mRNA    linear   EST 26-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_32551 Katze_MMBR Macaca mulatta cDNA clone
            IBIUW:12605 5' similar to Bases 1 to 965 highly similar to human
            TTRAP (Hs.210628), mRNA sequence.
ACCESSION   CN803230
VERSION     CN803230.1  GI:47699206
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 965)
  AUTHORS   Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
  TITLE     Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
  JOURNAL   Genome Biol. 6 (7), R60 (2005)
   PUBMED   15998449
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.04.02. 780 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 965   Std Error: 0.00
            Plate: CL000288  row: D  column: 11
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .965
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:12605"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

Alignment Scores:
Pred. No.:              1.68e-160       Length:         965
Score:                  1460.00         Matches:        275
Percent Similarity:     98.9%           Conservative:   6
Best Local Similarity:  96.8%           Mismatches:     3
Query Match:            76.8%           Indels:         0

DB:                     7               Gaps:           0

US-10-757-745-2 (1-362) x CN803230 (1-965)

Qy         79 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 TCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCT 61

Qy         99 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 118
              ||||||||||||||||||||||||||||||||||||    |||||||||      ||||||
Db         62 AAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGACAGCATGTTCTGCTTCATTACC 121

Qy        119 TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCC 181

Qy        139 TyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TACTTAGCGTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTAT 241

Qy        159 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178
              ||||||||||||||||||:::|||:::||||||||||||||||||||||||||||||||
Db        242 AGCTACCTAAAGAAGAGAGCAAGTGATTATGAGATTATTACAGGTCATGAAGAAGGATAT 301

Qy        179 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCT 361

Qy        199 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAAT 421

Qy        219 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATG 481

Qy        239 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 AATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATA 541

Qy        259 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 278
              ||||||||||||||||||||||||||:::||||||||:::|||||||||||||||||||
Db        542 TTTGCAGGAGATACAAATCTAAGGGATCAAGAGGTTACCAAATGTGGTGGTTTACCCAAC 601

Qy        279 AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAsp 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGAT 661

Qy        299 ThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 ACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAGCTTCGTTTTGATCGAATA 721

Qy        319 PhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGly 338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 TTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGA 781

Qy        339 LeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 TTGGAAAAACTCGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAATTTA 841

Qy        359 AspIleIleLeu 362
              :::::||||||
Db        842 AATGTAATATTG 853

RESULT 19
BM553049
LOCUS       BM553049                1192 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6542413 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742798
            5', mRNA sequence.
ACCESSION   BM553049
VERSION     BM553049.1  GI:18791437
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1192)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12761  row: i  column: 07
            High quality sequence stop: 747.
FEATURES             Location/Qualifiers
     source          1. .1192
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5742798"
                     /tissue_type="medulla"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_119"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source normal medulla from
                     anonymous male age 27.  Library  is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed  upon
                     cloning).  Average insert size 1.3 kb, insert size range
                     0.9-3 kb.  Library is normalized and enriched for
                     full-length clones and  was constructed by C. Gruber
                     (Invitrogen).  Research Genetics  tracking code 013. Note:
                     this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                2.07e-159      Length:          1192
Score:                    1452.00        Matches:         304
Percent Similarity:       85.8%          Conservative:    9
Best Local Similarity:    83.3%          Mismatches:      38
Query Match:              76.3%          Indels:          15
DB:                       3              Gaps:            5

US-10-757-745-2 (1-362) x BM553049 (1-1192)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 117

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 177

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 237

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 297

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 357

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 417

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 477

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 537

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 597

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 657

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 717

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 777

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        778 TTAAAAATGGTTTTTAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 837

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThr-ArgCysGlyGlyLeuProAsnAsnIl 280
              |||||||||||||||||||||||||||||||| ||||||||||||||||||      ::
Db        838 GGAGATACAAATCTAAGGGATCGAGAGGTTACCACGATGTGGTGGTTTACCCACAACATT 897

Qy        280 eValAspVal-Trp---GluPheLeuGlyLysProLysHisCys--GlnTyrThrTrpAs 298
              :   |||||| |||       |||   |||   ||||||||||  :::|||||||||
Db        898 GTGGGATGTCCTGGGGAGTTTTTGGGGGCAACCCTAAACATTGCCCAGGTATACATGGGG 957

Qy        298 pThrGlnMetAsnSerAsnLeuGly---------IleThrAlaAlaCysLysLeuArgPh 315
                          ::::::              :::   |||                ||
Db        958 ATACCACAAATGAAACTCCTATTCCTGGGAAAAACTGCCTGCTTGGAAAACCTTCCCTTT 1017

Qy        315 eAspArgIlePhePheArgAlaAlaAla------GluGluGlyHis------IleIlePr 331
              ||||      ||||||      |||         :::|||||:::      |||   ||
Db       1018 TGATTCGAAATTTTTTTCCCAAGCAACCCACCCCAAAGAAGGGAACCCCTTTATTTTCCC 1077

Qy        331 oArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHi 351
              |      :::||||||              |||         :::             |||
Db       1078 CCCGAAG------GTTTTGGGACCCCTTCCTTGTGATTTAAAAAAAAA-ACGTGGGACCG 1130

Qy        351 sTrpGly 353
               |||||||
Db       1131 GTGGGGG 1137

RESULT 20
CN641671
LOCUS       CN641671                1090 bp    mRNA    linear   EST 12-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_5159 Katze_MMBR Macaca mulatta cDNA clone IBIUW:5970
            5' similar to Bases 11 to 990 highly similar to human TTRAP
            (Hs.210628), mRNA sequence.
ACCESSION   CN641671
VERSION     CN641671.1  GI:47152681
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 1090)
  AUTHORS   Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
  TITLE     Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
  JOURNAL   Genome Biol. 6 (7), R60 (2005)
```

```
  PUBMED   15998449
COMMENT    Contact: C. Magness
           Illumigen Biosciences Inc.
           2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
           Tel: 2063780400
           Fax: 2063780408
           Email: cmagness@illumigen.com
           Sequenced on 2003.12.04. 732 Q20 bases.
           PCR PRimers
           FORWARD: CCCTCACTAAAGGGAACAAAA
           BACKWARD: CACTATAGGGCGAATTGGGTA
           Insert Length: 1090   Std Error: 0.00
           Plate: CL000043  row: E  column: 07
           Seq primer: CCCTCACTAAAGGGAACAAAA
           POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .1090
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:5970"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

Alignment Scores:
Pred. No.:              6.94e-159      Length:          1090
Score:                  1447.00        Matches:         286
Percent Similarity:     93.0%          Conservative:    6
Best Local Similarity:  91.1%          Mismatches:      17
Query Match:            76.1%          Indels:          5
DB:                     7              Gaps:            2

US-10-757-745-2 (1-362) x CN641671 (1-1090)

Qy          47 LeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProVal 66
               ||||||||||||||||||||||||||||||      ||||||||||||||||||||||||
Db          11 CTGGCCGAGAACGACTGGGAGATGGAAAGCCTCTCAAACTCCTACTTCGAGCCTCCGGTG 70

Qy          67 GluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAsp 86
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          71 GAGGAGAGCGCCTTGGAACGCCGCCCTGAAACCATCTCTGAGCCCAAGACCTATGTTGAC 130

Qy          87 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 106
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         131 CTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACT 190

Qy         107 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 126
               ||||||||||||   |||||||||   ||||||||||||||||||||||||||||||||
Db         191 CAGCAAGAAAATGACAGCATGTTCTGCTTCATTACCTGGAATATTGATGGATTAGATCTA 250

Qy         127 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAsp 146
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         251 AACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCGTTGTACAGCCCAGAT 310

Qy         147 ValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSer 166
               ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db         311 GTGATATTTCTACAGGAAGTTATTCCCCCTTATTATAGCTACCTAAAGAAGAGAGCAAGT 370

Qy         167 AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186
               :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         371 GATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAA 430

Qy         187 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 206
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         431 TCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGA 490

Qy         207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 226
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         491 AACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCAT 550

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         551 TTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAG 610

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         611 AAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGG 670

Qy         267 AspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPhe 286
               |||:::||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db         671 GATCAAGAGGTTACCAAATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTT 730

Qy         287 LeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGly 306
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         731 TTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGA 790

Qy         307 IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 326
               ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         791 ATAACTGCTGCTTGTAAGCTTCCTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAA 850

Qy         327 -GlyHisIleIleProArgSerLeu---AspLeuGluGlyLeuGluLysLeuAspCys-- 344
                ||||||||||||||||||||||||   ||||||||||||||||
Db         851 GGGACACATTATTCCCCCGAAGTTTGGAACCTTCCTTGGATTGGAAAAAACTCCGACTGG 910

Qy         345 -GlyArgPhePro--SerAspHisTrpGlyLeuLeu 355
                |||:::|||||||      ||||||   |||:::|||
Db         911 GGGAAAATTTCCCTAAGGGATCACCTGGGGGTCCTT 946

RESULT 21
AL555333
LOCUS       AL555333                 948 bp    mRNA    linear   EST 30-MAR-2004
DEFINITION  AL555333 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK007YK05 5-PRIME, mRNA sequence.
ACCESSION   AL555333
VERSION     AL555333.3  GI:45860070
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:31277141.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DK007AF03QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK007YK05"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
```

primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:

Pred. No.:	2.15e-157	Length:	948
Score:	1433.50	Matches:	291
Percent Similarity:	94.5%	Conservative:	0
Best Local Similarity:	94.5%	Mismatches:	15
Query Match:	75.4%	Indels:	4
DB:	1	Gaps:	1

US-10-757-745-2 (1-362) x AL555333 (1-948)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db         19 ATGGAGTTGGGGAGTTGCCTGGMGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 78

Qy         21 Glu-ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAl 40
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GAGTGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGC 138

Qy         40 aAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSe 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTC 198

Qy         60 rTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGl 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 CTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGA 258

Qy         80 uProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIl 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 GCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAAT 318

Qy        100 eSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAs 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAA 378

Qy        120 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTT 438

Qy        140 uAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTy 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 AGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTA 498

Qy        160 rLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheTh 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 CCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCAC 558

Qy        180 rAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 AGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCC 618

Qy        200 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 AAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCT 678

Qy        220 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCA 738

Qy        240 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 GTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGC 798

Qy        260 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 280
              ||||||||||   |||  ||||||||||||||||||||||||||||||||||      ||
Db        799 AGGAGATACAWATCTARGGGATCGAGAGGTTACCAGATGTGGTGGKTTACCAACAAM-AT 857

Qy        280 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGl 300
              |||||||   ||||||||||||||||         ||||||
```

```
Db        858 TGTKGATKTCTGGGAGTTTTTGGGCAAMCTTAA-CATTGCAGTATAAWB---GGATMAMA 913

Qy        300 nMetAsnSerAsnLeuGlyIle 307
                  |||||||||||||||||||
Db        914 AATGAACTCTAATCTYGGAATA 935
```

RESULT 22
BM555041

LOCUS	BM555041 1103 bp mRNA linear EST 20-FEB-2002
DEFINITION	AGENCOURT_6545705 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5737150 5', mRNA sequence.
ACCESSION	BM555041
VERSION	BM555041.1 GI:18795166
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1103)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12746 row: m column: 23 High quality sequence stop: 678.

FEATURES Location/Qualifiers

```
     source          1. .1103
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5737150"
                     /tissue_type="duodenal adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_88"
                     /note="Organ: small intestine; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                     oligo-dT primed. Average insert size 1.767 kb. Library
                     enriched for full-length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."
```

ORIGIN

Alignment Scores:

Pred. No.:	7.13e-155	Length:	1103
Score:	1413.00	Matches:	289
Percent Similarity:	93.0%	Conservative:	2
Best Local Similarity:	92.3%	Mismatches:	14
Query Match:	74.3%	Indels:	8
DB:	3	Gaps:	3

US-10-757-745-2 (1-362) x BM555041 (1-1103)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 89

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 149

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 209

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
```

```
Db        210 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 269

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        270 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 329

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        330 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 389

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        390 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 449

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        450 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 509

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db        510 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 569

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db        570 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 629

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220

Db        630 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 689

Qy        221 CysLeuMetThrSer-HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 240

Db        690 TGCCTTATGACATCCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCA 749

Qy        240 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 260

Db        750 GNTAAAAATGGTTTTAAAGAAAATGCAAGAAGCTCCAGAGTCACCTTACGTTATATTTGC 809

Qy        260 aGlyAspThrAsnLeuArg-AspArgGluValThrArg-CysGly-GlyLeuProAsnAs 279

Db        810 AGGAGATACAAATCTAAGGGGATCGAGAGGGTACCCGAATGTGGTGGGTTTACCCACAAC 869

Qy        279 nIleValAspValTrp------GluPheLeuGlyLysProLysHisCys---GlnTyrTh 296

Db        870 CATTGTGGAA---TGGCTGGGGAATTTTTTGGGCAAACCTAAACTTTGCCCAGAATACAT 926

Qy        296 rTrpAspThrGlnMetAsnSerAsnLeuGly 306

Db        927 GGGGAAACCCCAATGGAACCCTAATCTTGGG 957
```

RESULT 23
CR765451

LOCUS	CR765451 846 bp mRNA linear EST 23-SEP-2004
DEFINITION	DKFZp469E2434_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone DKFZp469E2434 5', mRNA sequence.
ACCESSION	CR765451
VERSION	CR765451.1 GI:52605526
KEYWORDS	EST.
SOURCE	Pongo pygmaeus (orangutan)
ORGANISM	Pongo pygmaeus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE	1 (bases 1 to 846)
AUTHORS	Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
TITLE	Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., et al.)
JOURNAL	Unpublished (2004)
COMMENT	Contact: MIPS MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany This is the 5' sequence of the clone insert. Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp469E2434) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469E2434 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES Location/Qualifiers

```
     source          1. .846
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp469E2434"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="469 (synonym: pkid1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
```

ORIGIN

Alignment Scores:			
Pred. No.:	2.38e-152	Length:	846
Score:	1390.00	Matches:	272
Percent Similarity:	97.8%	Conservative:	1
Best Local Similarity:	97.5%	Mismatches:	4
Query Match:	73.1%	Indels:	2
DB:	7	Gaps:	0

US-10-757-745-2 (1-362) x CR765451 (1-846)

```
Qy         34 SerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGlu 53

Db          3 TCGGTCGCGAGCTGCGATACCGCTGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAG 62

Qy         54 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 73

Db         63 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 122

Qy         74 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 93

Db        123 CGCCCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 182

Qy         94 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 113

Db        183 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 242

Qy        114 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 133

Db        243 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 302

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153

Db        303 CGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTT 362

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173

Db        363 ATTCCCCCATATTATAGCTACCTAAAGAAGAGTTCAAGTAATTATGAGATTATTACAGGT 422

Qy        174 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 193

Db        423 CATGAAGAAGGGTATTTCACAGCTATCATGTTGAAGAAATCAAGAGTGAAATTAAAAAGC 482

Qy        194 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 213

Db        483 CAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 542

Qy        214 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 233

Db        543 AGTGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 602
```

```
Qy        234 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAG 662

Qy        254 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGT 722

Qy        274 GlyGlyLeuProAsnAsnIleValAspValTrp-GluPheLeuGlyLysProLysHisCy 293
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        723 GGTGGTTTACCCAACAACATCGTGGATGTCTGGGGAGTTTTTGGGCAAACCTAAACATTG 782

Qy        293 s-GlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 310
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 GCCAGTNTACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCT 835
```

RESULT 24
BG740396

```
LOCUS       BG740396                870 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602634171F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4779449 5',
            mRNA sequence.
ACCESSION   BG740396
VERSION     BG740396.1  GI:14051049
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 870)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10636  row: m  column: 18
            High quality sequence stop: 826.
FEATURES             Location/Qualifiers
     source          1. .870
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4779449"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn3"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
```

Alignment Scores:

Pred. No.:	3.26e-152	Length:	870
Score:	1389.00	Matches:	281
Percent Similarity:	96.9%	Conservative:	2
Best Local Similarity:	96.2%	Mismatches:	4
Query Match:	73.0%	Indels:	6
DB:	2	Gaps:	0

US-10-757-745-2 (1-362) x BG740396 (1-870)

```
Qy         31 GluPheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsn 50
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAAC 61

Qy         51 AspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAla 70
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GACTGGGAGATGGAAAGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCC 121

Qy         71 LeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGlu 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAA 181

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAAT 241

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCA 301

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTA 361

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATT 421

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAA 481

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGT 541

Qy        211 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACC 601

Qy        231 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAG 661

Qy        251 AlaProGluSerAlaThrValIle-PheAlaGlyAspThrAsnLeuArgAspArgGluVa 270
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        662 GCTCCAGAGTCAGCTACAGTTATAATTCGCAGGAGATACAAATCTAAGGGATCGAGAGGT 721

Qy        270 lThrArgCysGlyGlyLeu--ProAsnAsnIleValAspValTrpGluPheLeuGlyLys 289
              |||||||||||||||||||  |||||||||||||||||||||||||   ||||||||||||
Db        722 TACCAGATGTGGTGGTTTACCCCAAACAACATTGTGGATGTCTGGGA-TTTTTGGGCAAA 780

Qy        290 ProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn-LeuGlyIleThrAl 309
              |||||||||||||||||||||||||||||::::::|||||||||| ||||||||||||||
Db        781 CCTAAACATTGCCAGTATACCTGGGATACCAAATTGAACTCTAATCCTTGGAATAACTGC 840

Qy        309 aAla-CysLysLeuArgPheAspArg 317
              |||| |||         |||||||||
Db        841 TGCTTTGTTAACCTTCGTTTGATCGG 866
```

RESULT 25
CN642527

```
LOCUS       CN642527                1087 bp    mRNA    linear   EST 12-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_6530 Katze_MMPL2 Macaca mulatta cDNA clone IBIUW:4699
            5' similar to Bases 1 to 1084 highly similar to human TTRAP
            (Hs.210628), mRNA sequence.
ACCESSION   CN642527
VERSION     CN642527.1  GI:47153537
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 1087)
  AUTHORS   Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
```

```
          Iadonato,S.P.
  TITLE   Analysis of the Macaca mulatta transcriptome and the sequence
          divergence between Macaca and human
  JOURNAL Genome Biol. 6 (7), R60 (2005)
  PUBMED  15998449
COMMENT   Contact: C. Magness
          Illumigen Biosciences Inc.
          2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
          Tel: 2063780400
          Fax: 2063780408
          Email: cmagness@illumigen.com
          Sequenced on 2003.12.24. 840 Q20 bases.
          PCR PRimers
          FORWARD: CCCTCACTAAAGGGAACAAAA
          BACKWARD: CACTATAGGGCGAATTGGGTA
          Insert Length: 1087   Std Error: 0.00
          Plate: CL000066  row: D  column: 10
          Seq primer: CCCTCACTAAAGGGAACAAAA
          POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .1087
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:4699"
                     /sex="male"
                     /dev_stage="newborn infant"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMPL2"
                     /note="Organ: placenta; Vector: Uni-ZAP XR; Site_1: EcoR
                     I; Site_2: Xho I; Created from Stratagene ZAP-cDNA
                     Synthesis kit (catalog #200400)  and ZAP-CDNA Gigapack III
                     Gold Cloning Kit (Catalog #200450)"
ORIGIN

Alignment Scores:
Pred. No.:                4.08e-151      Length:          1087
Score:                    1381.00        Matches:         260
Percent Similarity:       98.9%          Conservative:    5
Best Local Similarity:    97.0%          Mismatches:      3
Query Match:              72.6%          Indels:          0
DB:                       7              Gaps:            0

US-10-757-745-2 (1-362) x CN642527 (1-1087)

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114
              ||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db          2 TCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGACAGCATGTTC 61

Qy        115 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGCTTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 121

Qy        135 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGGGTGTGTTCCTACTTAGCGTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATT 181

Qy        155 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 174
              ||||||||||||||||||||||||||||||:::|||:::||||||||||||||||||||
Db        182 CCCCCATATTATAGCTACCTAAAGAAGAGAGCAAGTGATTATGAGATTATTACAGGTCAT 241

Qy        175 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAA 301

Qy        195 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAT 361

Qy        215 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCT 421

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCA 481

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
              ||||||||||||||||||||||||||||||||||||:::|||||||||:::||||||
Db        482 GCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCAAGAGGTTACCAAATGTGGT 541

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAG 601

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 TATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAGCTTCGT 661

Qy        315 PheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeu 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 721

Qy        335 AspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHisTrpGlyLeu 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GACCTTCTTGGATTGGAAAAACTCGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTT 781

Qy        355 LeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||:::||||||
Db        782 CTGTGCAATTTAGATGTAATATTG 805

RESULT 26
CX166335
LOCUS       CX166335                 810 bp    mRNA    linear   EST 23-DEC-2004
DEFINITION  HESC2_39_B09.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone
            IMAGE:7469923 5', mRNA sequence.
ACCESSION   CX166335
VERSION     CX166335.1  GI:56796415
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 810)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15771  row: d  column: 17
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 810.
FEATURES             Location/Qualifiers
     source          1. .810
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7469923"
                     /sex="male"
                     /tissue_type="human embryonic stem cells differentiated to
                     an early endodermal cell type"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
```

/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_258"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early endodermal cell type. Cell line id and NIH Registry designation is BGO1. Positive for GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP expression. Passage number 40. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Note: this is a Mammalian Gene Collection library."

ORIGIN

Alignment Scores:

Pred. No.:	1.13e-149	Length:	810
Score:	1367.00	Matches:	268
Percent Similarity:	99.6%	Conservative:	0
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	71.9%	Indels:	1
DB:	8	Gaps:	0

US-10-757-745-2 (1-362) x CX166335 (1-810)

```
Qy         16 GluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGluPheAlaSerVal 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTC 62

Qy         36 AlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 GCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAA 122

Qy         56 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCT 182

Qy         76 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCC 242

Qy         96 ThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 ACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCT 302

Qy        116 LeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGly 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGG 362

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 422

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 482

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 542

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 602

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 662
```

```
Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 722

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 782

Qy        276 LeuProAsnAsnIleValAspValTrp 284
              ||||||   ||||||||||||||||||
Db        783 TTACCCAC-AACATTGTGGATGTCTGG 808
```

RESULT 27
CN801806
LOCUS CN801806 922 bp mRNA linear EST 26-MAY-2004
DEFINITION ILLUMIGEN_MCQ_37066 Katze_MMBR Macaca mulatta cDNA clone IBIUW:15450 5' similar to Bases 18 to 922 highly similar to human TTRAP (Hs.210628), mRNA sequence.
ACCESSION CN801806
VERSION CN801806.1 GI:47697782
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 922)
AUTHORS Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B., Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and Iadonato,S.P.
TITLE Analysis of the Macaca mulatta transcriptome and the sequence divergence between Macaca and human
JOURNAL Genome Biol. 6 (7), R60 (2005)
PUBMED 15998449
COMMENT Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2004.05.13. 514 Q20 bases.
PCR PRimers
FORWARD: CCCTCACTAAAGGGAACAAAA
BACKWARD: CACTATAGGGCGAATTGGGTA
Insert Length: 922 Std Error: 0.00
Plate: CL000291 row: H column: 09
Seq primer: CCCTCACTAAAGGGAACAAAA
POLYA=No.
FEATURES Location/Qualifiers
source 1. .922
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/clone="IBIUW:15450"
/sex="female"
/dev_stage="adult"
/lab_host="E. coli SOLR"
/clone_lib="Katze_MMBR"
/note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I; Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis kit (catalog #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit (Catalog #200450)"
ORIGIN

Alignment Scores:

Pred. No.:	1.83e-149	Length:	922
Score:	1366.00	Matches:	265
Percent Similarity:	92.2%	Conservative:	8
Best Local Similarity:	89.5%	Mismatches:	23
Query Match:	71.8%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2 (1-362) x CN801806 (1-922)

```
Qy          46 PheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProPro 65
               ||||||   |||   :::|||   |||   ||||||||   |||   ||||||||||||
Db          18 TTCCTGGACGAGAAGGAGTGGCACATGACCAGGGCTCTGATCTCCTGATTCGAGCCTCCG 77

Qy          66 ValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrVal 85
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          78 GTGGAGGAGAGCGCCTTGGAACGCCGCCCTGAAACCATCTCTGAGCCCAAGACCTATGTT 137

Qy          86 AspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 105
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         138 GACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGAT 197

Qy         106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
               ||||||||||||||   |||||||||      |||||||||||||||||||||||:::
Db         198 ACTCAGCAAGAAAATGACAGCATGTTCTGCTTCATTACCTGGAATATTGATGGATTAAAT 257

Qy         126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         258 CTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCGTTGTACAGCCCA 317

Qy         146 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 165
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db         318 GATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGAGCA 377

Qy         166 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 185
               |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         378 AGTGATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAG 437

Qy         186 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 205
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         438 AAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATG 497

Qy         206 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSer 225
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         498 AGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCC 557

Qy         226 HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeu 245
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         558 CATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTA 617

Qy         246 LysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu 265
               |||:::|||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         618 AAGAGAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACATATCTA 677

Qy         266 ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGlu 285
               ||||||:::|||||||||::||||||  ||||||||||||||||||||||||||||||
Db         678 AGGGATCAAGAGGTTACCAAATGTGGTCGTTTACCCAACAACATTGTGGATGTCTGGGAG 737

Qy         286 PheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeu 305
               ||||||   ||||||||||||   ||||||||||||||||   ||||||||||||||||
Db         738 TTTTTGAGCAAACCTAAACATTTGCAGTATACATGGGATACATCAATGAACTCTAATCTG 797

Qy         306 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 325
                  |||   ||||||||||||||||||||||||||||||||||:::|||     |||
Db         798 GAAATACCTGCTGCTTGTAAGCTTCGTTTTGATCGAATATTTTTCAAAGCACCAACAGAA 857

Qy         326 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 341
                  |||   |||   ||||||||||||      ||||||||||||||||
Db         858 CGAGGAACCATTTATCCCCGAAGCTTGTACTTTCTTGGATTGGAAAAA 905
```

RESULT 28
BU159911
LOCUS BU159911 870 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7933863 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6144208 5', mRNA sequence.
ACCESSION BU159911
VERSION BU159911.1 GI:22673821
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 870)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM13468 row: n column: 17
High quality sequence stop: 677.
FEATURES Location/Qualifiers
source 1. .870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6144208"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."
ORIGIN

Alignment Scores:

Pred. No.:	1.28e-146	Length:	870
Score:	1341.50	Matches:	268
Percent Similarity:	96.4%	Conservative:	1
Best Local Similarity:	96.1%	Mismatches:	6
Query Match:	70.5%	Indels:	4
DB:	5	Gaps:	1

US-10-757-745-2 (1-362) x BU159911 (1-870)

```
Qy          23 LysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaVal 42
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           2 AAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTG 61

Qy          43 AlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPhe 62
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 GCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTC 121

Qy          63 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 82
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 GAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAG 181

Qy          83 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 102
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         182 ACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCA 241

Qy         103 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 122
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 TCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGAT 301

Qy         123 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         302 GGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTG 361

Qy         143 TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 162
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         362 TACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAG 421

Qy         163 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 182
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        422 AAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATA 481

Qy        183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 ATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACC 541

Qy        203 LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 AAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTT 601

Qy        223 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAA 661

Qy        243 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 262
              |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        662 ATGGTTTTAAAGAAANATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGAT 721

Qy        263 ThrAsnLeuArgAspArgGluValThr-ArgCys-GlyGlyLeuProAsnAsnIleValA 282
              |||||||||||||||||||||||||| ||||||    ||||||||||||||||||||||
Db        722 ACAAATCTAAGGGATCGAGAGGTTACCCAGATGTGNGTGGTTTACCCAACAACATTGTGG 781

Qy        282 spValTrp-GluPheLeuGlyLysProLysHisCys---GlnTyrThrTrp 297
              ||||||| ||||||||||:::                |||||||||||
Db        782 ATGTCTGGGGAGTTTTTGGGGCAAACCCTANACATTGGCCAGTATACATGG 832
```

RESULT 29
BM554324

```
LOCUS       BM554324               1035 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6546794 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742466
            5', mRNA sequence.
ACCESSION   BM554324
VERSION     BM554324.1  GI:18793827
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1035)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12760  row: k  column: 11
            High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1. .1035
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5742466"
                     /tissue_type="medulla"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_119"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source normal medulla from
                     anonymous male age 27.  Library  is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed  upon
                     cloning).  Average insert size 1.3 kb, insert size range
                     0.9-3 kb.  Library is normalized and enriched for
                     full-length clones and  was constructed by C. Gruber
                     (Invitrogen).  Research Genetics  tracking code 013. Note:
                     this is a NIH_MGC Library."
ORIGIN
```

```
Alignment Scores:
Pred. No.:                4.37e-145      Length:          1035
Score:                    1329.50        Matches:         273
Percent Similarity:       96.2%          Conservative:    5
Best Local Similarity:    94.5%          Mismatches:      9
Query Match:              69.9%          Indels:          5
DB:                       3              Gaps:            0

US-10-757-745-2 (1-362) x BM554324 (1-1035)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGA-GAAGAGGGCGAGCCT 232

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 292

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 352

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 412

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 472

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 532

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 592

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 652

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 712

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCN 772

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 832

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||  ||||||||     |||||||||
Db        833 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGG-CATGCTGCGGGAACGATGAATCAG 891

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAla-ProGluSerAlaThrValIlePheAl 260
              |||||||||||||||||||||||||||||| ||||||||||||||   ||||||||
Db        892 TT-AAAATGGTTTTAAAGAAAATGCAAGAAGCTTCCAGAGTCAGCTACAGNTATATTTGC 950

Qy        260 aGlyAspThrAsnLeuArgAspArgGluValThrArgCys-GlyGlyLeuProAsnAsnI 280
              ||||||||||||||||:::|||||||||||||||:::||| |||||||||||      :
Db        951 AGGAGATACAAATCTAAAGGATCGAGAAGTTACCCAATGTTGGGGGTTTACCCACAACAT 1010

Qy        280 leValAspValTrpGluPheLeu 287
              ::   :::||||||:::||||||
Db       1011 TGTGGAATGTCTGGGATTTTTTG 1033
```

RESULT 30
CD251503

```
LOCUS       CD251503                906 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  AGENCOURT_14212105 NIH_MGC_179 Homo sapiens cDNA clone
            IMAGE:30385385 5', mRNA sequence.
ACCESSION   CD251503
VERSION     CD251503.1  GI:31011969
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM452  row: f  column: 18
            High quality sequence stop: 588.
FEATURES             Location/Qualifiers
     source          1. .906
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30385385"
                     /tissue_type="Pituitary"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_179"
                     /note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV
                     (destroyed); Site_2: NotI; Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning).  Average insert size 1.1 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN

Alignment Scores:
Pred. No.:              6.95e-145       Length:         906
Score:                  1327.00         Matches:        276
Percent Similarity:     94.9%           Conservative:   2
Best Local Similarity:  94.2%           Mismatches:     7
Query Match:            69.8%           Indels:         8
DB:                     6               Gaps:           0

US-10-757-745-2 (1-362) x CD251503 (1-906)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 101

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 161

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 221

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 281

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 341

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 401

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 461

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 521

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 581

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 641

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAAGAAATGAGCTT 701

Qy        221 CysLeuMetThrSerHis-LeuGluSerThrArgGlyHisAla-AlaGluArgMetAsnG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 TGCCTTATGACATCCCATTTTGGAGAGCACCAGAGGGCATGCTTGCGAAACGAATGAATC 761

Qy        240 lnLeuLysMetValLeuLysLysMetGln-GluAlaProGluSerAlaThrValIle-Ph 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 AGTTAAAAATGTTTTTTAAGAAAATGCAAGGAGCCTCCAGAGTTCGCTACCGTTATTTTT 821

Qy        259 eAlaGlyAspThrAsn-LeuArgAspArgGluValThrArgCysGly--GlyLeuProAs 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 TGCAGGAGATACAAATCCTAAGGGATCGAGAGGGTACCCGAATGGGGTGGGTTTACCCAA 881

Qy        278 nAsnIle-ValAspValTrpGlu 285
              ||||||| |||:::||||||||||
Db        882 CAACATTGGTGGAAGTCTGGGAA 904

RESULT 31
BI908925
LOCUS       BI908925                838 bp    mRNA    linear   EST 16-OCT-2001
DEFINITION  603067028F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5215926 5',
            mRNA sequence.
ACCESSION   BI908925
VERSION     BI908925.1  GI:16172029
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11542  row: h  column: 07
            High quality sequence stop: 773.
FEATURES             Location/Qualifiers
     source          1. .838
```

```
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5215926"
                     /tissue_type="leukocyte"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_118"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                     (destroyed); RNA source leukocytes from anonymous pool of
                     non-activated adult donors. Library is oligo-dT primed
                     and directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.7 kb, insert size range
                     1.2-3.3 kb. Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen). Research Genetics tracking code 027. Note:
                     this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.06e-141       Length:         838
Score:                  1299.50         Matches:        264
Percent Similarity:     95.0%           Conservative:   2
Best Local Similarity:  94.3%           Mismatches:     9
Query Match:            68.3%           Indels:         5
DB:                     3               Gaps:           1

US-10-757-745-2 (1-362) x BI908925 (1-838)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 61

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 121

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 181

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 241

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 301

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 361

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 421

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 481

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 541

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 601

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly-AsnGluLe 220
              |||||||||||||||||||||    ||||||||||||||||||||||| ||||||||
Db        602 AGTACCAAAATGATGAGAAACCTTTCATGTGTGCATGTGAATGTGTCAGGACAATGAGCT 661

Qy        220 uCysLeuMetThrSerHisLeu---GluSerThrArgGlyHisAlaAlaGluArgMetAs 239
              |||||||||||||||||||||   |||||||||||||||||||||||||||||||||||
Db        662 TTGCCTTATGACATCCCATTTTGGAGAGCACCAGAGTGGCATTGCTGCGGAACGAATGAA 721

Qy        239 nGlnLeuLysMetValLeu--LysLysMetGlnGluAlaProGluSerAlaThrValIle 258
              ||||||||||||||||||  :::   ||||||||||||||||||||||||||||||||||
Db        722 TCAGTTAAAAATGGTTTTACAAGACACATGCAAGAGGCTCCAGAGTCAGCTACAGTTATA 781

Qy        259 PheAlaGlyAspThrAsnLeuArgAsp-ArgGluValThrArgCysGlyGly 275
              |||||||||  |||||||||::: ||||||||||||||||||||||||||
Db        782 TCTGCAGGAGATCCAAATCTAAGGGAATCGAGAGGTTACCAGATGTGGTGGT 833

RESULT 32
BG740339
LOCUS       BG740339                 932 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602635289F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4780318 5',
            mRNA sequence.
ACCESSION   BG740339
VERSION     BG740339.1  GI:14050992
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 932)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10639  row: a  column: 23
            High quality sequence stop: 759.
FEATURES             Location/Qualifiers
     source          1. .932
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4780318"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn3"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.5kb. Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.26e-141       Length:         932
Score:                  1299.50         Matches:        269
Percent Similarity:     89.3%           Conservative:   5
Best Local Similarity:  87.6%           Mismatches:     19
Query Match:            68.3%           Indels:         14
DB:                     2               Gaps:           6

US-10-757-745-2 (1-362) x BG740339 (1-932)

Qy         58 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 77
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 CTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACC 62

Qy         78 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 97
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACT 122

Qy         98 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        123 TCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATT 182

Qy        118 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCys 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGT 242

Qy        138 SerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyr 157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 TCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATAT 302

Qy        158 TyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGly 177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGA 362

Qy        178 TyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATT 422

Qy        198 ProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly 217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGA 482

Qy        218 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGA 542

Qy        238 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 ATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTT 602

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 ATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCC 662

Qy        278 AsnAsnIleValAspValTrpGluPhe-LeuGlyLysProLysHisCysGln-TyrThrT 297
              |||||||||||||||||||||||||| |||||||||||||||||||||| |||||||
Db        663 AACAACATTGTGGATGTCTGGGAGTTCCTTGGGCAAACCTAAACATTGCCAGTTATACAT 722

Qy        297 rpAspThrGlnMetAsnSerAsn-LeuGlyIleThrAlaAlaCys---LysLeuArgPhe 315
              |||||||||||||||||||||| ||||||||||||||||||||   ||||||
Db        723 GGGATACACAAATGAACTCTAATCTTGGGAATAACTGCTGACTGGTAAAACTTTCGATTT 782

Qy        316 ---AspArgIlePhePheArg------AlaAlaAlaGluGluGlyHisIle---IlePro 331
                  |||||||||:::         |||:::::::::||||||         |||
Db        783 TGGATCGAAATATTTTTCAAGAAGCAGACAGCCAGCAAAGACGGACACCATTAATTCCCC 842

Qy        332 ArgSerLeu------------AspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 347
              |||   |||            |||         |||      :::      |||   |||
Db        843 CGAAAGTTGGGACCTTCTTGTGATTCAACAACACTGCACCTGGTG------GGAAATTTC 896

Qy        348 ProSerAspHisTrp 352
                  |||   |||
Db        897 GTTAGGGATCCATGG 911

RESULT 33
BX365835
LOCUS       BX365835                783 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX365835 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
            cDNA clone CS0DB007YB11 5-PRIME, mRNA sequence.
ACCESSION   BX365835
VERSION     BX365835.2  GI:46288859
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 783)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 5, 2003 this sequence version replaced gi:30370946.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS1DB002ZF03QP1&c=3474.r.
FEATURES              Location/Qualifiers
     source           1. .783
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DB007YB11"
                      /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
                      /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:                5.54e-141       Length:          783
Score:                    1293.00         Matches:         251
Percent Similarity:       98.8%           Conservative:    3
Best Local Similarity:    97.7%           Mismatches:      3
Query Match:              68.0%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2 (1-362) x BX365835 (1-783)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         13 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 72

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              |||||||||||||||||||||||   :::||||||   ||||||||:::|||||||||
Db         73 GAGGTGAAAAAGCGGCGACTTCTGTTTTTGGAGTTTCCCTCGGTCGCAACCTGCGATGCC 132

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 GCAGTGGCTCAGTGCTTCCTGGCCGARAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 192

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 252

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              |||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        253 CCCAAAACCTATKTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 312

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db        313 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATTTTCTCTCTCATTACCTGGAAT 372

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 432

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 492

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 552

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        553 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 612

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 672

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 732

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTT 783

RESULT 34
BX374579
LOCUS       BX374579                780 bp    mRNA    linear   EST 27-APR-2004
DEFINITION  BX374579 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
            cDNA clone CS0DB007YB11 5-PRIME, mRNA sequence.
ACCESSION   BX374579
VERSION     BX374579.2  GI:46618956
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 780)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30452317.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DB007CA06QP1&c=3474.r.
FEATURES              Location/Qualifiers
     source           1. .780
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DB007YB11"
                      /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
                      /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:                1.24e-140       Length:          780
Score:                    1290.00         Matches:         251
Percent Similarity:       98.0%           Conservative:    0
Best Local Similarity:    98.0%           Mismatches:      5
Query Match:              67.8%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2 (1-362) x BX374579 (1-780)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         13 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 72

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         73 GAGGTGAAAAARCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCRATGCC 132

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              |||   |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        133 GCAKTGGCTCAKTGCTTCCTGGCCGARAACGACTGGGARATGGAAAGGGCTCTGAACTCC 192

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 TACTTCGAGCCTCCGGTGGAGGARAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 252

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        253 CCCAARACCTATKTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 312

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 372

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        373 ATTGATGGATTARATCTAAACAATCTGTCAGARAGGGCTCGAGGGGTGTGTTCCTACTTA 432

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 492

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 552

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 612

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 672

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 732

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 256
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACA 780

RESULT 35
BI553412
LOCUS       BI553412                858 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603193358F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5264582 5',
            mRNA sequence.
ACCESSION   BI553412
VERSION     BI553412.1  GI:15440724
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 858)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
```

```
           http://image.llnl.gov
           Plate: LLAM11667  row: c  column: 15
           High quality sequence stop: 829.
FEATURES             Location/Qualifiers
     source          1. .858
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5264582"
                     /tissue_type="hippocampus"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_95"
                     /note="Organ: brain; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average
                     insert size 2.5 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation). Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this
                     is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                1.66e-138      Length:          858
Score:                    1272.50        Matches:         258
Percent Similarity:       93.2%          Conservative:    3
Best Local Similarity:    92.1%          Mismatches:      16
Query Match:              66.9%          Indels:          5
DB:                       3              Gaps:            2

US-10-757-745-2 (1-362) x BI553412 (1-858)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 80

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 140

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 200

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 260

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 320

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 380

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 440

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 500

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 560

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 620

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsn---ValSerGlyAsnGlu 219
              ||||||||||||||||||||||||||||||||||||||||||   ||||||   ||||||
Db        621 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACTGTGTCAGCGA-AATGAG 679

Qy        220 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 CTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAAT 739

Qy        240 GlnLeuLys------MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
              |||:::          :::||||||      ||||||||||||      :::
Db        740 CAGGTTAACAATGGTTTTCA-ACACGAAAAATGCCAAGAGCTCCAGAGTCAGTACAGTTA 798

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro 277
              ||||||||||||||||    ||||||      ||||||      ||||||||||||
Db        799 TATTTTGCGGGAGATACAATCCTAAGGATCGCAGAGGTTCCAGATGTGGGTGGTTTACCC 858

RESULT 36
BU169945
LOCUS       BU169945                 883 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7913097 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6024760
            5', mRNA sequence.
ACCESSION   BU169945
VERSION     BU169945.1  GI:22683929
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13235  row: m  column: 17
            High quality sequence stop: 413.
FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6024760"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

Alignment Scores:
Pred. No.:                1.17e-134      Length:          883
Score:                    1240.00        Matches:         248
Percent Similarity:       95.4%          Conservative:    0
Best Local Similarity:    95.4%          Mismatches:      5
Query Match:              65.2%          Indels:          7
DB:                       5              Gaps:            1

US-10-757-745-2 (1-362) x BU169945 (1-883)

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 60
```

```
Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 120

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 180

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 240

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 300

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 360

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 420

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 480

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 540

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 600

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 660

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCys-GlyGlyLeuProAsnAsnIl 280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGAGATACAAATCTAAGGGATCGAGAGGGTACCAGATGGGGGGGGTTTACCCAACAACAT 720

Qy        280 eValAspValTrp--GluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
              |   |||   |||     |||||||||||||||||||              ||||||
Db        721 TGGGGATGGCTGGGGAGTTTTTGGGGAAACCTAAA-----------ACATGG 762
```

RESULT 37
BI258848
LOCUS BI258848 767 bp mRNA linear EST 17-JUL-2001
DEFINITION 602969663F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5109303 5', mRNA sequence.
ACCESSION BI258848
VERSION BI258848.1 GI:14815606
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 767)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11264 row: m column: 16
High quality sequence stop: 758.

```
FEATURES             Location/Qualifiers
     source          1. .767
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5109303"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_12"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
ORIGIN
```

Alignment Scores:

Pred. No.:	1.6e-134	Length:	767
Score:	1238.00	Matches:	247
Percent Similarity:	98.0%	Conservative:	3
Best Local Similarity:	96.9%	Mismatches:	4
Query Match:	65.1%	Indels:	2
DB:	2	Gaps:	0

US-10-757-745-2 (1-362) x BI258848 (1-767)

```
Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCA 60

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTAC 120

Qy         62 PheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCC 180

Qy         82 LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSer 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGC 240

Qy        102 ProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATT 300

Qy        122 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCT 360

Qy        142 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTA 420

Qy        162 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCT 480

Qy        182 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 540

Qy        202 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCys 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGC 600

Qy        222 LeuMetThrSerHisLeuGluSerThrArgGlyHisAla-AlaGluArgMetAsnGlnLe 241
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        601 CTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTTGCGGAACGAATGAATCAGTT 660
```

```
Qy        241 uLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGl 261
              |:::::|||||||||   ||||||||||||||||||||||||||||||||||   |||||
Db        661 ACAACTGGTTTTAAAGACAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATT-GCAGG 719

Qy        261 yAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              |||||||:::|||  ||||||||||  ||||||||||
Db        720 AGATACAAGTCTAGGGGATCGAGAGGTTCCCAGATGTGGGGGT 762

RESULT 38
BG541819
LOCUS       BG541819                937 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602569712F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4694066 5',
            mRNA sequence.
ACCESSION   BG541819
VERSION     BG541819.1  GI:13534052
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 937)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1517  row: h  column: 03
            High quality sequence stop: 714.
FEATURES             Location/Qualifiers
     source          1. .937
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4694066"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_77"
                     /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.9
                     kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                4.99e-134       Length:          937
Score:                    1235.00         Matches:         273
Percent Similarity:       90.3%           Conservative:    6
Best Local Similarity:    88.3%           Mismatches:      19
Query Match:              64.9%           Indels:          13
DB:                       2               Gaps:            1

US-10-757-745-2 (1-362) x BG541819 (1-937)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 76

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 136

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 196

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 256

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 316

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 376

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 436

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 496

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu-GluGlyTyrPheTh 180
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        497 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAACGAAGGATATTTCAC 556

Qy        180 rAlaIleMetLeuLysLysSer-ArgValLysLeuLysSerGlnGluIleIlePro-Phe 199
              |||||||||||||||||||| |||||||||||  ||||||||||||||||||| |||
Db        557 AGCTATAATGTTGAAGAAATCAACGAGTGAAATTACCAAGCCAAGAGATTATTCCTTTTT 616

Qy        200 ProSerThr-LysMetMetArgAsnLeuLeuCysVal-HisValAsnValSerGlyAsnG 219
              ||||||||| ||||||||||||||||||||||||||| ||||||||||||||||||:::|
Db        617 CCAAGTACCAAAAATGATGAGAAACCTTTTATGTGTTGCATGTGAATGTGTCAGGACATG 676

Qy        219 luLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg-Met 238
              ||||||||  ||||||  |||  ||||||||||||||||||||||||||||||| |||
Db        677 AGCTTTGCTT-ATGACCATCCATTGGGAGAGCACCAGAGGGCATGCTGCGGAACGGAATG 735

Qy        239 AsnGlnLeu-LysMetValLeuLys-LysMetGlnGluAla-ProGluSerAlaThrVal 257
              ||||||::: |||:::|||||||| ||||||||||||||    |||||||||
Db        736 AATCAGGTCAAAAATTGGTTTTAAAGAAAAATGCAAGAGGCTCAAGGAGTCAGCTAACAGT 795

Qy        258 IlePheAlaGlyAspThrAsn-LeuArgAspArgGluValThrArgCysGlyGlyLeuPr 277
              ||||||  |||||||||||| ||||||||||||||      :::|||      |||||
Db        796 ATATTTGGCGGAGATACCAATCTTAAGGGATCGAGAAGGTTACCAATGTTGGTGGTTACC 855

Qy        277 oAsnAsnIleValAspValTrp---GluPheLeuGlyLysProLysHisCysGlnTyrTh 296
              |     ||||||||||||||   :::|||  ||||||  |||      :::|||
Db        856 CAAACAAATTGTGGATGTTTGGGGAAAGTTTGG-GGAAAGCTTAAAATGTGGCGTTATAA 914

Qy        296 rTrpAsp 298
              ||||||
Db        915 CTGGGAT 921

RESULT 39
CN298922
LOCUS       CN298922                729 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600077717 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN298922
VERSION     CN298922.1  GI:47315336
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
```

```
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 729   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

Alignment Scores:
Pred. No.:              4.37e-134       Length:         729
Score:                  1234.00         Matches:        238
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            64.9%           Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-2 (1-362) x CN298922 (1-729)

Qy          4 GlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluProGluValLys 23
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAA 61

Qy         24 LysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaValAla 43
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 AAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCT 121

Qy         44 GlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGlu 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAG 181

Qy         64 ProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThr 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACC 241

Qy         84 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSer 103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCT 301

Qy        104 GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGly 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGA 361

Qy        124 LeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyr 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTAC 421

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 481

Qy        164 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 541

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAA 601

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMet 223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATG 661

Qy        224 ThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 ACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTA 715

RESULT 40
CN298924
LOCUS       CN298924                729 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000424496911 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN298924
VERSION     CN298924.1  GI:47315338
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 729   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, cell lines H1, H7, and
                     H9"
                     /clone_lib="GRN_ES"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from undifferentiated hES cell lines H1 (p32), H7 (p29),
                     and H9 (p26) maintained in feeder-free conditions"
ORIGIN

Alignment Scores:
Pred. No.:              4.37e-134       Length:         729
Score:                  1234.00         Matches:        240
Percent Similarity:     99.6%           Conservative:   0
Best Local Similarity:  99.6%           Mismatches:     1
Query Match:            64.9%           Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-2 (1-362) x CN298924 (1-729)

Qy         25 ArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaValAlaGln 44
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CGTCCGCTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAG 66

Qy         45 CysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluPro 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 TGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCT 126

Qy         65 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTAT 186
```

```
Qy         85 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 104
Db        187 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 246

Qy        105 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 124
Db        247 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 306

Qy        125 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 144
Db        307 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGC 366

Qy        145 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 164
Db        367 CCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGA 426

Qy        165 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 184
Db        427 TCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTG 486

Qy        185 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 204
Db        487 AAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATG 546

Qy        205 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr 224
Db        547 ATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACA 606

Qy        225 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 244
Db        607 TCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTT 666

Qy        245 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 264
Db        667 TTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAAT 726

Qy        265 Leu 265
Db        727 CTA 729
```

RESULT 41
CN298923

LOCUS	CN298923 711 bp mRNA linear EST 16-MAY-2004
DEFINITION	17000600002122 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION	CN298923
VERSION	CN298923.1 GI:47315337
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 711)
AUTHORS	Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.
TITLE	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL	Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED	15146197
COMMENT	Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@geron.com Insert Length: 711 Std Error: 0.00.

FEATURES	Location/Qualifiers
source	1. .711 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /tissue_type="embryonic stem cells, DMSO-treated H9 cell line" /clone_lib="GRN_PREHEP" /note="oligo dT primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"

ORIGIN

Alignment Scores:

Pred. No.:	9.55e-130	Length:	711
Score:	1197.00	Matches:	235
Percent Similarity:	99.2%	Conservative:	0
Best Local Similarity:	99.2%	Mismatches:	1
Query Match:	62.9%	Indels:	1
DB:	7	Gaps:	0

US-10-757-745-2 (1-362) x CN298923 (1-711)

```
Qy         13 AlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGluPhe 32
Db          2 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 61

Qy         33 AlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrp 52
Db         62 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 121

Qy         53 GluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGlu 72
Db        122 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 181

Qy         73 ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 92
Db        182 CGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 241

Qy         93 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 112
Db        242 ACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGC 301

Qy        113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132
Db        302 ATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGG 361

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
Db        362 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 421

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
Db        422 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 481

Qy        173 GlyHis-GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 192
Db        482 GGTCATGGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 541

Qy        192 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 212
Db        542 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 601

Qy        212 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 232
Db        602 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 661

Qy        232 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
Db        662 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATG 710
```

RESULT 42
CX760857

LOCUS	CX760857 731 bp mRNA linear EST 24-JAN-2005
DEFINITION	AGENCOURT_40963383 NIH_MGC_281 Homo sapiens cDNA clone IMAGE:7783600 3', mRNA sequence.
ACCESSION	CX760857
VERSION	CX760857.1 GI:58057513

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 731)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Meri Firpo
             cDNA Library Preparation: Express Genomics
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15944  row: b  column: 14
            High quality sequence stop: 574.
FEATURES             Location/Qualifiers
     source          1. .731
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7783600"
                     /tissue_type="pluripotent cell line derived from
                     blastocyst inner cell mass"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_281"
                     /note="Organ: Blastocyst; Vector: pExpress-1; Site_1:
                     EcoRV; Site_2: NotI; RNA obtained from pluripotent cell
                     line derived from blastocyst inner cell mass (cell line
                     HSF-6, NIH Registry designation UC06. Positive for OCT4
                     expression by rtPCR, positive for SSEA-3, SSEA-4,
                     Tra-1-81, Tra-1-60 by immunofluorescence. Negative for
                     SSEA-1 by immunofluorescence Passage 62. cDNA was primed
                     using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. Size-selection >1.25
                     kb resulted in an average insert size of 2.0 kb. This
                     primary library is normalized (non-normalized primary
                     library is NIH_MGC_280) and was constructed by Express
                     Genomics (Frederick, MD). Note: this is a Mammalian Gene
                     Collection library."
ORIGIN

Alignment Scores:
Pred. No.:                5.82e-129      Length:          731
Score:                    1190.50        Matches:         236
Percent Similarity:       97.9%          Conservative:    1
Best Local Similarity:    97.5%          Mismatches:      5
Query Match:              62.6%          Indels:          2
DB:                       8              Gaps:            0

US-10-757-745-2 (1-362) x CX760857 (1-731)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 64

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 124

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 184

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 244

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 304

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 364

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 424

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 484

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 544

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              |||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GCTATAATGTTGAACAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 604

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 664

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        665 TGCCTTATGACATCCCATTTGGAGAGCACCAT-AGGCATTCTGCGAT-ACGATGAATCAG 722

Qy        241 LeuLys 242
              ||||||
Db        723 TTAAAA 728

RESULT 43
CR767800
LOCUS       CR767800                749 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp469C058_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469C058 5', mRNA sequence.
ACCESSION   CR767800
VERSION     CR767800.1  GI:52610302
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,
            Deutschenbaur,S., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix
            (Martinsried/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp469C058) is available at
            the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
            Berlin, Germany. Please contact RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469C058
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
```

```
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp469C058"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="469 (synonym: pkid1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:                6.93e-129      Length:          749
Score:                    1190.00        Matches:         225
Percent Similarity:       98.7%          Conservative:    3
Best Local Similarity:    97.4%          Mismatches:      3
Query Match:              62.6%          Indels:          0
DB:                       7              Gaps:            0

US-10-757-745-2 (1-362) x CR767800 (1-749)

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 AGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAG 66

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db         67 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGTTCAAGTAATTATGAGATTATT 126

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 ACAGGTCTTGAAGAAGGGTATTTCACAGCTATCATGTTGAAGAAATCAAGAGTGAAATTA 186

Qy        192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTG 246

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 CATGTGAGTGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 306

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCT 366

Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CCAGAGTCAGCTACAGTTATATTTGCAGGAGATACGAATCTAAGGGATCGAGAGGTTACC 426

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 AGATGTGGTGGTTTACCCAACAACATCGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAA 486

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGT 546

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||
Db        547 AAGCTTCGTTTTGATCGAATATTTTTCAGAGCAGCCGCAGAAGAGGGACACATTGTTCCC 606

Qy        332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHis 351
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CAAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCAC 666

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||||||||||   |||||||||
Db        667 TGGGGTCTTCTGTGCAACTTANATATAATATTG 699

RESULT 44
BU661472
```

```
LOCUS       BU661472                710 bp    mRNA    linear   EST 30-SEP-2002
DEFINITION  cl72d11.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo
            sapiens cDNA clone cl72d11 5', mRNA sequence.
ACCESSION   BU661472
VERSION     BU661472.1  GI:23373654
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 710)
  AUTHORS   Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
  TITLE     Gene Expression in Human Erythroid Precursor Cells
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Jeffery L. Miller
            Laboratory of Chemical Biology
            National Institute of Diabetes and Digestive and Kidney Diseases
            Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
            20892, USA
            Tel: 301 402 2373
            Fax: 301 435 5148
            Email: jm7f@nih.gov
            The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
            Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
            analyses by National Institutes of Health Intramural Sequencing
            Center (NISC). More information available at:
            http://hembase.niddk.nih.gov
            Plate: 72  row: d  column: 11
            Seq primer: 5' lambda-TriplEx2 Sequencing Primer.
FEATURES             Location/Qualifiers
     source          1. .710
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl72d11"
                     /sex="unknown"
                     /tissue_type="blood"
                     /cell_type="Erythroid Precursor Cells"
                     /cell_line="Primary Culture of Peripheral Blood
                     Mononuclear Cells"
                     /dev_stage="Precursor erythroblasts; GPA++"
                     /lab_host="DH5alpha"
                     /clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl
                     library)"
                     /note="Organ: blood; Vector: pTriplEx2; Site_1: SfiI;
                     Site_2: SfiI; A complementary DNA (cDNA) library from
                     human erythroid precursor cells was constructed using
                     SMART PCR (polymerase chain reaction) cDNA Library
                     Construction Kit (Clontech, Palo Alto, CA) according to
                     the manufacturer's directions, but with slight
                     modifications. Briefly, reverse transcription was
                     performed in the presence of 1 umol/L peptide nucleic acid
                     (PNA) oligos
                     (N-terminal)-biotin-GTC-CAC-CCG-AAG-CTT-G-(C-terminal) and
                     (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)
                     . Synthesized cDNA was digested with SfiI and
                     size-selected on a 1% agarose gel (>800bp). Large-scale
                     sequencing of the library was performed by the NIH
                     Intramural Sequencing Center (NISC;
                     Http://www.nisc.nih.gov/)."
ORIGIN

Alignment Scores:
Pred. No.:                1.09e-128      Length:          710
Score:                    1188.00        Matches:         231
Percent Similarity:       99.6%          Conservative:    0
Best Local Similarity:    99.6%          Mismatches:      1
Query Match:              62.5%          Indels:          0
DB:                       5              Gaps:            0

US-10-757-745-2 (1-362) x BU661472 (1-710)
```

```
Qy         33 AlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrp 52
Db         13 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 72

Qy         53 GluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGlu 72
Db         73 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 132

Qy         73 ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 92
Db        133 CGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 192

Qy         93 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 112
Db        193 ACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGC 252

Qy        113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132
Db        253 ATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGG 312

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
Db        313 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 372

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
Db        373 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 432

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
Db        433 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 492

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
Db        493 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCCTTTATGTGTGCAT 552

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
Db        553 GTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 612

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
Db        613 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 672

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsn 264
Db        673 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAAT 708
```

RESULT 45
BI334820
LOCUS BI334820 704 bp mRNA linear EST 30-JUL-2001
DEFINITION 602998939F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141193 5', mRNA sequence.
ACCESSION BI334820
VERSION BI334820.1 GI:15019477
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 704)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11347 row: n column: 10
High quality sequence stop: 704.

FEATURES Location/Qualifiers
source 1. .704
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5141193"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

ORIGIN

Alignment Scores:

Pred. No.:	1.41e-128	Length:	704
Score:	1187.00	Matches:	231
Percent Similarity:	99.1%	Conservative:	1
Best Local Similarity:	98.7%	Mismatches:	1
Query Match:	62.4%	Indels:	1
DB:	3	Gaps:	0

US-10-757-745-2 (1-362) x BI334820 (1-704)

```
Qy          6 CysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluProGluValLysLysArg 25
Db          3 TGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGG 62

Qy         26 ArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCys 45
Db         63 CGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGC 122

Qy         46 PheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProPro 65
Db        123 TTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCG 182

Qy         66 ValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrVal 85
Db        183 GTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTT 242

Qy         86 AspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 105
Db        243 GACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGAT 302

Qy        106 ThrGlnGln-GluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAs 125
Db        303 ACTCAGCAAGCAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGA 362

Qy        125 pLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPr 145
Db        363 TCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCC 422

Qy        145 oAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSe 165
Db        423 AGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATC 482

Qy        165 rSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLy 185
Db        483 AAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTCGAA 542

Qy        185 sLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMe 205
Db        543 GAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGAT 602

Qy        205 tArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSe 225
Db        603 GAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATC 662

Qy        225 rHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
```

```
                 ||||||||||||||||||||||||||||||||||||||||
Db        663 CCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATG 702

RESULT 46
BI760756
LOCUS       BI760756                 757 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603044763F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5185334 5',
            mRNA sequence.
ACCESSION   BI760756
VERSION     BI760756.1  GI:15752334
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 757)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11462  row: m  column: 15
            High quality sequence start: 6
            High quality sequence stop: 755.
FEATURES             Location/Qualifiers
     source          1. .757
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5185334"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_116"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female.  Library is
                     oligo-dT primed and  directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.4 kb,
                     insert size range 1-3 kb.  Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber  (Invitrogen).  Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.59e-128       Length:         757
Score:                  1187.00         Matches:        243
Percent Similarity:     98.0%           Conservative:   0
Best Local Similarity:  98.0%           Mismatches:     4
Query Match:            62.4%           Indels:         4
DB:                     3               Gaps:           0

US-10-757-745-2 (1-362) x BI760756 (1-757)

Qy          37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGlu-MetGluAr 56
               ||||||||||||  ||||||||||||||||||||||||||||||||||| ||||||||
Db          17 AGCTGCGATGCCGGAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGGATGGAAAG 76

Qy          56 gAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGl 76
               ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db          77 GGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGC-CTGGAACGCCGACCTGA 135

Qy          76 uThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerTh 96
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         136 AACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCAC 195

Qy          96 rThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLe 116
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         196 CACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCT 255

Qy         116 uIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVa 136
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         256 CATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGT 315

Qy         136 lCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProPr 156
               ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db         316 GTGTTCCTACTTAGC-TTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCC 374

Qy         156 oTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGl 176
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         375 ATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGA 434

Qy         176 uGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIl 196
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         435 AGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGAT 494

Qy         196 eIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSe 216
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         495 TATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTC 554

Qy         216 rGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGl 236
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         555 AGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGA 614

Qy         236 uArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaTh 256
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         615 ACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTAC 674

Qy         256 rValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLe 276
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         675 AGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTT 734

Qy         276 uProAsnAsnIleValAspVal 283
               |||||||   ||||||||||||
Db         735 ACCCAACAA-ATTGTGGATGTC 755

RESULT 47
AU139147
LOCUS       AU139147                 752 bp    mRNA    linear   EST 02-AUG-2002
DEFINITION  AU139147 PLACE1 Homo sapiens cDNA clone PLACE1010031 5', mRNA
            sequence.
ACCESSION   AU139147
VERSION     AU139147.1  GI:11000668
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
```

```
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1010031"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="Vector: pME18SFL3"
ORIGIN

Alignment Scores:
Pred. No.:                6.1e-128      Length:          752
Score:                    1182.00       Matches:         238
Percent Similarity:       94.8%         Conservative:    1
Best Local Similarity:    94.4%         Mismatches:      10
Query Match:              62.1%         Indels:          3
DB:                       1             Gaps:            0

US-10-757-745-2 (1-362) x AU139147 (1-752)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArg-AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 660

Qy        220 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 240
              |||||||   ||||||   |||    :::|||   ||||||      ||||||||||||||
Db        661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy        240 lnLeuLysMetValLeu-LysLysMetGln 249
              ||   ||||||   |||    |||||||||
Db        721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750

RESULT 48
```

```
CR767236
LOCUS       CR767236                701 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp469E1236_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469E1236 5', mRNA sequence.
ACCESSION   CR767236
VERSION     CR767236.1  GI:52609173
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 701)
  AUTHORS   Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ansorge,W., Krieger,S., Regiert,T., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; rlin, Germany. Please contact
            RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469E1236
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES              Location/Qualifiers
     source           1. .701
                      /organism="Pongo pygmaeus"
                      /mol_type="mRNA"
                      /db_xref="taxon:9600"
                      /clone="DKFZp469E1236"
                      /tissue_type="kidney"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="469 (synonym: pkid1)"
                      /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:                9.37e-128     Length:          701
Score:                    1180.00       Matches:         229
Percent Similarity:       99.1%         Conservative:    1
Best Local Similarity:    98.7%         Mismatches:      2
Query Match:              62.0%         Indels:          0
DB:                       7             Gaps:            0

US-10-757-745-2 (1-362) x CR767236 (1-701)

Qy         26 ArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCys 45
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 CGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCGAGCTGCGATACCGCTGTGGCTCAGTGC 63

Qy         46 PheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProPro 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 TTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCG 123

Qy         66 ValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrVal 85
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 GTGGAGGAGAGCGCCTTGGAACGCCGCCCTGAAACCATCTCTGAGCCCAAGACCTATGTT 183

Qy         86 AspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGAT 243

Qy        106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGAT 303

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        304 CTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCA 363

Qy        146 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        364 GATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGTTCA 423

Qy        166 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AGTAATTATGAGATTATTACAGGTCATGAAGAAGGGTATTTCACAGCTATCATGTTGAAG 483

Qy        186 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 AAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATG 543

Qy        206 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSer 225
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        544 AGAAACCTTTTATGTGTGCATGTGAGTGTGTCAGGAAATGAGCTTTGCCTTATGACATCC 603

Qy        226 HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeu 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 CATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTA 663

Qy        246 LysLysMetGlnGluAlaProGluSerAlaThrVal 257
              |||||||||||||||||||||||||||||||||||
Db        664 AAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTT 699

RESULT 49
AL516230
LOCUS       AL516230                816 bp    mRNA    linear   EST 23-MAR-2004
DEFINITION  AL516230 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            CS0DA005YK22 5-PRIME, mRNA sequence.
ACCESSION   AL516230
VERSION     AL516230.3  GI:45652882
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 816)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:30490655.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DA005BF11QP1&c=3474.r.
FEATURES              Location/Qualifiers
     source           1. .816
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DA005YK22"
                      /tissue_type="NEUROBLASTOMA"
                      /clone_lib="Homo sapiens NEUROBLASTOMA"
                      /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                      with a NotI-oligo(dT) primer. Five prime end enriched,
                      double-strand cDNA was digested with Not I and cloned into
                      the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                      Library was not normalized."
ORIGIN

Alignment Scores:
Pred. No.:              1.38e-125       Length:         816
```

```
Score:                  1162.50         Matches:        249
Percent Similarity:     94.1%           Conservative:   4
Best Local Similarity:  92.6%           Mismatches:     6
Query Match:            61.1%           Indels:         10
DB:                     1               Gaps:           1

US-10-757-745-2 (1-362) x AL516230 (1-816)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeu-CysValGluPheAlaSerValAla-SerCysAspA 40
              |||||||||||||||||||||||| |||||||||||||||||||||||| ||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGGTGTGTGGAGTTTGCCTCGGTCGCGAAGCTGCGATG 145

Qy         40 laAlaValAlaGlnCysPhe----LeuAlaGluAsnAspTrpGluMetGluArgAlaLeu 58
              ||||||||||||||||||       :::||||||||||||||||||   ||||||||||||
Db        146 CCGCAGTGGCTCAGTGCGTGTCGCGTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTG 205

Qy         59 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIl 78
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        206 AACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGGCCGACCTGAAACCAT 265

Qy         78 eSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSe 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTC 325

Qy         98 rLysIleSerProSerGluAspThr--GlnGlnGluAsnGlySer-MetPheSerLeuIl 117
              |||||||||||||||||||||||:::  :::|||||||||||||| |||   ||||||||
Db        326 TAAAATCAGCCCATCTGAAGATAGCTCAGGCAAGAAAATGGCAGCGATGTKCTCTCTCAT 385

Qy        117 eThr-TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValC 137
              |::: ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 TAGCCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGT 445

Qy        137 ysSerTyrLeuAlaLeuTyrSerProAspValIlePhe-LeuGlnGluValIleProPro 156
              ||   ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        446 GTKCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTGCTACAGGAAGTTATTCCCCCA 505

Qy        157 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAA 565

Qy        177 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATT 625

Qy        197 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 ATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTSA 685

Qy        217 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 236
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 GGAAATGAGCGTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAA 745

Qy        237 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 CGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACA 805

Qy        257 ValIlePhe 259
              |||||||||
Db        806 GTTATATTT 814

RESULT 50
AK009089
LOCUS       AK009089                1289 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male tongue cDNA, RIKEN full-length enriched
            library, clone:2310002D07 product:Traf and Tnf receptor associated
            protein, full insert sequence.
ACCESSION   AK009089
VERSION     AK009089.1  GI:12843662
KEYWORDS    HTC; CAP trapper.
```

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1289)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES Location/Qualifiers
source 1. .1289
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2310002D07"
/db_xref="taxon:10090"
/clone="2310002D07"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature 55. .1160
/note="Traf and Tnf receptor associated protein (MGD|MGI:1860486, GB|NM_019551, evidence: BLASTN, 99%, match=1184)
putative"
polyA_signal 1260. .1265
/note="putative"
polyA_site 1289
/note="putative"
ORIGIN

Alignment Scores:

Pred. No.:	7.41e-125	Length:	1289
Score:	1159.00	Matches:	238
Percent Similarity:	80.1%	Conservative:	43
Best Local Similarity:	67.8%	Mismatches:	66
Query Match:	60.9%	Indels:	8
DB:	4	Gaps:	3

US-10-757-745-2 (1-362) x AK009089 (1-1289)

```
Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31
                |||||||||:::   |||  |||       ||||||:::|||||||||   |||:::
Db        121 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 177

Qy         32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51
                ||||||  |||       ||||||       :::           |||   |||||||||
Db        178 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 237

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                |||:::   ::::::||||||:::::|||||||||   |||   |||         :::
Db        238 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGA-GAACGACCAAGGGTG 296

Qy         72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91
                  ::|||   |||       :::   :::   ||||||||||||||||||||:::
Db        297 GCCGCGCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 356

Qy         92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111
                ||:::|||      :::   |||||||||       |||       |||:::
Db        357 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 413

Qy        112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131
                |||       |||   ||||||||||||||||||||||||||       ||||||   |||
Db        414 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 473

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
                ||||||||||||||||    |||||||||||||||||||||:::|||||||||
Db        474 AGGGCTCGAGGGGTGTGTCCTGC-CTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 532

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
                ||||||   |||||||||   :::|||||||||||||||:::::::::|||   ||||||
Db        533 GAAGTTAT-CCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 591

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191
                ||||||:::|||||||||||||||||||||||:::|||||||||   |||||||||
Db        592 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGA-AGAGTGAAATTT 650

Qy        192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211
                ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
```

```
Db        651 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 710

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
              :::|||:::::::    |||||||||    ||||||||||||||||||||||||||||||||
Db        711 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 770

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
                  |||:::|||||||||:::    |||||||||    ||||||    ||||||||||||||||
Db        771 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 830

Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271
              |||:::|||    ||||||||||||||||||||||||||||||||||:::||||||
Db        831 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 890

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
              :::|||||||||||||||:::|||:::    |||    ||||||||||||||||||||||||
Db        891 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 950

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
              ||||||||||||||||||||||||:::    |||:::||||||    |||    ||||||
Db        951 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1010

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              |||    ||||||||||||||||||||||||       ||||||||||||||:::||||||
Db       1011 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1064

Qy        332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHis 351
              :::|||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db       1065 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1124

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||||    |||:::::::|||
Db       1125 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1157

RESULT 51
CR791801
LOCUS       CR791801                739 bp    mRNA    linear   EST 01-OCT-2004
DEFINITION  DKFZp468J1522_r1 468 (synonym: phrt1) Pongo pygmaeus cDNA clone
            DKFZp468J1522 5', mRNA sequence.
ACCESSION   CR791801
VERSION     CR791801.1  GI:53710641
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,
            Deutschenbaur,S., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix
            (Martinsried/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp468J1522) is available at
            the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
            Berlin, Germany. Please contact RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp468J1522
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES              Location/Qualifiers
     source           1. .739
                      /organism="Pongo pygmaeus"
                      /mol_type="mRNA"
                      /db_xref="taxon:9600"
                      /clone="DKFZp468J1522"
                      /tissue_type="heart"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="468 (synonym: phrt1)"
                      /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:                8.96e-125       Length:          739
Score:                    1155.00         Matches:         226
Percent Similarity:       95.0%           Conservative:    1
Best Local Similarity:    94.6%           Mismatches:      12
Query Match:              60.7%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-2 (1-362) x CR791801 (1-739)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||    ||||||||||||||||||||||||||||||||||||||||
Db         19 ATGGAGTTGGGGAGTTTCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCC 78

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GAGGTGAAGAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCGAGCTGCGATACC 138

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||    ||||||||||||||||||||||||||    |||||||||
Db        139 GCTGTGGCTCAGTGCTTCCTGGNCGAGAACGACTGGGAGATGGAAAGGGNTCTGAACTCC 198

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGCCCTGAAACCATCTCTGAG 258

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 318

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||    ||||||||||||||||||||||||
Db        319 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCNTGTTCTCTCTCATTACCTGGAAT 378

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 438

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 GCTTTGTNCAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 498

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              |||||||||    ||||||    ||||||||||||||||||||||||||||||    ||||||
Db        499 CTAAAGAAGAGTTCAAGTNATTATGAGATTATTACAGGTCATGAAGAAGGGGATTTCACA 558

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              |||||||||    ||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 GCTATCATGTNGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 618

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||:::||||||    |||||||||
Db        619 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAGTGTGTCAAGAAATGAGCTT 678

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
              ||||||||||||||||||||||||||||||    ||||||||||||||||||||||
Db        679 TGCCTTATGACATCCCATTTGGAGAGCACCANAGGGCATGCTGCGGAACGAATGAAT 735

RESULT 52
BI754101
LOCUS       BI754101                774 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603027659F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5197838 5',
            mRNA sequence.
ACCESSION   BI754101
VERSION     BI754101.1  GI:15745679
KEYWORDS    EST.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 774)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11495  row: f  column: 15
            High quality sequence stop: 756.
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5197838"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_114"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
ORIGIN
```

Alignment Scores:

Pred. No.:	1.66e-124	Length:	774
Score:	1153.00	Matches:	232
Percent Similarity:	97.5%	Conservative:	1
Best Local Similarity:	97.1%	Mismatches:	6
Query Match:	60.6%	Indels:	2
DB:	3	Gaps:	0

US-10-757-745-2 (1-362) x BI754101 (1-774)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 119

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 179

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 239

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 299

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 359

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 419

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 479

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 539

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 599

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db        600 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAGAAAGCCAAGAGATTATTCCTTTTCCA 659

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              |||||||||||||||||||  |||||||||||||||||||||||  |||||||||
Db        660 AGTACCAAAATGATGAGAAACCT-TTATGTGTGCATGTGAATGTGTCAGG-AATGAGCTT 717

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 TTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGGGAACGGAATGAAT 774
```

```
RESULT 53
CN298920
LOCUS       CN298920                657 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600170525 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN298920
VERSION     CN298920.1  GI:47315334
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 657   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cell, retinoic acid and
                     mitogen-treated hES cell line H7"
                     /clone_lib="GRN_PRENEU"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from hES cell line H7 (p29) maintained in feeder-free
                     conditions. Embryoid bodies were generated in the presence
                     of all-trans retinoic and mitogens."
ORIGIN
```

Alignment Scores:

Pred. No.:	2.89e-122	Length:	657
Score:	1133.00	Matches:	218
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	59.6%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2 (1-362) x CN298920 (1-657)

```
Qy          16 GluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGluPheAlaSerVal 35
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           2 GAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTC 61

Qy          36 AlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu 55
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 GCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAA 121

Qy          56 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 75
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 AGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCT 181

Qy          76 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 95
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         182 GAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCC 241

Qy          96 ThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 115
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 ACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCT 301

Qy         116 LeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGly 135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         302 CTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGG 361

Qy         136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         362 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 421

Qy         156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         422 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 481

Qy         176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         482 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 541

Qy         196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         542 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 601

Qy         216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 233
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         602 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 655
```

RESULT 54
BG719113

```
LOCUS       BG719113                678 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602699045F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4831186 5',
            mRNA sequence.
ACCESSION   BG719113
VERSION     BG719113.1  GI:13998300
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10753  row: i  column: 11
            High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4831186"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN
```

Alignment Scores:

Pred. No.:	8.99e-122	Length:	678
Score:	1129.00	Matches:	218
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	59.4%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2 (1-362) x BG719113 (1-678)

```
Qy           1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          23 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 82

Qy          21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          83 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 142

Qy          41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         143 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 202

Qy          61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         203 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 262

Qy          81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         263 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 322

Qy         101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         323 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 382

Qy         121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         383 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 442

Qy         141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         443 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 502

Qy         161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         503 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 562

Qy         181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         563 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 622
```

```
Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAAC 676

RESULT 55
BE784416
LOCUS       BE784416                853 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601473891F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3876862 5',
            mRNA sequence.
ACCESSION   BE784416
VERSION     BE784416.1  GI:10205614
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 853)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9638  row: e  column: 23
            High quality sequence stop: 660.
FEATURES             Location/Qualifiers
     source          1. .853
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3876862"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

Alignment Scores:
Pred. No.:              2.24e-120       Length:         853
Score:                  1118.50         Matches:        233
Percent Similarity:     94.1%           Conservative:   5
Best Local Similarity:  92.1%           Mismatches:     14
Query Match:            58.8%           Indels:         4
DB:                     2               Gaps:           1

US-10-757-745-2 (1-362) x BE784416 (1-853)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 65

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db         66 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTT-GCCTCGGTCGCAAGCTGCGATGCC 124

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 184

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 244

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 304

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 364

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 424

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 484

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 544

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 604

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn---Glu 219
              ||||||||||||||||||||   ||||||||||||||||||||||||||          :::
Db        605 AGTACCAAAATGATGAGAAACCT-TTATGTGTGCATGTGAATGTGTCCAGGAAAATGAAG 663

Qy        220 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
              |||   |||||||||   ||||||          |||||          :::|||||||||
Db        664 CTTTGGCTTATGACATTCCATTTGGGAGAGCACAGAGGGATGGTGGGGAAACGAATGAAT 723

Qy        240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              |||   ||||||||||||:::::::|||:::|||   |||
Db        724 CAGTT-AAAATGGTTTTACAGGAAATGGAGGAGGGTCCG 761

RESULT 56
DN122208
LOCUS       DN122208                798 bp    mRNA    linear   EST 15-FEB-2005
DEFINITION  1122397 MARC 4PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   DN122208
VERSION     DN122208.1  GI:59816487
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 798)
  AUTHORS   Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
            Nonneman,D.J., Wray,J.E. and Keele,J.W.
  TITLE     Porcine EST collection using a normalized library constructed from
            embryos representing early developmental stages
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: HHY8024  row: L  column: 17
            Seq primer: GTAATACGACTCACTATAGGG.
FEATURES             Location/Qualifiers
     source          1. .798
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 4PIG"
```

/note="Vector: pcDNA3.1; Site_1: EcoRI; Site_2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."
ORIGIN

Alignment Scores:

Pred. No.:	3.96e-120	Length:	798
Score:	1116.00	Matches:	217
Percent Similarity:	91.9%	Conservative:	21
Best Local Similarity:	83.8%	Mismatches:	21
Query Match:	58.7%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2 (1-362) x DN122208 (1-798)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||      |||   ||||||   |||   |||      :::||||||||||||||||
Db         21 ATGGAGCAGAAAAGCGGCCTGGAGGATGGACCGGAGCTGGAGCAGGAGGAGGGAGAGCCC 80

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||:::||||||||||
Db         81 GAGGTGAAGAAGCGCCGACTTCTGTGTGTGGAGTTTGCCTCGGTAGCGACTTGCGACGCT 140

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        141 GCCGTGGCTCAGTGCTACCTGGCCGAGAACGATTGGGAGATGGAAAGAGCGCTGAACTCC 200

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||   ||||||||||:::|||||||:::   ||||||||||||   :::|||
Db        201 TACTTCGAGCCGGCCGTGGAGGAGAACGCTTTAAAAAGTCGTCCTGAGACCCCCAATGAG 260

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              |||::::::   ||||||||||||   |||||||||:::||||||   :::   |||
Db        261 CCCGAGTCCTGTGTTGACTTAACCAAGGAAGAAACAAGTGATTCCTTTAGTTGTAAAACC 320

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              |||   |||||||:::   |||||||||:::|||||||:::||||||   ||||||||||||
Db        321 AGCACATCTGAAAATGTTCAGCAAGAAGATGGCAGTGTGTTCTCTTTCATTACCTGGAAT 380

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||:::|||   ||||||||||||||||||||||||||||||
Db        381 ATAGATGGATTGGACCTAAACAGTCTGCTCGAGAGGGCTCGAGGGGTGTGTTCCTATTTA 440

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        441 GCATTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCGTATTATACCTAC 500

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              |||||||||:::::|||:::|||||||||||||||||||:::||||||||||||||||||
Db        501 CTTAAGAAGAAAGCAAGTAGTTATGAGATTATTACAGGTAATGAAGAAGGATATTTCACA 560

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db        561 GCTATAATGTTGAAGAAATCGAGAGTGAAATTTAAAAGCCAAGAGATTATTCCTTTTCCA 620

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              :::||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        621 AATACAAAAATGATGAGAAACCTTTTGTGTGTGCATGTGAGTGTGTCAGGAAATGAACTT 680

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        681 TGCCTTATGACTTCTCACTTGGAGAGCACCAGAGGGCATGCCGAGGAACGAATGAATCAG 740

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 CTGAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTT 797
```

RESULT 57
BG119064
LOCUS BG119064 927 bp mRNA linear EST 30-JAN-2001
DEFINITION 602347589F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4442399 5',
mRNA sequence.
ACCESSION BG119064
VERSION BG119064.1 GI:12612570
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 927)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10215 row: a column: 24
High quality sequence stop: 676.
FEATURES Location/Qualifiers
source 1. .927
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4442399"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:

Pred. No.:	1.3e-119	Length:	927
Score:	1112.50	Matches:	245
Percent Similarity:	85.1%	Conservative:	7
Best Local Similarity:	82.8%	Mismatches:	36
Query Match:	58.5%	Indels:	12
DB:	2	Gaps:	5

US-10-757-745-2 (1-362) x BG119064 (1-927)

```
Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAA 60

Qy         77 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 96
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACC 120

Qy         97 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTC 180

Qy        117 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVal 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTG 240

Qy        137 CysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProPro 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCA 300

Qy        157 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAA 360

Qy        177 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 196
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATT 420

Qy        197 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 216
                ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db        421 ATTCCTTTTCCAAGTACCAAAATGATGAGAAACTT-TTATGTGTGCATGTGAATGTGTCA 479

Qy        217 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 236
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAA 539

Qy        237 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 256
                ||||||||||||||||||||||||||||||||||   ||||||||||||||||   |||
Db        540 CGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCA-GAGGCTCCGGAGTCAGT-ACA 597

Qy        257 ValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeu 276
                ||||||   ||||||||||||||||   |||||||||||||||||||||||||||||||
Db        598 GTTATATCTGCAGGAGATACAAATCT-AGGGATCGAGAGGTTACCAGATGTGGTGGTTAC 656

Qy        277 ProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 296
                   ||||||||||||||||||||   ||||||                  :::
Db        657 ---CAAAACATTGTGGATGTCTGGGAGGTTTTGGGGAACTTAACATTGGCAGCATTC--- 710

Qy        297 TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAsp 316
                |||:::||||||||:::   |||   |||||||||         ||||||
Db        711 TGGGAGACACAAATGGACCCTAATTCTGGGATAACTGTGGGTGTGAAACTCGGTTGATCG 770

Qy        317 ArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeu 336
                |||   ||||||||||||   :::   |||   :::   ||||||
Db        771 AATATT---TTCAGAGCAGCGCGAAAAAGGGGA---CTTTTCCCGAGG------------ 812

Qy        337 LeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrp 352
                |||                  |||:::|||   :::      |||
Db        813 TTGAGCTCTTGTGTTAAAACGGCGGGGAAATTCACGAACGGGCTGTGG 860

RESULT 58
BG498689
LOCUS       BG498689                 770 bp    mRNA    linear   EST 27-MAR-2001
DEFINITION  602544301F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4666633 5',
            mRNA sequence.
ACCESSION   BG498689
VERSION     BG498689.1  GI:13460206
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 770)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1474  row: a  column: 02
            High quality sequence stop: 668.
FEATURES             Location/Qualifiers
     source          1. .770
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4666633"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_60"
                     /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI
                     (ggccattatggcc); Double-stranded cDNA was prepared from
                     cell line RNA.  5' and 3' adaptors were used in cloning as
                     follows: 5' adaptor sequence:  5'-CACGGCCATTATGGCC-3' and
                     3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.5
                     kb (range 0.9-4.0 kb). 14/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                1.67e-118       Length:          770
Score:                    1102.00         Matches:         219
Percent Similarity:       97.4%           Conservative:    2
Best Local Similarity:    96.5%           Mismatches:      1
Query Match:              57.9%           Indels:          5
DB:                       2               Gaps:            0

US-10-757-745-2 (1-362) x BG498689 (1-770)

Qy        140 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSer 159
                |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTTAGCTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGC 61

Qy        160 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 179
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTC 121

Qy        180 ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 199
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTT 181

Qy        200 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 219
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAG 241

Qy        220 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAAT 301

Qy        240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 259
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTT 361

Qy        260 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsn 279
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAAC 421

Qy        280 IleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThr 299
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACA 481

Qy        300 GlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe 319
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTCGATCGAATATTT 541

Qy        320 PheArgAlaAlaAlaGluGluGlyHisIleIleProArg-SerLeuAspLeuLeuGly-L 339
                ||||||||||||||||||||||||||||||||||||| |||||||||||||||||| |
Db        542 TTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGGAAGTTTGGACCTTCTTGGCAT 601

Qy        339 euGluLysLeuAsp-CysGlyArgPhePro-SerAspHisTrpGlyLeuLeuCys-AsnL 358
                |||||||||||||| ||||||||||||||    ||||||||||||||||||||| ||||
Db        602 TAGAAAAACTGGACCTGTGGTAGATTTCCTAGGTGATCACTGGGGTCTTCTGTGCAAACT 661

Qy        358 euAspIleIle 361
                ||||||||:::
Db        662 TAGATATACTA 672

RESULT 59
BG391213
```

```
LOCUS       BG391213                908 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  602417244F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4536635 5',
            mRNA sequence.
ACCESSION   BG391213
VERSION     BG391213.1  GI:13284661
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 908)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10460  row: h  column: 12
            High quality sequence stop: 732.
FEATURES             Location/Qualifiers
     source          1. .908
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4536635"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_92"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              6.42e-118       Length:         908
Score:                  1098.00         Matches:        247
Percent Similarity:     82.1%           Conservative:   6
Best Local Similarity:  80.2%           Mismatches:     41
Query Match:            57.7%           Indels:         16
DB:                     2               Gaps:           3

US-10-757-745-2 (1-362) x BG391213 (1-908)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              |||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTT-GCCTCGGTCGCAAGCTGCGATGCC 119

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 179

Qy         61 TyrPheGluProProValGluGluSerAla-LeuGluArgArgProGlu-ThrIleSerG 80
              |||||||||||||||||||||||||||||| ||||||||||||||||    |||||||||
Db        180 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCGTTGGAACGCCGACCTGTAAACCATCTCTG 239

Qy         80 luProLysThrTyrValAspLeu--ThrAsnGluGluThrThrAspSerThrThrSerLy 99
              ||||||||||||||||||||||    :::|||||||||||||||||||||||||||||||
Db        240 AGCCCAAGACCTATGTTGACCTAGACCAGATGAAGAAACAACTGATTCCACCACTTCTAA 299

Qy         99 sIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 AATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTG 359

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTA 419

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAG 479

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTT 539

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
              ||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        540 CACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAGAAGCCAAGAGATTATTCCTTT 599

Qy        199 eProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn-ValSerGlyAsnG 219
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        600 TCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACCGTGTCAGGAAATG 659

Qy        219 luLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla-AlaGluArgMet 238
              ||||||||   ||||||   |||   ||||||||||||||||   ||| |||   ||||||
Db        660 AGCTTTGCTT-ATGACATNCCATTGGGAGAGCACCAGAGGGGATGCTGGCGGGACGAATG 718

Qy        239 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258
              ||||||||||||          :::|||   :::          ||||||   |||   |||
Db        719 AATCAGTTAAAATGGTTCCAACGAAAATGGCGA---GGCTCCGAGTCAGTTACCGGTATA 775

Qy        259 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 278
              |||         :::         |||                 |||   |||
Db        776 TTGCGGGGA---------ATCAATTCTAGGGTCCAGGTTCCAGCGTGGGGGTTCCCCCAA 826

Qy        279 AsnIleValAspValTrp------------GluPheLeuGlyLysProLysHisCysGln 294
                          ||||||               |||      :::
Db        827 CTGTGGTTTCGGGTTTGGGCACCTACACTGCCGTTCTGGGACCAGGCCCTCTGGAACGCG 886

Qy        295 TyrThrTrpAsp 298
                 |||||||||
Db        887 CGAACGTGGGAT 898

RESULT 60
CB157906
LOCUS       CB157906                634 bp    mRNA    linear   EST 29-JAN-2003
DEFINITION  K-EST0217150 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-5-F06
            5', mRNA sequence.
ACCESSION   CB157906
VERSION     CB157906.1  GI:28143040
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 5  row: F  column: 06
            High quality sequence stop: 634.
FEATURES             Location/Qualifiers
     source          1. .634
```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L18POOL1n1-5-F06"
/cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
/lab_host="Top10F'"
/clone_lib="L18POOL1n1"
/note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI; The library was contributed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."

ORIGIN

Alignment Scores:

Pred. No.:	8.13e-118	Length:	634
Score:	1095.00	Matches:	211
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	57.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2 (1-362) x CB157906 (1-634)

```
Qy         17 GluGlyGluProGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAla 36
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCA 60

Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGG 120

Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAA 180

Qy         77 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 96
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACC 240

Qy         97 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTC 300

Qy        117 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVal 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTG 360

Qy        137 CysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProPro 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCA 420

Qy        157 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAA 480

Qy        177 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATT 540

Qy        197 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCA 600

Qy        217 GlyAsnGluLeuCysLeuMetThrSerHisLeu 227
              |||||||||||||||||||||||||||||||
Db        601 GGAAATGAGCTTTGCCTTATGACATCCCATTTG 633
```

RESULT 61
BG431303
LOCUS BG431303 650 bp mRNA linear EST 14-MAR-2001
DEFINITION 602499879F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4613587 5', mRNA sequence.
ACCESSION BG431303
VERSION BG431303.1 GI:13337809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 650)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM1363 row: n column: 20
High quality sequence stop: 648.
FEATURES Location/Qualifiers
source 1. .650
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4613587"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_75"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	8.46e-118	Length:	650
Score:	1095.00	Matches:	211
Percent Similarity:	98.1%	Conservative:	1
Best Local Similarity:	97.7%	Mismatches:	4
Query Match:	57.6%	Indels:	1
DB:	2	Gaps:	0

US-10-757-745-2 (1-362) x BG431303 (1-650)

```
Qy        145 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGA 61

Qy        165 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTG 121

Qy        185 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATG 181

Qy        205 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACA 241

Qy        225 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTT 301
```

```
Qy        245 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAAT 361

Qy        265 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGG 421

Qy        285 GluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAAT 481

Qy        305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCA 541

Qy        325 GluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCys 344
              |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        542 GAAGAGGGACACATTATTCCCCGAAG-TTGGACCTTCTTGGATTAGAAAAACTGGACTGT 600

Qy        345 GlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              |||||||||||||||||||||||||   :::||||||||||
Db        601 GGTAGATTTCCTAGTGATCACTGGGGGTCTTCTGTGCACTTAGATATA 648

RESULT 62
CK952220
LOCUS       CK952220                  680 bp    mRNA    linear   EST 15-MAR-2004
DEFINITION  4091710 BARC 10BOV Bos taurus cDNA clone 10BOV29_F21 5', mRNA
            sequence.
ACCESSION   CK952220
VERSION     CK952220.1  GI:45466600
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Sonstegard, T.S., Van Tassell, C.P., Matukumalli, L.K., Harhay,
            G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
  TITLE     Production of EST from cDNA libraries derived from immunologically
            activated bovine gut
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Tad S. Sonstegard
            Bovine Functional Genomics Laboratory
            Animal and Natural Resources Institute
            Bdlg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
            Tel: 3015048416
            Fax: 3015048414
            Email: tads@anri.barc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            0.000925 using options -trim_alt " -trim_fasta. Vector identified
            by cross_match using options -minmatch 12 -minscore 18
            Plate: 29  row: F  column: 21
            Seq primer: CCCAGTCACGACGTTGTAAAACG
            High quality sequence stop: 680.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /strain="Holstein"
                     /db_xref="taxon:9913"
                     /clone="10BOV29_F21"
                     /sex="Male"
                     /tissue_type="Pooled"
                     /dev_stage="Multiple"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="BARC 10BOV"
                     /note="Organ: Small Intestine; Vector: pAgen-1; Site_1:
                     EcoRV; Site_2: NotI; Equimolar amounts of mRNA extracted
                     from proximal jejunums of 18 and 21 wk old steers, and
                     distal ileums of 14 day old calves. proximal jejunum
                     exposed to C. oncophora for 3 and 6 weeks, and distal
                     ileum exposed to C. parvum for 7 days"
ORIGIN

Alignment Scores:
Pred. No.:                1.37e-116       Length:          680
Score:                    1085.00         Matches:         203
Percent Similarity:       96.0%           Conservative:    14
Best Local Similarity:    89.8%           Mismatches:      9
Query Match:              57.0%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-2 (1-362) x CK952220 (1-680)

Qy        120 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 139
              ||||||||||||||||||:::|||||||||   ||||||||||||||||||||||||||||
Db          1 AATATTGATGGATTGGACATGAACAATCTCCTAGAGAGGGCTCGAGGGGTGTGTTCCTAT 60

Qy        140 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSer 159
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db         61 TTAACTTTGTACAGCCCTGATGTGATATTTCTACAGGAAGTCATTCCCCCATATTATGCC 120

Qy        160 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 179
              ||||||||||||:::::::|||:::|||:::||||||||||||   ||||||||||||||
Db        121 TACCTAAAGAAGAAAGCAAGTAGTTACAAGATTATTACAGGTCGTGAAGAAGGATATTTC 180

Qy        180 ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 199
              ||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db        181 ACAGCTATAATGTTGAAGAAATCAAGAGTGAAGTTTAAAAGCCAAGAAATTATTCCTTTT 240

Qy        200 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 219
              |||:::|||:::|||||||||||||||||||||||||||||:::||||||||||||||||
Db        241 CCAAATACGCAAATGATGAGAAACCTTTTGTGTGTTCATGTGAGTGTGTCAGGAAATGAA 300

Qy        220 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
              ||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        301 CTTTGCCTTATGACTTCTCATTTGGAAAGCACCAGAGGGCACGCTAAGGAACGAATGAAT 360

Qy        240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 259
              ||||||||||||||||:::|||||||||||||||||||||   ||||||||||||||||||
Db        361 CAGTTGAAAATGGTTTTAGAGAAAATGCAAGAGGCTCCAGGATCAGCTACAGTTATATTT 420

Qy        260 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsn 279
              ||||||||||||||||||||||||:::||||||||:::||||||||||||||||||||||
Db        421 GCTGGAGATACGAATCTCAGGGATCAAGAAGTTACCAAATGTGGTGGTTTACCCAACAAC 480

Qy        280 IleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThr 299
              |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTTTGGATGTCTGGGAGTTTTTGGGCAAACCTAAGCATTGCCAGTATACATGGGATACA 540

Qy        300 GlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe 319
              ||||||||||||||||||||||||||   |||   ||||||||||||||||||||||||||
Db        541 CAAATGAACTCTAATCTTGGAATAGCTGCTAATTGTAAACTTCGTTTTGATCGAATATTT 600

Qy        320 PheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeu 339
              ||||||||||||||||   ||||||||||||||:::||||||||||||||||||||||||
Db        601 TTTAGAGCTGCAGCAGAAGGTGGCCACATTATTCCCCAAAGTTTGGACCTGCTTGGGCTG 660

Qy        340 GluLysLeuAspCysGly 345
              |||||||||||||||||
Db        661 GAAAAACTGGACTGTGGT 678

RESULT 63
BG719977
LOCUS       BG719977                  786 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602691335F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4823432 5',
            mRNA sequence.
ACCESSION   BG719977
VERSION     BG719977.1  GI:13999164
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Hominidae; Homo.
REFERENCE  1  (bases 1 to 786)
  AUTHORS  NIH-MGC http://mgc.nci.nih.gov/.
  TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL  Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
           Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM10733  row: f  column: 09
           High quality sequence stop: 784.
FEATURES             Location/Qualifiers
     source          1. .786
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4823432"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                3.9e-116        Length:         786
Score:                    1082.00         Matches:        213
Percent Similarity:       99.5%           Conservative:   0
Best Local Similarity:    99.5%           Mismatches:     1
Query Match:              56.9%           Indels:         1
DB:                       2               Gaps:           0

US-10-757-745-2 (1-362) x BG719977 (1-786)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 205

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 265

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 325

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 385

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 445

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAC 505

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 565

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 625

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 685

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 745

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 214
              ||||||||||||||||||||  ||||||||||||||||||
Db        746 AGTACCAAAATGATGAGAAACTT-TTATGTGTGCATGTGAAT 786

RESULT 64
BE892886
LOCUS       BE892886                995 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601435730F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920790 5',
            mRNA sequence.
ACCESSION   BE892886
VERSION     BE892886.1  GI:10353510
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 995)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9752  row: l  column: 07
            High quality sequence stop: 573.
FEATURES             Location/Qualifiers
     source          1. .995
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3920790"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

Alignment Scores:
Pred. No.:                7.52e-113       Length:         995
Score:                    1055.50         Matches:        232
Percent Similarity:       85.2%           Conservative:   10
Best Local Similarity:    81.7%           Mismatches:     31
Query Match:              55.5%           Indels:         12
DB:                       2               Gaps:           3

US-10-757-745-2 (1-362) x BE892886 (1-995)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
```

```
Db         30 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 89

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40

Db         90 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 149

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60

Db        150 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 209

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db        210 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 269

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        270 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 329

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        330 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 389

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        390 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 449

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        450 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 509

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db        510 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 569

Qy        181 AlaIleMet-LeuLysLysSerArgValLysLeu--LysSerGlnGluIleIlePro-Ph 199

Db        570 GCTATAATGTCGCACGAAATCAAGAGTGAAATTACCAAAGCCAAGGAGATCATTCCTTTT 629

Qy        199 eProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGl 219

Db        630 TCCAAGTACCAACATGATGAGCAACT--TTATGTGTGCATGTGACTGTGTCAGGGAATGA 687

Qy        219 uLeuCysLeuMetThrSerHisLeu---GluSerThrArgGlyHisAlaAlaGluArgMe 238

Db        688 GGCTGGCCTATGGGGTTCCCATTTGGGGGAAAGCACCCGACGGAATGCTGCG-------- 739

Qy        238 tAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIl 258

Db        740 ----AGCACGCACATGCATCACGCGTTCCACCACGTGGCTCTCCAAACCGCAACGATGGT 795

Qy        258 ePhe-AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProA 278

Db        796 CACAAGCGGGCCCCACAAGGATACAGTACAGGGCTATTCTC---TGGGGGAGCAATCCAA 852

Qy        278 sn 278

Db        853 AC 854
```

```
RESULT 65
CX756424
LOCUS       CX756424               799 bp    mRNA    linear   EST 24-JAN-2005
DEFINITION  AGENCOURT_41337854 NIH_MGC_281 Homo sapiens cDNA clone
            IMAGE:7779791 3', mRNA sequence.
ACCESSION   CX756424
VERSION     CX756424.1  GI:58053080
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 799)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Meri Firpo
             cDNA Library Preparation: Express Genomics
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15934  row: c  column: 21
            High quality sequence stop: 579.
FEATURES             Location/Qualifiers
     source          1. .799
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7779791"
                     /tissue_type="pluripotent cell line derived from
                     blastocyst inner cell mass"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_281"
                     /note="Organ: blastocyst; Vector: pExpress-1; Site_1:
                     EcoRV; Site_2: NotI; RNA obtained from pluripotent cell
                     line derived from blastocyst inner cell mass (cell line
                     HSF-6, NIH Registry designation UC06. Positive for OCT4
                     expression by rtPCR, positive for SSEA-3, SSEA-4,
                     Tra-1-81, Tra-1-60 by immunofluorescence. Negative for
                     SSEA-1 by immunofluorescence Passage 62. cDNA was primed
                     using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. Size-selection >1.25
                     kb resulted in an average insert size of 2.0 kb. This
                     primary library is normalized (non-normalized primary
                     library is NIH_MGC_280) and was constructed by Express
                     Genomics (Frederick, MD). Note: this is a Mammalian Gene
                     Collection library."
ORIGIN
```

```
Alignment Scores:
Pred. No.:                2.05e-112       Length:          799
Score:                    1050.50         Matches:         220
Percent Similarity:       93.6%           Conservative:    0
Best Local Similarity:    93.6%           Mismatches:      13
Query Match:              55.2%           Indels:          5
DB:                       8               Gaps:            2

US-10-757-745-2 (1-362) x CX756424 (1-799)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20

Db         18 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 77

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40

Db         78 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 137

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60

Db        138 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 197

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db        198 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 257

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        258 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 317
```

```
Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 377

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTTCC 617

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||   ||||||||||||   ||||||||||||||   |||||||||      |||
Db        618 AGTACCAA-ATGATGAGAAACCT-TTATGTGTGCATGTGAT-GTGTCAGGAATGAGCTTG 674

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
                      |||||||||   ||||||      |||
Db        675 ---CCTTATGAATCCCATTTGAGACAC---AGAGGCATGCTTGCG 713
```

RESULT 66
CB069952
LOCUS CB069952 621 bp mRNA linear EST 21-JAN-2003
DEFINITION is31e10.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6554035 5' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION CB069952
VERSION CB069952.1 GI:27814472
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 621)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other_ESTs: is31e10.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40UP from Gibco
High quality sequence stop: 476.
FEATURES Location/Qualifiers
source 1. .621
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6554035"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."
ORIGIN

Alignment Scores:

Pred. No.:	6.07e-112	Length:	621
Score:	1045.00	Matches:	202
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	54.9%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2 (1-362) x CB069952 (1-621)

```
Qy          3 LeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluProGluVal 22
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         13 TTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTG 72

Qy         23 LysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaVal 42
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 AAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTG 132

Qy         43 AlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPhe 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 GCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTC 192

Qy         63 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAG 252

Qy         83 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 ACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCA 312

Qy        103 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 TCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGAT 372

Qy        123 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 GGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTG 432

Qy        143 TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 TACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAG 492

Qy        163 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 AAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATA 552

Qy        183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 202
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        553 ATGTTGAAGAAATCAAGAGTGAAATTAAAAAGNCAAGAGATTATTCCTTTTCCCAGTACC 612

Qy        203 LysMetMet 205
              |||||||||
Db        613 AAAATGATG 621
```

RESULT 67
BI915865
LOCUS BI915865 915 bp mRNA linear EST 16-OCT-2001
DEFINITION 603184527F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5248495 5', mRNA sequence.
ACCESSION BI915865

```
VERSION     BI915865.1  GI:16179808
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11627  row: e  column: 08
            High quality sequence stop: 696.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5248495"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_121"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
                     fetal brains, female age 20 weeks, female age 24 weeks,
                     and male age 26 weeks.  Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed  upon
                     cloning).  Average insert size 1.7 kb, insert size range
                     0.7-3.5 kb.  Library is normalized and enriched for
                     full-length clones and  was constructed by C. Gruber
                     (Invitrogen).  Research Genetics  tracking code 017. Note:
                     this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.3e-111        Length:         915
Score:                  1044.50         Matches:        240
Percent Similarity:     83.5%           Conservative:   13
Best Local Similarity:  79.2%           Mismatches:     37
Query Match:            54.9%           Indels:         15
DB:                     3               Gaps:           5

US-10-757-745-2 (1-362) x BI915865 (1-915)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 67

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 127

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 187

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 247

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 307

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 367

Qy        121 IleAsp-GlyLeuAspLeuAsnAsnLeuSer---GluArgAlaArgGlyValCysSerTy 139
              |||||| |||||||||||||||||||||||    |||   :::   |||:::||||||||
Db        368 ATTGATCGGATTAGATCTAAACAATCTGTCAGACGAGGGCTCGACGGGGCTGTGTTCCTA 427

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAG 487

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIle-IleThrGlyHisGluGluGlyTyrP 179
              |||||||||||||||||||||||||||||||| ||||||||||:::|||||||||||
Db        488 CTACCTAAAGAAGAGATCAAGTAATTATGAGATTCATTACAGGTCATGACGAAGGATATT 547

Qy        179 heThrAlaIleMetLeuLys---LysSerArgValLysLeu---LysSerGlnGluIleI 197
              ||||||||         :::   ||||||||||||   |||       |||||||||||||
Db        548 TCACAGCCTATAACTGTCGAAGCAAATCAAGAGTGAA-CTCTACCACAGCCAAGAGATCA 606

Qy        197 leProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal-AsnValSer 216
              ||||||||||||||||   ||||||||||||||   |||||||||||| |||||||||
Db        607 TTCCTTTTCCAAGTACCACAATGATGAGAAACCTTCA-TGTGTGCATGTGCAATGTGTCA 665

Qy        217 GlyAsnGluLeuCysLeuMetThrSerHisLeuGlu-SerThrArg-GlyHisAla-Ala 235
              ||||||||||||  |||||||||||||||||||||| |||||||| |||||||| |||
Db        666 GGAAATGAGCTTCGCCTTATGACATCCCATTTGGAGCAGCACCAGAGGGCCATGCCTGCG 725

Qy        236 GluArgMetAsnGlnLeuLysMetVal-LeuLysLysMetGlnGluAlaProGluSerAl 255
              ||||||:::       |||:::   :::    :::   ||||||||||||||    :::
Db        726 GAACGACTGACTCCGTTACAATGGCTCCCCACGACACATGCAAGAGGCTCCAGCAACTCA 785

Qy        255 aThrVal------IlePheAlaGlyAspThrAsnLeuArgAsp---ArgGluValThrAr 272
                    :::          :::|||||||||   |||   |||   :::
Db        786 GCCTACACGCAATCTTTGCACGGAGAATACAAATCTACACGATCCACGAGGTCTCCCAGA 845

Qy        272 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 290
              |||   |||:::|||                              |||||||||
Db        846 ATTTCGTGGCTATCCCCACAACACTAGTGCCTCCCGGCGAATCGACGGCAAACCT 900

RESULT 68
CV023369
LOCUS       CV023369                 634 bp    mRNA    linear   EST 20-AUG-2004
DEFINITION  288 Full Length cDNA from the Mammalian Gene Collection Homo
            sapiens cDNA 5' similar to BC017553, mRNA sequence.
ACCESSION   CV023369
VERSION     CV023369.1  GI:51481130
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
            Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
            Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
            Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
            Vandenhaute,J., Cusick,M.E., Albala,J.S., Hill,D.E. and Vidal,M.
  TITLE     Human ORFeome Version 1.1: a Platform for Reverse Proteomics
  JOURNAL   Genome Res. (2004) In press
COMMENT     Contact: Vidal M
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 5739
            Email: Marc_Vidal@dfci.harvard.edu
            ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
            results from a PCR reaction using an MGC full-length cDNA as
            template DNA and ORF specific primers
            PCR PRimers
            FORWARD: ATGGAGTTGGGGAGTTGCCTGGA
            BACKWARD: TACAATATTATATCTAAGTTGCACAGAAG
```

```
           Insert Length: 634   Std Error: 32.00
           Plate: 11008  row: 05  column: F
           Seq primer: ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
           High quality sequence start: 97
           High quality sequence stop: 633
           POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .634
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="mixed"
                     /clone_lib="Full Length cDNA from the Mammalian Gene
                     Collection"
                     /note="Vector: mixed; The ORFs were PCR amplified from the
                     MGC (Mammalian Gene Collection) as of April 2004 and
                     cloned by recombinational Gateway cloning into pDONR223
                     Donor vector. Reference : MGC (Mammalian Gene Collection)
                     Program Team, Generation and Initial Analysis of more than
                     15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
                     2002, 99(26), 16899-16903"
ORIGIN

Alignment Scores:
Pred. No.:                8.23e-112      Length:          634
Score:                    1044.00        Matches:         203
Percent Similarity:       97.2%          Conservative:    2
Best Local Similarity:    96.2%          Mismatches:      6
Query Match:              54.9%          Indels:          0
DB:                       7              Gaps:            0

US-10-757-745-2 (1-362) x CV023369 (1-634)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysVal 211
                          :::   |||||||||:::
Db        601 GTACCAAATGATGAGAAACTTTTATTGTGCATG 633

RESULT 69
BU928818
LOCUS       BU928818                909 bp    mRNA    linear   EST 18-OCT-2002
DEFINITION  AGENCOURT_10422426 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:6654806
            5', mRNA sequence.
ACCESSION   BU928818
VERSION     BU928818.1  GI:24117548
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2907  row: a  column: 14
            High quality sequence stop: 543.
FEATURES             Location/Qualifiers
     source          1. .909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6654806"
                     /tissue_type="glioblastoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_57"
                     /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.55 kb (range 0.9-4.0 kb).  12/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
ORIGIN

Alignment Scores:
Pred. No.:                1.93e-111      Length:          909
Score:                    1043.00        Matches:         198
Percent Similarity:       99.0%          Conservative:    1
Best Local Similarity:    98.5%          Mismatches:      2
Query Match:              54.8%          Indels:          0
DB:                       5              Gaps:            0

US-10-757-745-2 (1-362) x BU928818 (1-909)

Qy        162 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCT 62

Qy        182 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 122

Qy        202 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCys 221
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 ACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGC 182

Qy        222 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 241
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTA 242

Qy        242 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 261
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 AAAATGGTTTTGAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGA 302

Qy        262 AspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleVal 281
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTG 362

Qy        282 AspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATG 422

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 AACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGA 482

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 341
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAANAA 542

Qy        342 LeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
                ||||||||||||||||||||||||||||||||||||||||||  :::||||||||||||
Db        543 CTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGNCACTTAGATATAATA 602

Qy        362 Leu 362
                |||
Db        603 TTG 605
```

```
RESULT 70
BM476771
LOCUS       BM476771               1110 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6481623 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5555311
            5', mRNA sequence.
ACCESSION   BM476771
VERSION     BM476771.1  GI:18525813
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1110)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12275  row: e  column: 08
            High quality sequence stop: 582.
FEATURES             Location/Qualifiers
     source          1. .1110
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5555311"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

Alignment Scores:
Pred. No.:              1.77e-110       Length:         1110
Score:                  1036.00         Matches:        195
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            54.5%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2 (1-362) x BM476771 (1-1110)

Qy        168 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 187
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 TATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCA 64

Qy        188 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 207
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 AGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAAC 124

Qy        208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 CTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTG 184

Qy        228 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 247
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 GAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAA 244

Qy        248 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 267
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGAT 304

Qy        268 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 CGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTG 364

Qy        288 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 307
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 GGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATA 424

Qy        308 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 327
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 ACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGA 484

Qy        328 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 347
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 CACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 544

Qy        348 ProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
                ||||||||||||||||||||||||||||||||||||||||||||
Db        545 CCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 589
```

```
RESULT 71
AL703449
LOCUS       AL703449               602 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686G0621_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686G0621 5', mRNA sequence.
ACCESSION   AL703449
VERSION     AL703449.1  GI:19686804
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   Ansorge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Ansorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
```

```
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686G0621) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .602
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686G0621"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

Alignment Scores:
Pred. No.:              3.37e-110      Length:         602
Score:                  1030.00        Matches:        199
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            54.2%          Indels:         0
DB:                     1              Gaps:           0

US-10-757-745-2 (1-362) x AL703449 (1-602)

Qy          2 GluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluProGlu 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 GAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAG 64

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCA 124

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 GTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTAC 184

Qy         62 PheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCC 244

Qy         82 LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSer 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 AAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGC 304

Qy        102 ProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 CCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATT 364

Qy        122 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 GATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCT 424

Qy        142 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTA 484

Qy        162 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 AAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCT 544

Qy        182 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 ATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 601
```

```
RESULT 72
BG533717
LOCUS       BG533717                793 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602562373F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4700059 5',
            mRNA sequence.
ACCESSION   BG533717
VERSION     BG533717.1  GI:13525257
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 793)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
            cDNA Library Preparation: CLONETECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1533  row: a  column: 20
            High quality sequence stop: 666.
FEATURES             Location/Qualifiers
     source          1. .793
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4700059"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_76"
                     /note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.85
                     kb (range 1.0-4.0 kb). 15/15 colonies contained  inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.19e-108      Length:         793
Score:                  1018.50        Matches:        229
Percent Similarity:     89.8%          Conservative:   8
Best Local Similarity:  86.7%          Mismatches:     18
Query Match:            53.5%          Indels:         11
DB:                     2              Gaps:           1

US-10-757-745-2 (1-362) x BG533717 (1-793)

Qy         55 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 74
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 75

Qy         75 ProGluThrIleSerGluProLysThrTyrValAspLeuThr-AsnGluGluThrThrAs 94
              |||||||||||||||||||||||||||||||||||||||||| :::|||||||||||||
Db         76 CCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAAGTGAAGAAACAACTGA 135

Qy         94 pSerThrThrSerLysIleSerProSerGlu-AspThrGlnGlnGluAsnGlySerMetP 114
              |||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||
Db        136 TTCCACCACTTCTAAAATCAGCCCATCTGTTAGATACTCAGCAAGAAAATGGCAGCATGT 195

Qy        114 heSerLeuIleThrTrpAsnIleAspGly-LeuAspLeuAsnAsnLeuSerGluArgAla 133
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
```

```
Db        196 TCTCTCTCATTACCTGGAATATTGATGGAGTTAGATCTAAACAATCTGTCAGAGAGGGCT 255

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153

Db        256 CGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTT 315

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173

Db        316 ATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 375

Qy        174 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 193

Db        376 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGCGTGAAATTAAAAAGC 435

Qy        194 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 213

Db        436 CAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 495

Qy        214 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 233

Db        496 AACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 555

Qy        234 AlaAlaGluArgMetAsnGlnLeuLys-MetValLeuLysLysMetGlnGluAlaProGl 253

Db        556 GCTGCGGAACGAATGACTCAGTTAAAACATGGTTTAGAAGAAAATGCAAGAGGT-CCAGA 614

Qy        253 uSerAlaThrValIlePheAlaGlyAsp-ThrAsnLeuArg---AspArgGluValThrA 272

Db        615 GTCAGCTACAGTTATATTTGCAGGAGAATACAAATCTAAGGGGATCGAAGAGGGTAACCA 674

Qy        272 rg-CysGlyGly-LeuProAsnAsnIleVal-AspValTrpGluPheLeuGlyLysProL 291

Db        675 GAATGTGGTGGTATTAAGGAAAAACATTGTGAGATGTCTGGGAAGA-TGGGGCAGACGTA 733

Qy        291 ysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 309

Db        734 AAAATTGGCAATATACATGGGATTAACGAAATGACATCTATCTGGGAATAAATGCT 789
```

```
RESULT 73
BI161201
LOCUS       BI161201                959 bp    mRNA    linear   EST 05-JUL-2001
DEFINITION  602865659F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5019639 5',
            mRNA sequence.
ACCESSION   BI161201
VERSION     BI161201.1  GI:14621202
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 959)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1835  row: e  column: 16
            High quality sequence stop: 551.
FEATURES             Location/Qualifiers
     source          1. .959
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5019639"
                     /tissue_type="epithelioid carcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_42"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                  1.61e-108       Length:          959
Score:                      1018.50         Matches:         231
Percent Similarity:         79.3%           Conservative:    11
Best Local Similarity:      75.7%           Mismatches:      50
Query Match:                53.5%           Indels:          14
DB:                         2               Gaps:            6

US-10-757-745-2 (1-362) x BI161201 (1-959)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20

Db         19 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 78

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40

Db         79 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 138

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60

Db        139 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 198

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db        199 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 258

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        259 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 318

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        319 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 378

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        379 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTTCCTACTTA 438

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        439 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 498

Qy        161 LeuLysLysArgSerSerAsnTyrGluIle-IleThrGlyHisGluGluGlyTyrPheTh 180

Db        499 CTAAAGAAGAGATCAAGTAATTATGAGATTCATTACAGGTCATGAAGAAGGATATTGCGC 558

Qy        180 rAlaIleMetLeuLysLys---SerArgValLysLeu---LysSerGln-GluIleIleP 198

Db        559 AGCTATAGATGTTGAAAGAAACTCACGAGTGAGACTTCAAAAAGGCCAGGGAGATTATTC 618

Qy        198 roPheProSerThrLys-MetMetArgAsnLeuLeuCysValHisVal-AsnValSerGl 217

Db        619 CTTTTCCAAGTACCAACAATGATGAGAAACCTTTATGTGTGCCATGTGAAATGCGTCAGG 678

Qy        217 yAsn---GluLeuCysLeuMetThrSerHisLeu---GluSerThrArgGlyHisAlaAl 235

Db        679 CACACTGAGGCTCCGCCTTATGACATCCCATTTTGGAGACGCACCAAGAGGGCCTGCTGC 738

Qy        235 a-GluArgMetAsnGlnLeuLys---MetValLeuLysLysMetGlnGluAlaProGluS 254
```

```
Db        739 GGGAACGAATGACTCAGGG-CGAACACTGGTCCTAAAGGAACAACAGCCAAGGAAGCGTC 797

Qy        254 erAlaThr---------ValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValT 271

Db        798 CCGGAGTTCAGGAACAAGCCATCACTCGCGGGAGCACACCAGTCAAAGGGAACGAAGACG 857

Qy        271 hrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProL 291

Db        858 GCACCAAAGGAGAGGACAACCGGCAGAAAACGGGGACGCAGAGAACAAGGGGCACACCCA 917

Qy        291 ysHis 292

Db        918 AGCAT 922

RESULT 74
CR753214
LOCUS       CR753214                652 bp    mRNA    linear   EST 01-SEP-2004
DEFINITION  DKFZp469I195_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469I195 5', mRNA sequence.
ACCESSION   CR753214
VERSION     CR753214.1  GI:51845629
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 652)
  AUTHORS   Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ansorge,W., Krieger,S., Regiert,T., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; lin, Germany. Please contact
            RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469I195
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp469I195"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="469 (synonym: pkid1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:              1.14e-108       Length:         652
Score:                  1017.50         Matches:        203
Percent Similarity:     96.2%           Conservative:   2
Best Local Similarity:  95.3%           Mismatches:     7
Query Match:            53.5%           Indels:         2
DB:                     7               Gaps:           0

US-10-757-745-2 (1-362) x CR753214 (1-652)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20

Db         18 ATGGAGTTGGGGAGTTTCCTGGAGGGCGGGAGGGAAGCGGCGGAGGAAGAGGGCGAGCCC 77

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40

Db         78 GAGGTGAAGAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCGAGCTGCGATACC 137

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60

Db        138 GCTGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 197

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db        198 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGCCCTGAAACCATCTCTGAG 257

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        258 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 317

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        318 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 377

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGln-GluValIleProProTyrTyrSerTy 160

Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGGAAGTTATTCCCCCATATTATAGCTA 497

Qy        160 rLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheTh 180

Db        498 CCTAAAGAAGAGTTCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGGTATTTCAC 557

Qy        180 rAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePr 200

Db        558 AGCTATCATGTTGAAGAATC--AGGAGTGAAATTAAAAGCCAAGAGATTTTTCCTTTTCC 615

Qy        200 oSerThrLysMetMetArgAsnLeuLeuCysValHis 212

Db        616 AAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 652

RESULT 75
BP262101
LOCUS       BP262101                581 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262101 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI05893, mRNA sequence.
ACCESSION   BP262101
VERSION     BP262101.1  GI:52177332
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI05893"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
```

Pred. No.:	1.08e-108	Length:	581
Score:	1017.00	Matches:	193
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	53.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2 (1-362) x BP262101 (1-581)

```
Qy        125 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 144
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGC 60

Qy        145 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGA 120

Qy        165 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTG 180

Qy        185 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATG 240

Qy        205 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACA 300

Qy        225 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTT 360

Qy        245 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAAT 420

Qy        265 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGG 480

Qy        285 GluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAAT 540

Qy        305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArg 317
              |||||||||||||||||||||||||||||||||||||||
Db        541 CTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGA 579
```

Search completed: March 11, 2006, 03:41:03
Job time : 9737.3 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 03:44:14 ; Search time 1047.64 Seconds
(without alignments)
797.263 Million cell updates/sec

Title: US-10-757-745-2
Perfect score: 1902
Sequence: 1 MELGSCLEGGREAAEEEGEP.....DCGRFPSDHWGLLCNLDIIL 362

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7673375 seqs, 1153648444 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170339_4330/app_query.fasta_1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1676_@runat_09032006_170339_4330 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	299.5	15.7	634	6	US-09-925-065A-783536	Sequence 783536,
	2	296	15.6	606	6	US-09-925-065A-783537	Sequence 783537,
c	3	196	10.3	952	8	US-10-750-185-25698	Sequence 25698, A
c	4	196	10.3	952	8	US-10-750-623-25698	Sequence 25698, A
	5	140	7.4	2928	7	US-10-330-773-785	Sequence 785, App
	6	128	6.7	1769	7	US-10-330-773-788	Sequence 788, App
	7	126	6.6	47040	7	US-10-330-773-784	Sequence 784, App
	8	123	6.5	49355	7	US-10-330-773-787	Sequence 787, App
	9	109.5	5.8	2080	12	US-11-140-625-2	Sequence 2, Appli
	10	108	5.7	1434	7	US-10-932-182A-76279	Sequence 76279, A
	11	108	5.7	1434	7	US-10-932-182A-76279	Sequence 76279, A
c	12	107	5.6	196716	9	US-11-114-798-53	Sequence 53, Appl
	13	106.5	5.6	2505	7	US-10-932-182A-52	Sequence 52, Appl
	14	106.5	5.6	2505	7	US-10-932-182A-52	Sequence 52, Appl
	15	105.5	5.5	1830	12	US-11-078-951-1	Sequence 1, Appli
	16	105.5	5.5	1939	12	US-11-078-951-10	Sequence 10, Appl
	17	105	5.5	70708	7	US-10-330-773-582	Sequence 582, App
c	18	104	5.5	190276	8	US-10-661-966-1	Sequence 1, Appli
	19	104	5.5	645179	8	US-10-995-561-13293	Sequence 13293, A
c	20	103	5.4	110469	7	US-10-330-773-314	Sequence 314, App
	21	102	5.4	1829	9	US-11-096-568A-5188	Sequence 5188, Ap
c	22	102	5.4	154548	12	US-11-121-086-33	Sequence 33, Appl
c	23	102	5.4	194534	7	US-10-330-773-458	Sequence 458, App
	24	102	5.4	1691140	12	US-11-091-018-1	Sequence 1, Appli
	25	101.5	5.3	81210	8	US-10-995-561-13295	Sequence 13295, A
	26	101.5	5.3	171732	12	US-11-121-086-98	Sequence 98, Appl
c	27	101	5.3	24861	7	US-10-330-773-248	Sequence 248, App
	28	101	5.3	69648	7	US-10-330-773-911	Sequence 911, App
	29	100	5.3	5238	9	US-11-072-175-59	Sequence 59, Appl
	30	100	5.3	176760	12	US-11-121-086-51	Sequence 51, Appl
	31	99.5	5.2	440684	7	US-10-330-773-114	Sequence 114, App
c	32	99	5.2	118544	7	US-10-330-773-76	Sequence 76, Appl
	33	99	5.2	1080000	8	US-10-928-446A-1	Sequence 1, Appli
	34	99	5.2	1080000	8	US-10-928-446A-181	Sequence 181, App
	35	99	5.2	1080000	8	US-10-928-446A-183	Sequence 183, App
	36	99	5.2	1080000	8	US-10-928-446A-185	Sequence 185, App
	37	99	5.2	1080000	8	US-10-928-446A-187	Sequence 187, App
	38	99	5.2	1080000	8	US-10-928-446A-189	Sequence 189, App
	39	99	5.2	1080000	8	US-10-928-446A-191	Sequence 191, App
	40	99	5.2	1080000	8	US-10-928-446A-193	Sequence 193, App
	41	99	5.2	1080000	8	US-10-928-446A-195	Sequence 195, App
	42	99	5.2	1080000	8	US-10-928-446A-197	Sequence 197, App
	43	99	5.2	1080000	8	US-10-928-446A-199	Sequence 199, App
	44	99	5.2	1080000	8	US-10-928-446A-201	Sequence 201, App
	45	98.5	5.2	2257	9	US-11-072-512-1677	Sequence 1677, Ap
	46	98.5	5.2	11185	12	US-11-122-329-100	Sequence 100, App
c	47	98.5	5.2	98862	12	US-11-121-086-76	Sequence 76, Appl
	48	98	5.2	561	6	US-09-925-065A-124743	Sequence 124743,
	49	98	5.2	633	6	US-09-925-065A-911490	Sequence 911490,
	50	98	5.2	639	6	US-09-925-065A-911273	Sequence 911273,
	51	98	5.2	3828	8	US-10-821-234-746	Sequence 746, App
c	52	98	5.2	184000	12	US-11-121-086-37	Sequence 37, Appl
c	53	98	5.2	191797	12	US-11-121-086-13	Sequence 13, Appl
	54	97.5	5.1	1466	9	US-11-096-568A-22794	Sequence 22794, A
	55	97.5	5.1	2203	12	US-11-167-856-11	Sequence 11, Appl
	56	97.5	5.1	2991	7	US-10-932-182A-76957	Sequence 76957, A
	57	97.5	5.1	2991	7	US-10-932-182A-76957	Sequence 76957, A
	58	97	5.1	791	6	US-09-925-065A-918150	Sequence 918150,
	59	97	5.1	791	6	US-09-925-065A-944994	Sequence 944994,
c	60	97	5.1	440684	7	US-10-330-773-114	Sequence 114, App
c	61	96.5	5.1	185393	12	US-11-121-086-101	Sequence 101, App
c	62	96.5	5.1	199130	8	US-10-995-561-13233	Sequence 13233, A
c	63	96	5.0	228835	7	US-10-330-773-398	Sequence 398, App
	64	95.5	5.0	981	9	US-11-096-568A-29182	Sequence 29182, A
c	65	95.5	5.0	74828	7	US-10-330-773-969	Sequence 969, App
	66	95.5	5.0	170837	12	US-11-121-086-97	Sequence 97, Appl
	67	95	5.0	165883	12	US-11-112-908-18	Sequence 18, Appl
c	68	95	5.0	645179	8	US-10-995-561-13293	Sequence 13293, A
	69	93.5	4.9	4251	7	US-10-932-182A-3111	Sequence 3111, Ap
	70	93.5	4.9	4251	7	US-10-932-182A-3111	Sequence 3111, Ap
	71	93.5	4.9	171423	12	US-11-121-086-85	Sequence 85, Appl
	72	93.5	4.9	215308	12	US-11-121-086-77	Sequence 77, Appl
c	73	93	4.9	597	6	US-09-925-065A-871294	Sequence 871294,
	74	93	4.9	180574	12	US-11-121-086-70	Sequence 70, Appl
	75	93	4.9	212805	12	US-11-112-908-19	Sequence 19, Appl

ALIGNMENTS

```
RESULT 1
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783536

Alignment Scores:
Pred. No.:                1.97e-22        Length:          634
Score:                    299.50          Matches:         63
Percent Similarity:       83.3%           Conservative:    2
Best Local Similarity:    80.8%           Mismatches:      10
Query Match:              15.7%           Indels:          3
DB:                       6               Gaps:            1

US-10-757-745-2 (1-362) x US-09-925-065A-783536 (1-634)

Qy          85 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 104
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         265 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 324

Qy         105 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         325 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 384

Qy         125 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         385 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTAAGTA 444

Qy         145 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 162
               :::::|||          |||           |||   |||   |||
Db         445 TTATCACTTCTATTTAATAGTAATGTG---------TCTTATGTATATACAAAG 489

RESULT 2
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783537

Alignment Scores:
Pred. No.:                4.44e-22        Length:          606
Score:                    296.00          Matches:         58
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              15.6%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2 (1-362) x US-09-925-065A-783537 (1-606)

Qy          85 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 104
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         414 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 473

Qy         105 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         474 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 533

Qy         125 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         534 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTG 587

RESULT 3
US-10-750-185-25698/c
; Sequence 25698, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25698
;   LENGTH: 952
;   TYPE: DNA
;   ORGANISM: Bovine  19866881510436
US-10-750-185-25698

Alignment Scores:
Pred. No.:                5e-11           Length:          952
Score:                    196.00          Matches:         38
Percent Similarity:       91.3%           Conservative:    4
Best Local Similarity:    82.6%           Mismatches:      4
Query Match:              10.3%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2 (1-362) x US-10-750-185-25698 (1-952)

Qy         169 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 188
               :::|||       |||   ||||||||||||||||||||||||||||||||||||||||
```

```
Db        800 CAAATACCTGAAGGTCGTGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 741

Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
              ||||||   ||||||||||||||||||||||||||||:::|||:::|||||||||||||||
Db        740 GTGAAGTTTAAAAGCCAAGAAATTATTCCTTTTCCAAATACGCAAATGATGAGAAACCTT 681

Qy        209 LeuCysValHisValAsn 214
              ||||||||||||||:::
Db        680 TTGTGTGTTCATGTAAGT 663

RESULT 4
US-10-750-623-25698/c
; Sequence 25698, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25698
;   LENGTH: 952
;   TYPE: DNA
;   ORGANISM: Bovine  19866881510436
US-10-750-623-25698

Alignment Scores:
Pred. No.:                5e-11          Length:          952
Score:                    196.00         Matches:         38
Percent Similarity:       91.3%          Conservative:    4
Best Local Similarity:    82.6%          Mismatches:      4
Query Match:              10.3%          Indels:          0
DB:                       8              Gaps:            0

US-10-757-745-2 (1-362) x US-10-750-623-25698 (1-952)

Qy        169 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 188
              :::|||      |||  ||||||||||||||||||||||||||||||||||||||||||||
Db        800 CAAATACCTGAAGGTCGTGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 741

Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
              ||||||   ||||||||||||||||||||||||||:::|||:::|||||||||||||||
Db        740 GTGAAGTTTAAAAGCCAAGAAATTATTCCTTTTCCAAATACGCAAATGATGAGAAACCTT 681

Qy        209 LeuCysValHisValAsn 214
              ||||||||||||||:::
Db        680 TTGTGTGTTCATGTAAGT 663

RESULT 5
US-10-330-773-785
; Sequence 785, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785
;   LENGTH: 2928
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2928)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-785

Alignment Scores:
Pred. No.:                0.000236       Length:          2928
Score:                    140.00         Matches:         89
Percent Similarity:       35.0%          Conservative:    55
Best Local Similarity:    21.7%          Mismatches:      163
Query Match:              7.4%           Indels:          104
DB:                       7              Gaps:            19

US-10-757-745-2 (1-362) x US-10-330-773-785 (1-2928)

Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArg------------------ 25
              :::||||||        |||          :::   |||
Db        666 CAGGCGGCGGCGGCCTCGGGCGCCGCGAGGTCGCGTCCCCGAACAGTGAGTTCCATGGGA 725

Qy         26 ------------ArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
                        ||||||                 :::|||   |||      ||||||
Db        726 AACGGCACCAGTCGACTCTACAGTGCTCTCGCCAAGACCGTCAACAGCAGTGCAGCTGCT 785

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
                    :::|||          |||   |||                          :::
Db        786 CAGCACCCGGAGTACTTGGTGTCAACTGACCCCGAA--------------------CAT 824

Qy         62 PheGluPro------ProValGlu--------------GluSerAlaLeuGluArgArg 74
              ||||||      |||  |||                  :::   |||      |||
Db        825 CTGGAGCCCATCGATCCTAAGGAACTTCTGGAGGAATGCAGGGCTGTTCTGCACACTCGG 884

Qy         75 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 94
              |||              :::   :::||||||||||                |||||
Db        885 CCACCCCGCTACCAG------CGGGATTTTGTGGACCTG------------AGGACAGAT 926

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114
                 ::::::|||                                          |||
Db        927 TGCTCCAGCAGCCAC-----------------------------------TCTCCCATT 950

Qy        115 SerLeuIleThrTrpAsnIle------------------AspGlyLeuAsp--------- 125
                 ::::::  |||||||||                ::|||   |||
Db        951 CGCGTCATGCAGTGGAACATCCTCGCCCAAGCTCTCGGAGAAGGCAAAGACAACTTTGTG 1010

Qy        126 --------------LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
                           ||||||          :::         :::
Db       1011 CAGTGCCCCGTGGAAGCGCTCAAATGGGAAGAGAGGAAGTGCCTGATCCTGGAGGAGATC 1070

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              |||   ||||||::::::   ||||||||||             |||:::
Db       1071 CTGGCTTACCAGCCAGACATACTGTGCCTCCAGGAAGTG---GACCACTACTTTGACACC 1127

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe--- 179
                 :::          |||                |||::::::      :::|||
Db       1128 TTCCAGCCACTCCTCAGTAGA-----------CTGGGCTACCAAGGCACGTTTTTCCCC 1175

Qy        180 ------------------------------------------------ThrAlaIle 182
                                                               |||:::
Db       1176 AAGCCCTGGTCACCATGTCTAGATGTGGAACACAACAACGGTCCAGATGGCTGTGCCTTA 1235

Qy        183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 202
                  ::::::|||   ||||||   |||        |||        :::
Db       1236 TTTTTTCTCCAAAACCGATTCAAGCTTATCAGCAGCACCAATATTAGGCTGACAGCCATG 1295

Qy        203 LysMetMetArgAsnLeuLeuCysVal-----------HisValAsnValSerGlyAsn 218
                 :::      |||   :::   :::                        ||||||
Db       1296 ACCCTAAAAACCAATCAGGTGGCCATCGCACAGACCCTGGAGTGCAAGGAGTCTGGGCGA 1355

Qy        219 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
              :::   |||:::      :::||||||::::::      |||             |||
Db       1356 CAGTTCTGCATTGCTGTCACCCACTTAAAAGCCCGCACTGGCTGGGAGCGGTTCCGGTCA 1415
```

```
Qy        239 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal--- 257
Db       1416 GCTCAGGGCTGTGACCTCCTCCAGAACCTGCAGAACATCACCCAGGGAGCAAAGATCCCC 1475

Qy        258 ---IlePheAlaGlyAspThrAsnLeuArgAspArgGlu------ValThrArgCysGly 274
Db       1476 CTGATCGTCTGCGGGGACTTCAACGCAGAGCCAACCGAAGAGGTCTACAAACACTTTGCG 1535

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------GlyLysProLys 291
Db       1536 TCCTCCAGCCTCAACCTCAACAGCGCCTACAAGCTGCTGAGTCCCGATGGACAGTCGGAG 1595

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
Db       1596 CCTCCCTACACAACCTGG-----------------AAGATCCGGACCTCAGGCGAGTGT 1637

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
Db       1638 CGCCACACGCTGGACTATATCTGGTACTCCAGACATGCTCTGAGT---------GTCACG 1688

Qy        332 ArgSerLeuAspLeuLeu--------------GlyLeuGluLysLeuAspCysGlyArg 346
Db       1689 TCCGCCCTGGATCTACTCACTGAAGAACAGATTGGGCCCAACCGGCTCCCATCCTTCCAT 1748

Qy        347 PheProSerAspHisTrpGlyLeuLeuCysAsn 357
Db       1749 TACCCCTCGGACCACCTGTCCCTGGTGTGTGAC 1781
```

RESULT 6
US-10-330-773-788
; Sequence 788, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 788
; LENGTH: 1769
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-788

Alignment Scores:

Pred. No.:	0.00244	Length:	1769
Score:	128.00	Matches:	83
Percent Similarity:	33.0%	Conservative:	55
Best Local Similarity:	19.9%	Mismatches:	164
Query Match:	6.7%	Indels:	116
DB:	7	Gaps:	18

US-10-757-745-2 (1-362) x US-10-330-773-788 (1-1769)

```
Qy         20 ProGluValLysLysArgArgLeuLeuCysValGluPheAla---------SerValAla 36
Db        308 CCCAGGCCCGCATCCCCCCGGCTGCTGGCGGCGGCCTCGGCGGCCTCGGGCGCCGCGAGG 367

Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
Db        368 TCGTGTTCCCGAACAGTG------TGTTCCATGGGAACCGGTACAAGCAGACTCTATAGT 421

Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
Db        422 GCTCTC------------GCCAAGACACTGAACAGCAGCGCTGCCTCCCAGCACCCAGAG 469

Qy         77 ThrIleSerGluPro---------------------------------------------- 81
Db        470 TATTTGGTGTCACCTGACCCAGAGCATCTGGAGCCCATTGATCCTAAAGAGCTTCTTGAG 529

Qy         82 ----------------------------------------LysThrTyrValAspLeu 87
Db        530 GAATGCAGGGCCGTCCTGCACACCCGACCTCCCCGGTTCCAGAGGGATTTTGTGGATCTG 589

Qy         88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 107
Db        590 AGGACAGAT------TGCCCTAGTACCCACCCACCTATC---------------------- 622

Qy        108 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle------------------ 121
Db        623 ---------------------AGGGTTATGCAATGGAACATCCTCGCCCAAGCTCTTGGA 661

Qy        122 AspGlyLeuAsp-----------------------LeuAsnAsnLeuSerGluArgAla 133
Db        662 GAAGGCAAAGACAACTTTGTACAGTGCCCTGTTGAAGCACTCAAATGGGAAGAAAGGAAA 721

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
Db        722 TGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTCCAAGAGGTG 781

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
Db        782 ---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA------------CTAGGC 826

Qy        174 HisGluGluGlyTyrPhe----------------------------------------- 179
Db        827 TATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAACACAACAAT 886

Qy        180 ------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
Db        887 GGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTCAACAGTGCC 946

Qy        196 IleIleProProProSerThrLysMetMetArgAsnLeuLeuCysVal------------ 211
Db        947 AATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCACAGACCCTG 1006

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
Db       1007 GAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAAGCACGCACT 1066

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
Db       1067 GGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTGCAAAACATC 1126

Qy        252 ProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
Db       1127 ACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAGCCAACAGAA 1186

Qy        270 ------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
Db       1187 GAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTACAAGCTGCTG 1246

Qy        288 ---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304
Db       1247 AGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG-----------------AAG 1288

Qy        305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 324
Db       1289 ATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCTAAACATGCT 1348

Qy        325 GluGluGlyHisIleIleProArgSerLeuAspLeuLeu---------------GlyLeu 339
Db       1349 CTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAGATTGGACCC 1399

Qy        340 GluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsn 357
Db       1400 AACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGTGAC 1453
```

RESULT 7
US-10-330-773-784
; Sequence 784, Application US/10330773
; Publication No. US20060040262A1

```
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 784
;   LENGTH: 47040
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(47040)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-784

Alignment Scores:
Pred. No.:                0.269          Length:         47040
Score:                    126.00         Matches:        57
Percent Similarity:       37.9%          Conservative:   42
Best Local Similarity:    21.8%          Mismatches:     114
Query Match:              6.6%           Indels:         48
DB:                       7              Gaps:           10

US-10-757-745-2 (1-362) x US-10-330-773-784 (1-47040)

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ||||||          :::        :::      |||   ||||||::::::   |||
Db      35153 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 35212

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              |||||||||         |||:::        :::        |||
Db      35213 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 35260

Qy        171 IleThrGlyHisGluGluGlyTyrPhe-------------------------------- 179
                   |||::::::     :::|||
Db      35261 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 35317

Qy        180 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
                                  |||:::        ::::::|||   ||||||
Db      35318 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 35377

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 211
              |||         |||        :::      :::      |||   :::   :::
Db      35378 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 35437

Qy        212 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
                       ||||||   :::   |||:::      :::|||||:::
Db      35438 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 35497

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
              :::      |||        ||        |||         :::|||:::   :::
Db      35498 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 35557

Qy        249 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 266
              |||         :::   |||   :::      |||      ||||||   |||
Db      35558 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 35617

Qy        267 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
                  |||                                 |||:::         :::
Db      35618 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 35677

Qy        285 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
              :::  |||           |||:::  :::          ||||||
Db      35678 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 35725

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321
                    :::      :::      |||:::          |||   |||::::::
Db      35726 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 35779

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu------------ 337
                  |||          :::      :::|||||||||||||
Db      35780 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 35830

Qy        338 ---GlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356
                 |||       :::|||             :::||||||||||||      |||:::|||
Db      35831 ATTGGGCCCAACCGGCTCCCATCCTTCCATTACCCCTCGGACCACCTGTCCCTGGTGTGT 35890

Qy        357 Asn 357
                 :::
Db      35891 GAC 35893

RESULT 8
US-10-330-773-787
; Sequence 787, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 787
;   LENGTH: 49355
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(49355)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-787

Alignment Scores:
Pred. No.:                0.603          Length:         49355
Score:                    123.00         Matches:        56
Percent Similarity:       37.9%          Conservative:   43
Best Local Similarity:    21.5%          Mismatches:     114
Query Match:              6.5%           Indels:         48
DB:                       7              Gaps:           10

US-10-757-745-2 (1-362) x US-10-330-773-787 (1-49355)

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ||||||          :::        :::      |||   ||||||::::::   |||
Db      38299 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 38358

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              |||||||||         |||:::        :::        |||
Db      38359 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 38406

Qy        171 IleThrGlyHisGluGluGlyTyrPhe-------------------------------- 179
                   |||::::::     :::|||
Db      38407 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 38463

Qy        180 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
                                  |||:::        ::::::|||   ||||||
Db      38464 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 38523

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 211
              :::         |||        :::      :::      |||   :::   :::
Db      38524 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 38583

Qy        212 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
                       ||||||   :::   |||:::      :::|||||:::
Db      38584 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 38643

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
              :::      |||        |||        |||         :::|||:::   :::
Db      38644 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 38703
```

```
Qy        249 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 266
Db      38704 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 38763

Qy        267 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
Db      38764 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 38823

Qy        285 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
Db      38824 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 38871

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321
Db      38872 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 38925

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu------------ 337
Db      38926 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 38976

Qy        338 ---GlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356
Db      38977 ATTGGACCCAACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGT 39036

Qy        357 Asn 357
Db      39037 GAC 39039
```

RESULT 9
US-11-140-625-2
; Sequence 2, Application US/11140625
; Publication No. US20060026706A1
; GENERAL INFORMATION:
; APPLICANT: Van Wezel, Gilles P.
; APPLICANT: Vijgenboom, Erik
; TITLE OF INVENTION: A method for marker-less integration of a sequence of interest
; TITLE OF INVENTION: into the genome of a cell
; FILE REFERENCE: P62377SUS00
; CURRENT APPLICATION NUMBER: US/11/140,625
; CURRENT FILING DATE: 2005-05-27
; PRIOR APPLICATION NUMBER: EP 02080000.9
; PRIOR FILING DATE: 2002-11-28
; PRIOR APPLICATION NUMBER: PCT/NL2003/000835
; PRIOR FILING DATE: 2003-11-27
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 2080
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Sequence of the glkA region amplified from the S. coelicolor M145
; OTHER INFORMATION: genome
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2080)
; OTHER INFORMATION: Sequence starts at nucleotide position 37 and ends at nucleotide
; OTHER INFORMATION: position 2116
US-11-140-625-2

Alignment Scores:

Pred. No.:	0.299	Length:	2080
Score:	109.50	Matches:	61
Percent Similarity:	36.1%	Conservative:	53
Best Local Similarity:	19.3%	Mismatches:	129
Query Match:	5.8%	Indels:	73
DB:	12	Gaps:	14

US-10-757-745-2 (1-362) x US-11-140-625-2 (1-2080)

```
Qy         75 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 94
Db        991 CCGGCAGTCGCCCAGGAACCCCGCGCC-----------CCGACCCATGAGGCGCCTGAT 1038

Qy         95 -------------------------SerThrThrSerLysIleSerProSerGluAsp 105
Db      1039 CATGCGTTCCCGTAAATTGACGGCATGGCAACATCCCTCCCGGCGCTCCCCAACTCCCGC 1098

Qy        106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
Db      1099 ACGGAACCCGACGGTTCCGCGGTCATCCGCGTCCTGAGCTACAACATC------------ 1146

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
Db      1147 ---CGCTCACTGCGCGACGACACCGACGCCCTCGCCCGCGTCATCAAGGCCTGCGCGCCC 1203

Qy        146 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys------ 163
Db      1204 GACCTGGTCCTCCTCCAGGAGGCC---CCCCGCTTCTTCCGCTGGCGCAAGAAGATCACC 1260

Qy        164 ---------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
Db      1261 CGGCTCGCCGCCGCCGGCGATCTGGTCCTGCTCTCCGGCGGCGGAACCGCCGCCGGCCCC 1320

Qy        181 AlaIleMetLeuLys---LysSerArgValLysLeuLysSerGlnGluIleIleProPhe 199
Db      1321 GCGCTGCTCTGCTCCCTGCGGGCCACCGTGGAGCGCACCGAGGACGTGCTGCTCCCGCTC 1380

Qy        200 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 219
Db      1381 ACCCCCGGCCGGCACCGCCGGGGCATCGCCGCCGCCGTGGTCCGGATCGGCGGCGCCCGC 1440

Qy        220 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 239
Db      1441 ATCGGCGTACTCAGCTCCCACCTGTCCCTG---------GACGCGGACGAGCGCCACGAA 1491

Qy        240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 259
Db      1492 CAGGCCGGGCTGCTCCTCGACCACCTG---GCCGCCCTCGGCGTGAAGCACGCCGTGGCG 1548

Qy        260 AlaGlyAspThrAsnLeuArg------AspArgGluValThrArgCysGlyGlyLeuPro 277
Db      1549 GGCGGCGACCTGAACGAGCGCCCCGGCGGCCGCACCTTCCGGCGGCTCGGC--------- 1599

Qy        278 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db      1600 GAAGGGCTGCGCGACTGCTGG------ACCGCCGCGCCCTGGGGCGGCGAGTACACCTTC 1653

Qy        298 AspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArg 317
Db      1654 CCC-------------------------GCGACCGCCCCGGACCGCCGGATCGACGCG 1686

Qy        318 IlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu 337
Db      1687 GTCTTCGTCACCGAG---------------------------------------------- 1701

Qy        338 GlyLeuGluLysLeuAspCysGly---------------------------------- 345
Db      1702 GGCATCGAGGTGCTGGGCTGCGGAGTGCCGTCGGGACTGGCCGGGGTGGCCGAGGACGAC 1761

Qy        346 ---ArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
Db      1762 CTGAGGGCGGCCACGGACCACCTTCCGGTCCTGACCGCCCTCCGCGTA 1809
```

RESULT 10
US-10-932-182A-76279
; Sequence 76279, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043

```
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76279
;   LENGTH: 1434
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-76279

Alignment Scores:
Pred. No.:                0.27          Length:          1434
Score:                    108.00        Matches:         78
Percent Similarity:       35.2%         Conservative:    47
Best Local Similarity:    22.0%         Mismatches:      124
Query Match:              5.7%          Indels:          106
DB:                       7             Gaps:            20

US-10-757-745-2 (1-362) x US-10-932-182A-76279 (1-1434)

Qy         90 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 109
              :::|||          |||                ::::::   :::             :::
Db         37 AAGGAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCAATTATGTCCGAT 96

Qy        110 AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGlyLeuAspLeu--- 126
              |||          :::|||:::      :::|||:::|||      ||||||
Db         97 AATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGGTTAAAATACGTC 156

Qy        127 ---AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
                 :::     ||||||  |||  :::       ||||||   :::|||
Db        157 TCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCGGGCCACTCAACG 216

Qy        145 ------------------------------------------------------------ 145

Db        217 CTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGCAATGAAAACGAAGAT 276

Qy        146 ---AspValIlePheLeuGlnGluVal---IleProProTyrTyrSerTyrLeuLysLys 163
                 |||||||||   |||||||||:::               :::   ||||||
Db        277 TACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAGTATCTAGCTTCT 336

Qy        164 Arg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr------ 180
                 :::|||   |||                          |||      |||
Db        337 GCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATTCTGACGGGGCCT 396

Qy        181 ------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198
                    ||||||:::          |||:::|||   :::::|||       :::
Db        397 GGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACCTTTCTTTACCGG 447

Qy        199 PhePro---------------------------------SerThrLysMetMetArgAsn 207
                 ||||||                                 :::   ::::::
Db        448 TTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTACGTAGGGAAATCT 507

Qy        208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
              :::          :::|||              :::   :::|||   ||||||:::
Db        508 ATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATGATCAGTCACATG 567

Qy        228 Glu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln------------ 240
                       :::   :::|||   ||||||       ::::::
Db        568 CACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGATCTTGTCAGGCC 627

Qy        241 ------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258
                    ::::::|||      ::::::|||                  ||||||
Db        628 TGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGTTATGCGGTGATT 681

Qy        259 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 278
                 ||||||   |||                 :::|||   |||   ||||||:::
Db        682 GTGGTGGGTGACTTAAAC----------------TCCAGACCGGGCTCACTGCCCCAC 723

Qy        279 --------------------AsnIleValAspValTrpGluPheLeu---GlyLysPro 290
                             :::||||||   ||||||   |||   ||||||
Db        724 AAATTCCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTGCATGGGAAGCAA 783

Qy        291 ----------------------------------------------LysHisCysGln--- 294

Db        784 GACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTCAAGGGCTGTACCACG 843

Qy        295 ---------------TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 309
                            ||||||   |||      :::
Db        844 TGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAGCCCGAT-----------GAG 891

Qy        310 AlaCysLysLeuArgPhe---------AspArgIlePhePheArgAlaAlaAlaGluGlu 326
              ||||||:::|||   :::         |||      |||   :::
Db        892 GCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACAGTAGACGCAGGT 951

Qy        327 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 346
                          |||                  :::      |||||||||
Db        952 GTCAGGTTCACTGAACGG-----------------ATCCCTCACCTGGACTGCAGT--- 990

Qy        347 PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
                 |||||||||:::          |||   |||:::|||:::
Db        991 ---GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 1032

RESULT 11
US-10-932-182A-76279
; Sequence 76279, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76279
;   LENGTH: 1434
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-76279

Alignment Scores:
Pred. No.:                0.27          Length:          1434
Score:                    108.00        Matches:         78
Percent Similarity:       35.2%         Conservative:    47
Best Local Similarity:    22.0%         Mismatches:      124
Query Match:              5.7%          Indels:          106
DB:                       7             Gaps:            20

US-10-757-745-2 (1-362) x US-10-932-182A-76279 (1-1434)

Qy         90 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 109
              :::|||          |||                ::::::   :::             :::
Db         37 AAGGAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCAATTATGTCCGAT 96

Qy        110 AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGlyLeuAspLeu--- 126
              |||          :::|||:::      :::|||:::|||      ||||||
Db         97 AATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGGTTAAAATACGTC 156

Qy        127 ---AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
                 :::     ||||||  |||  :::       ||||||   :::|||
Db        157 TCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCGGGCCACTCAACG 216

Qy        145 ------------------------------------------------------------ 145

Db        217 CTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGCAATGAAAACGAAGAT 276

Qy        146 ---AspValIlePheLeuGlnGluVal---IleProProTyrTyrSerTyrLeuLysLys 163
                 |||||||||   |||||||||:::               :::   ||||||
Db        277 TACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAGTATCTAGCTTCT 336

Qy        164 Arg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr------ 180
                 :::|||   |||                          |||      |||
```

```
Db        337 GCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATTCTGACGGGGCCT 396
Qy        181 ------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198
Db        397 GGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACCTTTCTTTACCGG 447
Qy        199 PhePro----------------------------------SerThrLysMetMetArgAsn 207
Db        448 TTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTACGTAGGGAAATCT 507
Qy        208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
Db        508 ATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATGATCAGTCACATG 567
Qy        228 Glu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln------------ 240
Db        568 CACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGATCTTGTCAGGCC 627
Qy        241 ------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258
Db        628 TGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGTTATGCGGTGATT 681
Qy        259 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 278
Db        682 GTGGTGGGTGACTTAAAC----------------TCCAGACCGGGCTCACTGCCCCAC 723
Qy        279 --------------------AsnIleValAspValTrpGluPheLeu---GlyLysPro 290
Db        724 AAATTCCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTGCATGGGAAGCAA 783
Qy        291 ---------------------------------------------LysHisCysGln--- 294
Db        784 GACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTCAAGGGCTGTACCACG 843
Qy        295 --------------TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 309
Db        844 TGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAGCCCGAT------------GAG 891
Qy        310 AlaCysLysLeuArgPhe---------AspArgIlePhePheArgAlaAlaAlaGluGlu 326
Db        892 GCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACAGTAGACGCAGGT 951
Qy        327 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 346
Db        952 GTCAGGTTCACTGAACGG----------------ATCCCTCACCTGGACTGCAGT--- 990
Qy        347 PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
Db        991 ---GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 1032
```

RESULT 12
US-11-114-798-53/c
; Sequence 53, Application US/11114798
; Publication No. US20060035246A1
; GENERAL INFORMATION:
; APPLICANT: WU, RINA
; APPLICANT: MARQUEZ, ABBEY
; TITLE OF INVENTION: CHROMOGENIC IN SITU HYBRIDIZATION METHODS, KITS, AND
; TITLE OF INVENTION: COMPOSITIONS
; FILE REFERENCE: 0618.011.0004
; CURRENT APPLICATION NUMBER: US/11/114,798
; CURRENT FILING DATE: 2005-04-26
; PRIOR APPLICATION NUMBER: 10/173,525
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/952,851
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 53
; LENGTH: 196716
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: nucleotide sequence
US-11-114-798-53

Alignment Scores:

Pred. No.:	189	Length:	196716
Score:	107.00	Matches:	55
Percent Similarity:	44.0%	Conservative:	44
Best Local Similarity:	24.4%	Mismatches:	92
Query Match:	5.6%	Indels:	35
DB:	9	Gaps:	9

US-10-757-745-2 (1-362) x US-11-114-798-53 (1-196716)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db     194444 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 194386
Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db     194385 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 194332
Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db     194331 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCGTCAGTGTGCTGTATT 194275
Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db     194274 CAGGAA----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 194233
Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db     194232 ATAAAAGGATGGAGGAAGAGCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 194173
Qy        191 LeuLysSerGlnGluIleIleProProProSerThrLysMetMetArgAsnLeuLeuCys 210
Db     194172 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 194113
Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
Db     194112 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 194053
Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     194052 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 194008
Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db     194007 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 193951
Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db     193950 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 193891
Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     193890 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 193834
Qy        293 CysGlnTyrThrTrp 297
Db     193833 ACAGAATATACATTT 193819
```

RESULT 13
US-10-932-182A-52
; Sequence 52, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS

```
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 52
;   LENGTH: 2505
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-52

Alignment Scores:
Pred. No.:                0.8          Length:          2505
Score:                    106.50       Matches:         92
Percent Similarity:       33.5%        Conservative:    72
Best Local Similarity:    18.8%        Mismatches:      143
Query Match:              5.6%         Indels:          183
DB:                       7            Gaps:            27

US-10-757-745-2 (1-362) x US-10-932-182A-52 (1-2505)

Qy         20 ProGluValLysLys---------ArgArgLeuLeuCysValGluPheAlaSerValAla 36
Db       1111 CCGGAGATTAAGCGTTTGGACAACCTGCGTGTCTTGGACTTGTCCCATAATAGGTTGACC 1170

Qy         37 SerCysAspAlaAlaValAlaGlnCysPhe---------------------------- 46
Db       1171 GCGCTACCAGCGGAACTGGGCTCCTGCTTCCAGTTGAAGTACCTGTATTTCTTCGATAAC 1230

Qy         47 ---LeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProPro 65
Db       1231 ATGGTCACCACGCTACCATGGGAATTCGGCAATCTGTACAACTTGCAGTTT---CTCGGT 1287

Qy         66 ValGluGluSerAlaLeuGluArg---------------------------------- 73
Db       1288 GTAGAGGGCAATCCCTTGGAGAGGCAGCTTTTGAAGATCCTCACAGAAAAATCTGTCACG 1347

Qy         74 --------------------------ArgProGlu---ThrIleSerGluProLysThr 83
Db       1348 GGACTGATCTTCTACCTAAGAGACAACAGACCGGAGATCACCTTACCTCACGAGCGCAGG 1407

Qy         84 TyrValAspLeuThrAsn--------------------------------GluGluThr 92
Db       1408 TTCATCGAGATCAATACTGATGGGGAACCACAAAGAGAATATGACTCCTTGCAACAATCG 1467

Qy         93 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 112
Db       1468 ACGGAACACTTGACCGCGGATCTGGCCAAGAGATCGTTCACCGTGCTATCGTACAATACT 1527

Qy        113 MetPheSer-----------------------------------------LeuIleThr 118
Db       1528 CTGTGCCAACACTACGCCACTCCCAAAATGTACCGTTACACACCATCATGGGCATTGAGT 1587

Qy        119 TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 138
Db       1588 TGGGAT---------TACAGACGTAACAAGCTAAAGGAACAG------------------ 1620

Qy        139 TyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 158
Db       1621 ---ATTTTGTCATACAGTAGTGACGTGTTGTGTTTACAAGAAGTC--------------- 1662

Qy        159 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178
Db       1663 ------GAATCCAAGACGTTTGAAGACTACTGGGTGCCCCTGTTAGAAAAACACGGTTAT 1716

Qy        179 ---------PheThrAla----------------------------------------- 181
Db       1717 ACCGGTATTTTCCATGCAAAGGCAAGAGCAAAGACCATGCACTCCAAGGACTCCAAGAAA 1776

Qy        182 --------------IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 196
Db       1777 GTCGATGGGTGTTGTATTTTCTTTAAAAGAGACCAATTCAAACTGGTCAACAAAGACTCC 1836

Qy        197 IleProPheProSerThrLysMetMetArgAsn-------------------------- 207
Db       1837 ATGGATTTCAGCGGTGCTTGGATGAAGCACAAGAAGTTTCAAAGAACAGAAGACTATTTG 1896

Qy        208 ----------------------------------LeuLeuCysValHisValAsnVal 215
Db       1897 AACCGCGCTATGAACAAGGACAACGTCGCGCTGTTTTTGAAACTTCAGCACGTCTCC--- 1953

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
Db       1954 AGCGGCGATACCATTTGGGTGGTCACCACACATTTA--------------CACTGGGAT 1998

Qy        236 GluArgMetAsnGlnLeuLys--------------MetValLeuLysLysMetGlnGlu 250
Db       1999 CCCAAATTCAACGACGTGAAGACGTTCCAAGTGGGTGTCTTACTAGACCATCTAGAA--- 2055

Qy        251 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu--- 269
Db       2056 ------------------ACGTTGTTGAAAGAGGACACAACGCTCAATTCCAGACAGGAC 2097

Qy        270 ---------------ValThrArgCysGlyGlyLeuProAsnAsnIleValAsp------ 282
Db       2098 ATCAAGAAATCCCCCGTGCTCATTTGCGGTGACTTC---AATTCGTACATCAACTCCGCC 2154

Qy        283 ValTrpGluPheLeu-----------------------------GlyLysProLysHis 292
Db       2155 GTATACGAATTGATAAGTACAGGCCGCGTTCAAGTACACCAAGAGGGCAGCAGCAGAGAT 2214

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla------ 310
Db       2215 TTCGGCTACATGTCAGAGAAAAACTTCTCGCACAATTTGGCCCTAAAATCAAGCTACAAC 2274

Qy        311 Cys------LysLeuArgPhe--------------------------------AspArg 317
Db       2275 TGTATCGGAGAGTTGCCATTCACCAATTTCACACCTTCATTCACAGACGTCATTGACTAC 2334

Qy        318 IlePhePhe--------------ArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
Db       2335 ATATGGTTCTCCACGCACGCGCTAAGAGTGCGTGGGCTATTGGGCGAAGTGGACCCTGAG 2394

Qy        333 SerLeuAsp---LeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHis 351
Db       2395 TACGTGGATAAGTTTATCGGGTTCCCCAAT---------GACAAATTCCCCAGTGACCAT 2445

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIle 361
Db       2446 ATACCATTGCTAGCAAGGTTTGAGTTCATG 2475

RESULT 14
US-10-932-182A-52
; Sequence 52, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 52
;   LENGTH: 2505
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-52

Alignment Scores:
Pred. No.:                0.8          Length:          2505
Score:                    106.50       Matches:         92
```

```
Percent Similarity:     33.5%          Conservative:  72
Best Local Similarity:  18.8%          Mismatches:    143
Query Match:            5.6%           Indels:        183
DB:                     7              Gaps:          27

US-10-757-745-2 (1-362) x US-10-932-182A-52 (1-2505)

Qy         20 ProGluValLysLys---------ArgArgLeuLeuCysValGluPheAlaSerValAla 36
Db       1111 CCGGAGATTAAGCGTTTGGACAACCTGCGTGTCTTGGACTTGTCCCATAATAGGTTGACC 1170

Qy         37 SerCysAspAlaAlaValAlaGlnCysPhe----------------------------- 46
Db       1171 GCGCTACCAGCGGAACTGGGCTCCTGCTTCCAGTTGAAGTACCTGTATTTCTTCGATAAC 1230

Qy         47 ---LeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProPro 65
Db       1231 ATGGTCACCACGCTACCATGGGAATTCGGCAATCTGTACAACTTGCAGTTT---CTCGGT 1287

Qy         66 ValGluGluSerAlaLeuGluArg---------------------------------- 73
Db       1288 GTAGAGGGCAATCCCTTGGAGAGGCAGCTTTTGAAGATCCTCACAGAAAAATCTGTCACG 1347

Qy         74 --------------------------ArgProGlu---ThrIleSerGluProLysThr 83
Db       1348 GGACTGATCTTCTACCTAAGAGACAACAGACCGGAGATCACCTTACCTCACGAGCGCAGG 1407

Qy         84 TyrValAspLeuThrAsn--------------------------------GluGluThr 92
Db       1408 TTCATCGAGATCAATACTGATGGGGAACCACAAAGAGAATATGACTCCTTGCAACAATCG 1467

Qy         93 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 112
Db       1468 ACGGAACACTTGACCGCGGATCTGGCCAAGAGATCGTTCACCGTGCTATCGTACAATACT 1527

Qy        113 MetPheSer-----------------------------------------LeuIleThr 118
Db       1528 CTGTGCCAACACTACGCCACTCCCAAAATGTACCGTTACACACCATCATGGGCATTGAGT 1587

Qy        119 TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 138
Db       1588 TGGGAT---------TACAGACGTAACAAGCTAAAGGAACAG------------------ 1620

Qy        139 TyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 158
Db       1621 ---ATTTTGTCATACAGTAGTGACGTGTTGTGTTTACAAGAAGTC--------------- 1662

Qy        159 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178
Db       1663 ------GAATCCAAGACGTTTGAAGACTACTGGGTGCCCCTGTTAGAAAAACACGGTTAT 1716

Qy        179 ---------PheThrAla---------------------------------------- 181
Db       1717 ACCGGTATTTTCCATGCAAAGGCAAGAGCAAAGACCATGCACTCCAAGGACTCCAAGAAA 1776

Qy        182 ---------------IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 196
Db       1777 GTCGATGGGTGTTGTATTTTCTTTAAAAGAGACCAATTCAAACTGGTCAACAAAGACTCC 1836

Qy        197 IleProPheProSerThrLysMetMetArgAsn-------------------------- 207
Db       1837 ATGGATTTCAGCGGTGCTTGGATGAAGCACAAGAAGTTTCAAAGAACAGAAGACTATTTG 1896

Qy        208 -----------------------------------LeuLeuCysValHisValAsnVal 215
Db       1897 AACCGCGCTATGAACAAGGACAACGTCGCGCTGTTTTTGAAACTTCAGCACGTCTCC--- 1953

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
Db       1954 AGCGGCGATACCATTTGGGTGGTCACCACACATTTA--------------CACTGGGAT 1998

Qy        236 GluArgMetAsnGlnLeuLys--------------MetValLeuLysLysMetGlnGlu 250
```

```
Db       1999 CCCAAATTCAACGACGTGAAGACGTTCCAAGTGGGTGTCTTACTAGACCATCTAGAA--- 2055

Qy        251 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu--- 269
Db       2056 -----------------ACGTTGTTGAAAGAGGACACAACGCTCAATTCCAGACAGGAC 2097

Qy        270 --------------ValThrArgCysGlyGlyLeuProAsnAsnIleValAsp------ 282
Db       2098 ATCAAGAAATCCCCCGTGCTCATTTGCGGTGACTTC---AATTCGTACATCAACTCCGCC 2154

Qy        283 ValTrpGluPheLeu----------------------------GlyLysProLysHis 292
Db       2155 GTATACGAATTGATAAGTACAGGCCGCGTTCAAGTACACCAAGAGGGCAGCAGCAGAGAT 2214

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla------ 310
Db       2215 TTCGGCTACATGTCAGAGAAAAACTTCTCGCACAATTTGGCCCTAAAATCAAGCTACAAC 2274

Qy        311 Cys------LysLeuArgPhe--------------------------------AspArg 317
Db       2275 TGTATCGGAGAGTTGCCATTCACCAATTTCACACCTTCATTCACAGACGTCATTGACTAC 2334

Qy        318 IlePhePhe--------------ArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
Db       2335 ATATGGTTCTCCACGCACGCGCTAAGAGTGCGTGGGCTATTGGGCGAAGTGGACCCTGAG 2394

Qy        333 SerLeuAsp---LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
Db       2395 TACGTGGATAAGTTTATCGGGTTCCCCAAT---------GACAAATTCCCCAGTGACCAT 2445

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIle 361
Db       2446 ATACCATTGCTAGCAAGGTTTGAGTTCATG 2475
```

```
RESULT 15
US-11-078-951-1
; Sequence 1, Application US/11078951
; Publication No. US20050272119A1
; GENERAL INFORMATION:
;  APPLICANT: Kubota, Kazuishi; Nakahara, Kaori
;  APPLICANT:  Hara, Aayako; Ozeki, Yohei
;  APPLICANT:  Iijima, Yasuteru
;  TITLE OF INVENTION: 2',5'-Oligoadenylate phosphodiesterase
;  FILE REFERENCE: 05090C/HG
;  CURRENT APPLICATION NUMBER: US/11/078,951
;  CURRENT FILING DATE:  2005-03-11
;  PRIOR APPLICATION NUMBER: JP 2002-267797
;  PRIOR FILING DATE: 2002-09-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1830
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1827)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
US-11-078-951-1

Alignment Scores:
Pred. No.:              0.686          Length:        1830
Score:                  105.50         Matches:       57
Percent Similarity:     37.5%          Conservative:  47
Best Local Similarity:  20.6%          Mismatches:    100
Query Match:            5.5%           Indels:        73
DB:                     12             Gaps:          12

US-10-757-745-2 (1-362) x US-11-078-951-1 (1-1830)

Qy        143 TyrSerProAspValIlePheLeuGlnGluVal------------------------- 153
```

```
Db       1024 TACAACGCCGATGTCATCTGTTTGCAGGAGGTTGACCGCGCAGTGTTTTCTGACAGCTTG 1083

Qy        154 IleProProTyrTyrSerTyr--------------------LeuLysLys--------- 163

Db       1084 GTACCCGCCCTAGAGGCCTTCGGGCTCGAGGGGGTGTTTCGAATCAAGCAGCACGAAGGC 1143

Qy        164 --------------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178

Db       1144 CTGGCCACTTTCTACCGAAAGTCTAAGTTCAGCCTTCTTAGCCAGCATGACATTTCATTC 1203

Qy        179 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198

Db       1204 TACGAAGCCCTCGAGTCCGACCCACTTCACAAAGAACTG---CTGGAGAAACTAGTTTTG 1260

Qy        199 PheProSerThr------LysMetMetArgAsnLeuLeuCysValHisValAsnVal--- 215

Db       1261 TACCCATCAGCGCAGGAGAAGGTGCTCCAGAGATCTTCTGTTCTTCAGGTTTCAGTTCTT 1320

Qy        216 --------------SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 230

Db       1321 CAGTCTACAAAGGACTCTTCTAAAAGGATATGTGTTGCTAATACCCATCTTTACTGGCAT 1380

Qy        231 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 250

Db       1381 CCTAAAGGTGGGTATATTCGCCTCATTCAAATGGCAGTAGCCTTGGCTCACATAAGACAT 1440

Qy        251 AlaPro------------GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266

Db       1441 GTTTCATGTGATCTGTATCCTGGCATACCAGTTATATTTTGTGGGGACTTTAATAGTACA 1500

Qy        267 AspArg------------GluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 282

Db       1501 CCATCAACAGGAATGTATCATTTTGTCATCAATGGCAGCATTCCAGAGGATCATGAAGAC 1560

Qy        283 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 302

Db       1561 ---TGGGCTTCCAATGGGGAGGAGGAAAGATGC-------------------------- 1590

Qy        303 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArg----------------------- 314

Db       1591 ---AATATGTCTCTTACACATTTCTTCAAGCTGAAAAGTGCTTGTGGTGAACCTGCTTAC 1647

Qy        315 ------------------------------PheAspArgIlePhePheArgAlaAlaAla 324

Db       1648 ACAAATTATGTTGGTGGCTTTCATGGATGTCTAGATTACATTTTCATTGACTTAAATGCT 1707

Qy        325 ---------GluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 341

Db       1708 TTAGAGGTTGAACAGGTGATTCCATTACCTAGTCATGAAGAAGTTACCACCCACCAGGCC 1767

Qy        342 LeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358

Db       1768 TTACCTAGTGTTTCCCATCCCTCTGATCACATAGCACTTGTATGTGATTTA 1818
```

RESULT 16
US-11-078-951-10
; Sequence 10, Application US/11078951
; Publication No. US20050272119A1
; GENERAL INFORMATION:
; APPLICANT: Kubota, Kazuishi; Nakahara, Kaori
; APPLICANT: Hara, Aayako; Ozeki, Yohei
; APPLICANT: Iijima, Yasuteru
; TITLE OF INVENTION: 2',5'-Oligoadenylate phosphodiesterase
; FILE REFERENCE: 05090C/HG
; CURRENT APPLICATION NUMBER: US/11/078,951
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: JP 2002-267797
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1939
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-078-951-10

Alignment Scores:

Pred. No.:	0.739	Length:	1939
Score:	105.50	Matches:	57
Percent Similarity:	37.5%	Conservative:	47
Best Local Similarity:	20.6%	Mismatches:	100
Query Match:	5.5%	Indels:	73
DB:	12	Gaps:	12

US-10-757-745-2 (1-362) x US-11-078-951-10 (1-1939)

```
Qy        143 TyrSerProAspValIlePheLeuGlnGluVal-------------------------- 153

Db       1095 TACAACGCCGATGTCATCTGTTTGCAGGAGGTTGACCGCGCAGTGTTTTCTGACAGCTTG 1154

Qy        154 IleProProTyrTyrSerTyr--------------------LeuLysLys--------- 163

Db       1155 GTACCCGCCCTAGAGGCCTTCGGGCTCGAGGGGGTGTTTCGAATCAAGCAGCACGAAGGC 1214

Qy        164 --------------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178

Db       1215 CTGGCCACTTTCTACCGAAAGTCTAAGTTCAGCCTTCTTAGCCAGCATGACATTTCATTC 1274

Qy        179 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198

Db       1275 TACGAAGCCCTCGAGTCCGACCCACTTCACAAAGAACTG---CTGGAGAAACTAGTTTTG 1331

Qy        199 PheProSerThr------LysMetMetArgAsnLeuLeuCysValHisValAsnVal--- 215

Db       1332 TACCCATCAGCGCAGGAGAAGGTGCTCCAGAGATCTTCTGTTCTTCAGGTTTCAGTTCTT 1391

Qy        216 --------------SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 230

Db       1392 CAGTCTACAAAGGACTCTTCTAAAAGGATATGTGTTGCTAATACCCATCTTTACTGGCAT 1451

Qy        231 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 250

Db       1452 CCTAAAGGTGGGTATATTCGCCTCATTCAAATGGCAGTAGCCTTGGCTCACATAAGACAT 1511

Qy        251 AlaPro------------GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266

Db       1512 GTTTCATGTGATCTGTATCCTGGCATACCAGTTATATTTTGTGGGGACTTTAATAGTACA 1571

Qy        267 AspArg------------GluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 282

Db       1572 CCATCAACAGGAATGTATCATTTTGTCATCAATGGCAGCATTCCAGAGGATCATGAAGAC 1631

Qy        283 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 302

Db       1632 ---TGGGCTTCCAATGGGGAGGAGGAAAGATGC-------------------------- 1661

Qy        303 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArg----------------------- 314

Db       1662 ---AATATGTCTCTTACACATTTCTTCAAGCTGAAAAGTGCTTGTGGTGAACCTGCTTAC 1718

Qy        315 ------------------------------PheAspArgIlePhePheArgAlaAlaAla 324

Db       1719 ACAAATTATGTTGGTGGCTTTCATGGATGTCTAGATTACATTTTCATTGACTTAAATGCT 1778

Qy        325 ---------GluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 341

Db       1779 TTAGAGGTTGAACAGGTGATTCCATTACCTAGTCATGAAGAAGTTACCACCCACCAGGCC 1838

Qy        342 LeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358

Db       1839 TTACCTAGTGTTTCCCATCCCTCTGATCACATAGCACTTGTATGTGATTTA 1889
```

RESULT 17
US-10-330-773-582
; Sequence 582, Application US/10330773

; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 582
; LENGTH: 70708
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-582

Alignment Scores:

Pred. No.:	84	Length:	70708
Score:	105.00	Matches:	52
Percent Similarity:	43.1%	Conservative:	45
Best Local Similarity:	23.1%	Mismatches:	94
Query Match:	5.5%	Indels:	35
DB:	7	Gaps:	8

US-10-757-745-2 (1-362) x US-10-330-773-582 (1-70708)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||  ||||||     :::      |||:::                   :::
Db      17862 GAAGAAACTGCATCAACTAACGAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 17920

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||      :::::::::|||   |||::::::||||||          |||
Db      17921 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTCCA 17971

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::        :::   :::::::::             |||  |||      :::
Db      17972 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 18031

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||           :::::::|||      |||:::::::::
Db      18032 CAGGAA-----------------ACACATCTCACGTGCAGAAACACACATAGGCTCAAA 18073

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                 ||||||:::|||||||
Db      18074 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAAGGGTTGCA 18133

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::   :::          |||     ||||||:::    |||:::
Db      18134 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 18193

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::            ||||||   ::::::      :::
Db      18194 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATAT 18253

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||           |||   |||         :::|||   ||||||
Db      18254 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 18298

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||      :::|||   |||:::|||      ||||||   |||
Db      18299 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCCCA 18355

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                  :::|||   :::
Db      18356 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 18415

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              :::::::::|||::::::      |||      ||||||
Db      18416 TCAGCTCTGCACCAAGCGGATCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 18472

Qy        293 CysGlnTyrThrTrp 297
              :::|||||:::
Db      18473 ACAGAATATACATTT 18487
```

RESULT 18
US-10-661-966-1/c
; Sequence 1, Application US/10661966
; Publication No. US20050277118A1
; GENERAL INFORMATION:
; APPLICANT: Roth, Richard B.
; APPLICANT: Nelson, Matthew Roberts
; APPLICANT: Braun, Andreas
; TITLE OF INVENTION: METHODS FOR IDENTIFYING SUBJECTS AT RISK
; TITLE OF INVENTION: OF MELANOMA AND TREATMENTS THEREOF
; FILE REFERENCE: 524592003800
; CURRENT APPLICATION NUMBER: US/10/661,966
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: 60/410,595
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: 60/422,344
; PRIOR FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 190276
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-661-966-1

Alignment Scores:

Pred. No.:	383	Length:	190276
Score:	104.00	Matches:	58
Percent Similarity:	42.0%	Conservative:	50
Best Local Similarity:	22.6%	Mismatches:	102
Query Match:	5.5%	Indels:	48
DB:	8	Gaps:	10

US-10-757-745-2 (1-362) x US-10-661-966-1 (1-190276)

```
Qy         72 GluArgArgProGluThrIleSerGlu----------------------ProLysThrTyr 84
              :::|||  ||||||     ::::::                        |||   |||
Db     163396 AAAAGAACTCCTGAAGGAAGCACTAAACATGGAAAGGAACAAGTGGTACCAGCCACTGCA 163337

Qy         85 ValAspLeuThrAsnGluGluThrThrAspSer-----------------ThrThrSer 98
                        |||   :::|||   ::::::                        ||||||
Db     163336 AAATCATGCCAAAATGTAAAGACCATCGAGACTAGGAAGAAACTGCATCAACTAACGAGC 163277

Qy         99 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 118
              ||||||   |||:::                 :::   |||       ::::::::::|||
Db     163276 AAAATAA--CCAGCTAATATCATAATGACAGGATCAAATTCACACATAACAATATTAACT 163219

Qy        119 TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 138
                 |||:::::::||||||          |||         :::     :::   :::
Db     163218 TTAAACGTAAATGGACTG---------AATGCTCCAATTAAAAGACACAGACTGGCAAAC 163168

Qy        139 TyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 158
              ::::::            |||   |||      :::|||||
Db     163167 TGGATAAAGAGTCAAGACCCATCAGTGTGCTGCATACAGGAA----------------- 163126

Qy        159 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178
              :::::|||     |||   ::::::                 |||   |||       :::   |||
Db     163125 ACCCATCTCACATGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTAC 163066

Qy        179 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198
                   :::          ||||||:::   |||   :::         :::
Db     163065 CAAGCAAGTGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGAC 163006

Qy        199 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly--- 217
              |||      ||||||:::   |||:::                ::::::   |||   |||
Db     163005 TTTAAACCAACAAAGATCAAAAGAGACCAAGAAGGCCATTACATAATGGTAAAGGGACCA 162946

Qy        218 ---------AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 234
                        ||||||   ::::::    :::      :::|||
Db     162945 ATTCAACAAGAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGA 162895

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
```

```
Db     162894 GCACCCAGATTC------ATAAAGCAAGTCCTTAGTGACCTACAAAGAGACTTAGACTCC 162841

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeu-------------------------- 265
Db     162840 CACACAATAATAATG---GGAGACTTTAACATCCCACTGTCAACATTAGACAGATCAACG 162784

Qy        266 ---------------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
Db     162783 CGACAGAAAGTTAATAAGGACACCCAGGAATTGAACTCAGCTCTGCACCAAGGGGACCTA 162724

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db     162723 ATAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 162676

RESULT 19
US-10-995-561-13293
; Sequence 13293, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13293
;   LENGTH: 645179
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-13293

Alignment Scores:
Pred. No.:              1.83e+03        Length:         645179
Score:                  104.00          Matches:        54
Percent Similarity:     43.1%           Conservative:   43
Best Local Similarity:  24.0%           Mismatches:     94
Query Match:            5.5%            Indels:         35
DB:                     8               Gaps:           8

US-10-757-745-2 (1-362) x US-10-995-561-13293 (1-645179)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db     123129 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATCATGACAGGATCA 123187

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db     123188 AATTCACACATAACAATATTAACTTTAAATATAAATGGACTA---------AATTCTGCA 123238

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db     123239 ATTAAAAGACACAAACTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 123298

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db     123299 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 123340

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db     123341 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 123400

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db     123401 ATCCTAGTCTCTGACAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 123460

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db     123461 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 123520

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     123521 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 123565

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db     123566 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 123622

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db     123623 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 123682

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     123683 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 123739

Qy        293 CysGlnTyrThrTrp 297
Db     123740 ACAGAATATACATTT 123754

RESULT 20
US-10-330-773-314/c
; Sequence 314, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 314
;   LENGTH: 110469
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-314

Alignment Scores:
Pred. No.:              245             Length:         110469
Score:                  103.00          Matches:        55
Percent Similarity:     44.0%           Conservative:   44
Best Local Similarity:  24.4%           Mismatches:     92
Query Match:            5.4%            Indels:         35
DB:                     7               Gaps:           9

US-10-757-745-2 (1-362) x US-10-330-773-314 (1-110469)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db      32253 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 32195

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db      32194 AATTCACACATAACAATATTAACTTTAAATGTAAACGGA------CTAAATTCTCCAATT 32141

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      32140 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 32084

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db      32083 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 32042

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db      32041 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 31982

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db      31981 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 31922
```

```
Qy       211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
Db     31921 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATATAT 31862

Qy       227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     31861 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 31817

Qy       247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db     31816 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 31760

Qy       266 ------------------------------------ArgAspArgGluValThrArg 272
Db     31759 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 31700

Qy       273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     31699 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 31643

Qy       293 CysGlnTyrThrTrp 297
Db     31642 ACAGAATATACATTT 31628
```

RESULT 21
US-11-096-568A-5188
; Sequence 5188, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 5188
; LENGTH: 1829
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1829)
; OTHER INFORMATION: Ceres Seq. ID no. 14307363
US-11-096-568A-5188

Alignment Scores:

Pred. No.:	1.64	Length:	1829
Score:	102.00	Matches:	81
Percent Similarity:	34.4%	Conservative:	58
Best Local Similarity:	20.0%	Mismatches:	115
Query Match:	5.4%	Indels:	150
DB:	9	Gaps:	23

US-10-757-745-2 (1-362) x US-11-096-568A-5188 (1-1829)

```
Qy        17 GluGlyGluProGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAla 36
Db       144 AAAGGAACACCCAATGTGAAAATGCACCCCATCTGGTGCTTA----------------- 185

Qy        37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
Db       186 ---TGTAAAGCC-----------------ATTCAAACTAATTCATATGAAATTTCC--- 221

Qy        57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAla----------------- 70
Db       222 ------AATGGAAGCTATCATCCCACTGTGCAACAAACAACAGAAAAGAAGACAATGGAA 275

Qy        71 ---LeuGluArgArgProGluThrIleSerGlu---ProLysThrTyrValAspLeuThr 88
Db       276 AAGCTATCCAGGAAAGTTGAATCTATTGGGGCATTTCAAAAGTTGCCCATAGTGATGCCA 335

Qy        89 AsnGluGluThrThrAspSerThr---------ThrSerLysIleSerProSerGluAsp 105
```

```
Db       336 TCTATTGACATTATTGGCTCTGCACTCAGAAAGGCGAGGAAGGTTTCTCCAACGAAA--- 392

Qy       106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
Db       393 ----------------------------------------------------------GGA 395

Qy       126 LeuAsnAsnLeuSer------GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyr 143
Db       396 ATTGCCAACATTGCGAAAAGGGAGAAAAATAAGGGTGCAAAGCAACTTGATGCGTTG--- 452

Qy       144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
Db       453 -----------------ATGAAAGAAATTGCTGTTCCCTTGAGGACGTATGTGGAGAAT 494

Qy       164 ------------------------------ArgSerSerAsnTyrGluIleIleThrGly 173
Db       495 TTTCCTAACAAAATGCATCTTCATCCTTATGAGAGATCTCTTATTGAGCTGACACTTGGA 554

Qy       174 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 193
Db       555 ------GATGGATACTACGAAATGGTGCTACGAAAGGTAGACAGTTTAAGGAAGAGG--- 605

Qy       194 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 213
Db       606 ----------------------------------------------GTAGTATCTGTT 617

Qy       214 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 233
Db       618 GGAAAGGAACATGCTTCACTTTGCGCTAAGTCT-----------TCATCAAAGCGAGAA 665

Qy       234 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 253
Db       666 GCAGAGGAAAGGCTGAATGAG---------GGGCTTAAGAAAATTGAAGAG--------- 707

Qy       254 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp------------------ 267
Db       708 -----------ATTTTTGCTCAAGAACGAAAAGTTGTTGATGATTTGTTAGGCATAGCC 755

Qy       268 --------------------------------ArgGluValThrArgCys---GlyGly 275
Db       756 AAGACTCTCAGGGCAATGCCTGTAATTAATTTGGAAACACCAACTCTCTGCCTTGTTGGA 815

Qy       276 LeuProAsn-----------AsnIleValAspValTrpGluPheLeuGlyLysProLys 291
Db       816 GCTCCTAATGTAGGCAAATCTTCTTTGGTCCATGTACTCTCT---ACTGGCAAGCCTGAG 872

Qy       292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla--- 310
Db       873 ATTTGTAACTATCCCTTCACAACAAGA--------------GGAATTCTTATGGGTCAT 917

Qy       311 CysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 330
Db       918 ATTATTTTAAACTTTCAGAAG------TTTCAGGTAACAGATACTCCGGGCCTGCTAAGG 971

Qy       331 ProArgSerLeuAspLeuGlyLeuGluLysLeu----------------------- 342
Db       972 AGACATGATGAGGACAGGAATAATTTGGAAAAGTTGACTCTGGCTGTCCTTTCGCACCTG 1031

Qy       343 ------------------------------------AspCysGlyArgProSerAsp 350
Db      1032 CCGACAGCAGTTTTGTATGTCCATGACCTTTCTGGGGAGTGTGGCACCTCACCTTCTGAT 1091

Qy       351 HisTrpGlyLeu 354
Db      1092 CAGTTTTCAATA 1103
```

RESULT 22
US-11-121-086-33/c
; Sequence 33, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.

```
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33
;   LENGTH: 154548
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-33

Alignment Scores:
Pred. No.:                482             Length:         154548
Score:                    102.00          Matches:        55
Percent Similarity:       43.6%           Conservative:   43
Best Local Similarity:    24.4%           Mismatches:     93
Query Match:              5.4%            Indels:         35
DB:                       12              Gaps:           9

US-10-757-745-2 (1-362) x US-11-121-086-33 (1-154548)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||  ||||||     :::   ||||||:::                 :::
Db      35036 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 34978

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      ::::::   |||   |||::::::|||      ||||||:::
Db      34977 AATTCACACATAACAATATTMACTTTAAATGTAAATGGA------CTAAATTCTGCAATT 34924

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::             |||  |||      :::
Db      34923 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 34867

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                :::::|||       |||   ::::::
Db      34866 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 34825

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                ||||||:::   |||
Db      34824 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 34765

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||       ||||||:::   |||:::
Db      34764 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 34705

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::            ||||||   ::::::      :::
Db      34704 CATTACATAATGGTAAAGGGATCAATYCAACAAGAGGAGCTAACTATCCTAAATATTTAT 34645

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||           |||   |||          :::||||   ||||||
Db      34644 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 34600

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db      34599 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 34543

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                  :::|||   :::
Db      34542 CTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 34483

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db      34482 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 34426

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db      34425 ACAGAATATACATTT 34411
```

```
RESULT 23
US-10-330-773-458/c
; Sequence 458, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 458
;   LENGTH: 194534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(194534)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-458

Alignment Scores:
Pred. No.:                648             Length:         194534
Score:                    102.00          Matches:        53
Percent Similarity:       42.2%           Conservative:   42
Best Local Similarity:    23.6%           Mismatches:     96
Query Match:              5.4%            Indels:         35
DB:                       7               Gaps:           8

US-10-757-745-2 (1-362) x US-10-330-773-458 (1-194534)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||  ||||||     :::   ||||||:::                 :::
Db      25864 GAGGAAACCGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 25806

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      ::::::::::|||   |||::::::||||||          |||
Db      25805 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTCCA 25755

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::      :::   |||::::::                |||  |||      :::
Db      25754 ATTAAAAGACACAGACTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 25695

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                :::::|||       |||   ::::::
Db      25694 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 25653

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                ||||||:::   |||
Db      25652 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 25593

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg------------ 206
              :::         :::         |||       ||||||:::   |||
Db      25592 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 25533

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 226
              :::   |||   |||      ::::::          ||||||   ::::::      :::
Db      25532 CATTATTTAATGGTAAAGGGATCAATTCAACAAGAAGAGCTACCTATCCTAAATATATAT 25473

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||           |||   |||          :::|||   ||||||
Db      25472 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 25428

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db      25427 GACCTACAAAGAGAGTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 25371

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                  :::|||   :::
Db      25370 CCGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 25311
```

```
Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          ::::::::::|||::::::      |||      ||||||
Db      25310 TCAGCTTTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 25254

Qy        293 CysGlnTyrThrTrp 297
                ::::||||||:::
Db      25253 ACAGAATATACATTT 25239

RESULT 24
US-11-091-018-1
; Sequence 1, Application US/11091018
; Publication No. US20050287551A1
; GENERAL INFORMATION:
;  APPLICANT: Gretarsdottir, Solveig
;  APPLICANT:  Thorleifsson, Gudmar
;  APPLICANT:  Gulcher, Jeffrey R.
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR HUMAN STROKE;
;  TITLE OF INVENTION:  METHODS OF TREATMENT
;  FILE REFERENCE: 2345.2010-016
;  CURRENT APPLICATION NUMBER: US/11/091,018
;  CURRENT FILING DATE:  2005-03-25
;  PRIOR APPLICATION NUMBER: PCT/US03/29906
;  PRIOR FILING DATE: 2003-09-25
;  PRIOR APPLICATION NUMBER: 10/255,120
;  PRIOR FILING DATE: 2002-09-25
;  PRIOR APPLICATION NUMBER: 10/419,723
;  PRIOR FILING DATE: 2003-04-18
;  PRIOR APPLICATION NUMBER: 10/650,120
;  PRIOR FILING DATE: 2003-08-27
;  PRIOR APPLICATION NUMBER: 10/067,514
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 09/811,352
;  PRIOR FILING DATE: 2001-03-19
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1691140
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1691140)
;   OTHER INFORMATION: n=A,T,C or G
US-11-091-018-1

Alignment Scores:
Pred. No.:                1.03e+04        Length:          1691140
Score:                    102.00          Matches:         55
Percent Similarity:       46.9%           Conservative:    51
Best Local Similarity:    24.3%           Mismatches:      83
Query Match:              5.4%            Indels:          38
DB:                       12              Gaps:            9

US-10-757-745-2 (1-362) x US-11-091-018-1 (1-1691140)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                  :::
Db     311762 GAAGAAACTGCATCAACTAACAAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 311820

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||       ::::::::::|||   |||:::::|||||||                |||
Db     311821 AATTCACACATAACAATATTAACCTTAAATGTAAATGGGCTA---------ACTGCTTCA 311871

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                 :::   :::  :::::::::             |||  |||   ::::::
Db     311872 ATTAAAAGACACAGACTGGCAAACTGGATAAAGAGTCAAGACCCATCAGTGTGCTATATT 311931

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::|||      |||   ::::::
Db     311932 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 311973

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                  ||||||:::   |||

Db     311974 ATAAAGGGATGGAGGAAGATCTACCAAGCAAATGGAAAGCAAAAAAAGGCAGGGGTTGCA 312033

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
                 :::     :::          |||      ||||||:::   |||:::
Db     312034 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGCC 312093

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                 ::::::   |||   |||:::          ||||||  ::::::      :::
Db     312094 CATTACATAATGGTAAAGGGATCAATTCAACAGGAAGAGCTAACTATCCTAAATATATAT 312153

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                ::||||             |||          :::|||   ||||||:::
Db     312154 GCACCTAATACA---------GGAGTACCCAGATTC------ATAAAGCAAGTCCTTAGA 312198

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp------------ 262
                  ::::::         ::::::   |||::::::      ||||||
Db     312199 GACCTAGAAAGAGACTTAGACGCCCATACAATAATG------GGAGACTTTAACACCCCA 312252

Qy        263 ------------ThrAsnLeuArgAspArgGluValThrArg-CysGlyGlyLeu---- 276
                         |||:::   |||||||||:::::||||||  |||     |||:::
Db     312253 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTTAACAAGGATGTCCAGGAATTGAAC 312312

Qy        277 ----------ProAsn-AsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                       |||::: ::::::::::|||::::::      |||      ||||||
Db     312313 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCT 312369

Qy        293 CysGlnTyrThr 296
                ::::||||||
Db     312370 ACAGAATATACA 312381

RESULT 25
US-10-995-561-13295
; Sequence 13295, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13295
;   LENGTH: 81210
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-13295

Alignment Scores:
Pred. No.:                239             Length:          81210
Score:                    101.50          Matches:         52
Percent Similarity:       42.4%           Conservative:    43
Best Local Similarity:    23.2%           Mismatches:      96
Query Match:              5.3%            Indels:          34
DB:                       8               Gaps:            8

US-10-757-745-2 (1-362) x US-10-995-561-13295 (1-81210)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   |||:::                  :::
Db      22604 GAAGAAACTGCATCAACTAACGAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 22662

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||       ::::::::::|||   |||:::::||||||                |||
Db      22663 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTTCA 22713

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                 :::   :::  :::::::::             |||  |||       :::
Db      22714 ATCAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 22773

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
```

```
Db      22774 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 22815

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db      22816 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 22875

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db      22876 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 22935

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db      22936 CATTACATAATGGTAAAGGGATCAATTCAACAGGAAGAGCTAACTATCCTAAATATATAT 22995

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db      22996 GCACCCAACAATACAGGA------GCACCCAGATTC------ATAAAGCAAGTCCTGAGT 23043

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db      23044 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATT---GGAGACTTTAACACCCCA 23100

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db      23101 CTGTCAATATTAGACAGATCAACGAGACAAAAAGTTAACAAAGATAGCCAGGAATTGAAC 23160

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db      23161 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 23217

Qy        293 CysGlnTyrThr 296
Db      23218 ACAGAATATACA 23229

RESULT 26
US-11-121-086-98
; Sequence 98, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 98
;   LENGTH: 171732
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-98

Alignment Scores:
Pred. No.:              625          Length:         171732
Score:                  101.50       Matches:        52
Percent Similarity:     42.4%        Conservative:   43
Best Local Similarity:  23.2%        Mismatches:     96
Query Match:            5.3%         Indels:         34
DB:                     12           Gaps:           8

US-10-757-745-2 (1-362) x US-11-121-086-98 (1-171732)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db     153193 GAAGAAACTGCATCAACTAACGAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 153251

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db     153252 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTTCA 153302

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db     153303 ATCAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 153362

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db     153363 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 153404

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db     153405 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 153464

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db     153465 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 153524

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db     153525 CATTACATAATGGTAAAGGGATCAATTCAACAGGAAGAGCTAACTATCCTAAATATATAT 153584

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     153585 GCACCCAACAATACAGGA------GCACCCAGATTC------ATAAAGCAAGTCCTGAGT 153632

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db     153633 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATT---GGAGACTTTAACACCCCA 153689

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db     153690 CTGTCAATATTAGACAGATCAACGAGACAAAAAGTTAACAAAGATAGCCAGGAATTGAAC 153749

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     153750 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 153806

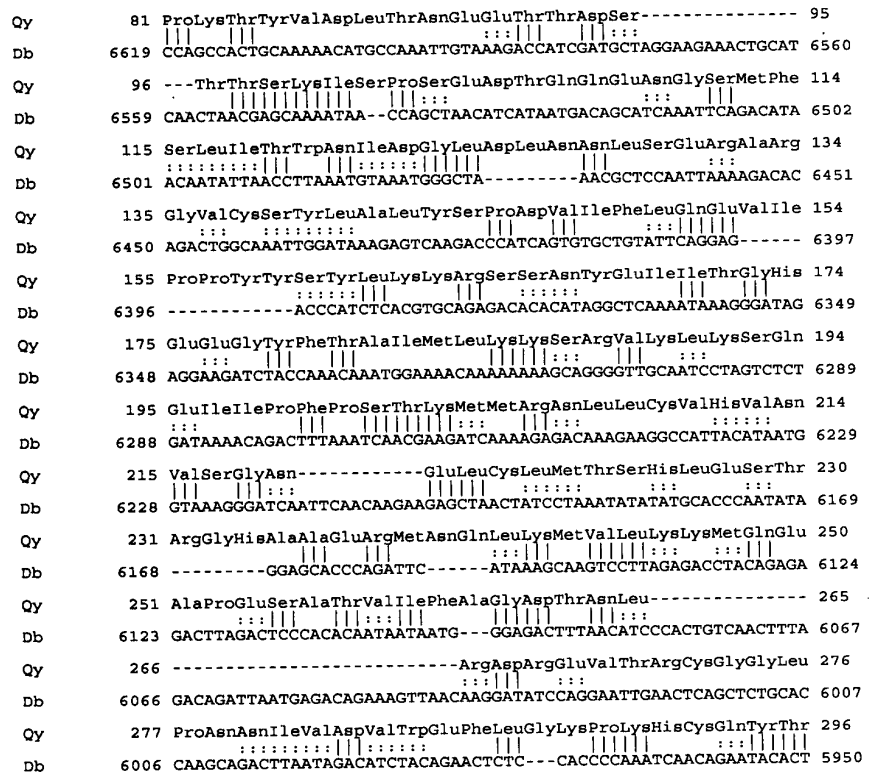
Qy        293 CysGlnTyrThr 296
Db     153807 ACAGAATATACA 153818

RESULT 27
US-10-330-773-248/c
; Sequence 248, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 248
;   LENGTH: 24861
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-248

Alignment Scores:
Pred. No.:              59.5         Length:         24861
Score:                  101.00       Matches:        60
Percent Similarity:     42.3%        Conservative:   50
Best Local Similarity:  23.1%        Mismatches:     103
Query Match:            5.3%         Indels:         48
DB:                     7            Gaps:           10

US-10-757-745-2 (1-362) x US-10-330-773-248 (1-24861)

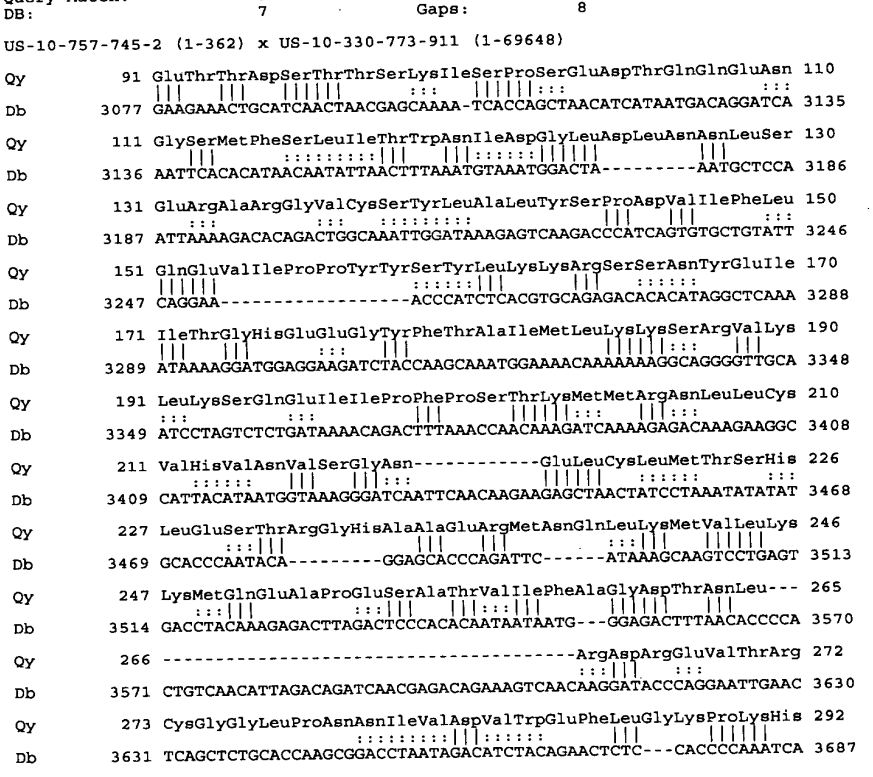
Qy         68 GluSerAlaLeuGluArgArgProGluThrIleSerGlu---------------------- 80
Db       6679 CAGGCCTGCCTTACAAGAGCTCCTGAAGGAAGCACTAAACATGGAAAGGAACAACTGGTA 6620
```



RESULT 28
US-10-330-773-911
; Sequence 911, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 911
; LENGTH: 69648
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-911

Alignment Scores:

Pred. No.:	223	Length:	69648
Score:	101.00	Matches:	52
Percent Similarity:	42.7%	Conservative:	44
Best Local Similarity:	23.1%	Mismatches:	95



RESULT 29
US-11-072-175-59
; Sequence 59, Application US/11072175
; Publication No. US20060029944A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273A CIP
; CURRENT APPLICATION NUMBER: US/11/072,175
; CURRENT FILING DATE: 2005-03-05
; PRIOR APPLICATION NUMBER: US 60/406,385
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US 10/648,593
; PRIOR FILING DATE: 2003-08-26
; NUMBER OF SEQ ID NOS: 571
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 59
; LENGTH: 5238

```
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-11-072-175-59

Alignment Scores:
Pred. No.:              10.4          Length:          5238
Score:                  100.00        Matches:         61
Percent Similarity:     36.4%         Conservative:    46
Best Local Similarity:  20.7%         Mismatches:      117
Query Match:            5.3%          Indels:          70
DB:                     9             Gaps:            12

US-10-757-745-2 (1-362) x US-11-072-175-59 (1-5238)

Qy         17 GluGlyGluProGluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAla 36
              |||   |||       :::|||||||   :::|||||||||:::::::
Db       1141 GAAGCAGAGAATGCTCTCAAAAAGGCAAAATTATTATGCATGCAACGTCAAGAT------ 1194

Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp-------------- 51
                 :::   |||   :::       |||   ||||||   :::
Db       1195 GAATATGAGAAAGCAAAGTCTTCCATGTTTCGTGCAGAAGAGGAGCATCTGTCTTCAAGT 1254

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProPro------ValGluGluSer 69
                 :::   :::   ||||||       |||            :::||||||
Db       1255 GGCGGATTAGCAAAAAATCTCAACAAGCAACTAGAAAAAAAGCGAAGGTTGGAAGAGGAG 1314

Qy         70 AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsn 89
              ||||||::::::       |||            |||       |||
Db       1315 GCTCTCCAAAAAGTAGAAGAAGCAGATGAACTTTACAAAGTTTGTGTGACAAATGTTGAA 1374

Qy         90 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 109
              |||           |||                        |||:::|||:::::::|||
Db       1375 GAAAGAAGAAATGAT----------------------GTAGAAAATACCAAAAGAGAA 1410

Qy        110 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp---------------------- 122
                    :::           :::   :::       |||
Db       1411 ATTTTAGCACAACTCCGGACACTTGTTTTCCAGTGTGATCTTACCCTTAAAGCGGTAACA 1470

Qy        123 --------------------GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGly 135
                                   |||           :::|||::::::|||   :::
Db       1471 GTTAACCTCTTCCACATGCAGCATCTGCAGGCTGCTTCCCTTGCAGACAGATTACAGTCT 1530

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              :::|||             ||||||   |||                 ||||||
Db       1531 CTCTGTGGTAGTGCCAAACTCTATGACCCAGGC------------CAAGAG--------- 1569

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSer--------------------------- 166
                 |||       ::::::|||      :::|||            )
Db       1570 ---TACAGTGAATTTGTCAAGGCCACAAATTCAACTGAAGAAGAAAAAGTTGATGGAAAT 1626

Qy        167 ---AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 185
                 |||       :::   :::              |||:::       |||       |||
Db       1627 GTAAATAAACATTTAAATAGTTCCCAACCTTCAGGATTTGGACCTGCCAACTCTTTA--- 1683

Qy        186 LysSerArgValLysLeuLysSerGlnGluIleIleProPhePro------SerThrLys 203
                                  ::::::::::::       |||       |||   |||
Db       1684 ----------------------GAGGATGTTGTACGCCTTCCTGACAGTTCTAATAAA 1719

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu--- 222
              :::       :::                  :::::::::|||
Db       1720 ATTGAAGAGGACAGATGCTCTAACAGTGCAGATATAACAGGTCCTTCCTTTATAAGATCA 1779

Qy        223 ------------MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
                        |||   |||           |||||||||   |||   :::       |||
Db       1780 TGGACATTTGGGATGTTTAGTGATTCTGAGAGCACTGGAGGGAGCAGCGAATCTAGATCT 1839

Qy        239 AsnGlnLeuLysMetValLeu--------------LysLysMetGlnGluAlaProGlu 253
                    :::   :::                    :::|||:::       |||
Db       1840 CTGGATTCAGAATCTATAAGTCCAGGAGACTTTCATCGAAAACTTCCACGAACACCATCC 1899

Qy        254 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 267
              |||   |||:::       |||   |||   :::   |||:::
Db       1900 AGTGGAACTATGTCCTCTGCAGATGATCTAGATGAAAGAGAG 1941
```

```
RESULT 30
US-11-121-086-51
; Sequence 51, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51
;   LENGTH: 176760
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-51

Alignment Scores:
Pred. No.:              942           Length:          176760
Score:                  100.00        Matches:         57
Percent Similarity:     42.4%         Conservative:    49
Best Local Similarity:  22.8%         Mismatches:      100
Query Match:            5.3%          Indels:          45
DB:                     12            Gaps:            10

US-10-757-745-2 (1-362) x US-11-121-086-51 (1-176760)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||       :::   ||||||:::                 :::
Db     81278 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 81336

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||       ::::::::::|||   |||::::::||||||           |||
Db     81337 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTCCA 81387

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                 :::         :::   :::::::::             |||   |||      :::
Db     81388 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 81447

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                 ::::::|||       |||   ::::::
Db     81448 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 81489

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                 ||||||:::   |||
Db     81490 ATAAAAGGATGGGGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 81549

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
                 :::         :::         |||       ||||||:::   |||:::
Db     81550 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAAGC 81609

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                 ::::::   |||   |||:::           ||||||   ::::::      :::
Db     81610 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATAT 81669

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                    :::|||               |||   |||         :::|||   ||||||
Db     81670 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 81714

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                 :::|||         :::|||   |||:::|||       ||||||   |||
Db     81715 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 81771

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db     81772 CTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 81831

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
```

```
Db      81832 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 81888

Qy        293 CysGlnTyrThrTrp--------------------AspThrGlnMetAsnSerAsnLeu 305

Db      81889 ACAGAATATACATTTTTTTCAGCACCACACCACACCTGTTCCAAAATGGACCACATACTT 81948

Qy        306 Gly---------IleThrAlaAlaCysLys 312

Db      81949 GGAAGTAAAGCTCTCCTCAGCAAATGTAAA 81978
```

RESULT 31
US-10-330-773-114
; Sequence 114, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 440684
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(440684)
; OTHER INFORMATION: n = A,T,C or G
US-10-330-773-114

Alignment Scores:

Pred. No.:	3.44e+03	Length:	440684
Score:	99.50	Matches:	51
Percent Similarity:	44.0%	Conservative:	48
Best Local Similarity:	22.7%	Mismatches:	91
Query Match:	5.2%	Indels:	36
DB:	7	Gaps:	10

US-10-757-745-2 (1-362) x US-10-330-773-114 (1-440684)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110

Db     366764 GAAGAAACTGCATCGACTAACAAGCAAAATAA-CCAACTAACATCATAATGACAGGATCA 366822

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db     366823 AATTCACACATAACAATATTAACCTTAAATGTAAATGGGCTA---------AATGCTCCA 366873

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db     366874 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTGTT 366933

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db     366934 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACAAATAGGCTCAGA 366975

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db     366976 ATAAAGGGATGGAGGAAGATCTAC---CAAGCAAATGGAAAACAAAAGGCAGGGGTTGCA 367032

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg------------ 206

Db     367033 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGGTCAAAAGAGAGAAAGAAGGC 367092

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 226

Db     367093 CATTACATAATGGTAAAGGGTTCAATTCAACAAGAAGAGCTAACTATCCTAAATATATAT 367152

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db     367153 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTTAGA 367197

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266

Db     367198 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTCTAACACCCCA 367254

Qy        267 ----------------------------------AspArgGluValThrArgCysGly 274

Db     367255 CTGTCAACATTAGACAGATCAACAAGACAGAAAGTTAACAAGGATATCCAGGAATTGAGC 367314

Qy        275 GlyLeuProAsn------AsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db     367315 TCAGCTCTTAACCAAGCAGACCTAATTGACATCTACAGAACTCTC---CACCCCAAATCA 367371

Qy        293 CysGlnTyrThrTrp 297

Db     367372 GCAGAATATACATTC 367386
```

RESULT 32
US-10-330-773-76/c
; Sequence 76, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 118544
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-76

Alignment Scores:

Pred. No.:	724	Length:	118544
Score:	99.00	Matches:	55
Percent Similarity:	45.8%	Conservative:	48
Best Local Similarity:	24.4%	Mismatches:	88
Query Match:	5.2%	Indels:	35
DB:	7	Gaps:	9

US-10-757-745-2 (1-362) x US-10-330-773-76 (1-118544)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110

Db      98861 GAAGAAACTGCATCAACTAATGAGCAAAATAA-CCAGCTACCATCATAATGACAGGATCA 98803

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db      98802 AATTCACACATAACAATATTAACCTTAAATGTAAATGGGCTA---------AATGCTTCA 98752

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db      98751 ATTAAAAGACACAGACTGGCAAACTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 98692

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db      98691 TAGGAA-----------------ACCCATCTCATATGCAGAGACACACATAGGCTCAAA 98650

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db      98649 ATAAAGGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 98590

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db      98589 ATCCTAGTCTCTGATAAAACAGACTTTAAACCGACAAAGATCAAAAGACACAAAGAAGGC 98530

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
```

```
                ::::::   |||   |||:::             ||||||   ::::::      :::
Db      98529 TATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTGAATATATAT 98470

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                ::|||             |||   |||          ::|||   ||||||:::
Db      98469 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTTAGA 98425

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::      |||:::|||     |||||   |||
Db      98424 GACCTACAAAGAGACTTAGACTACCACACAATAATAATG---GGAGACTTTAACACCCCA 98368

Qy        266 ------------------ArgAspArgGluValThrArgCysGly-GlyLeu-------- 276
                             :::||||||:::::||||||       |||:::
Db      98367 CTGTCAACATTAGATCAACAAGACAGAAAGATAACAAGGATATCCAGGAATTGAACTCAG 98308

Qy        277 -------ProAsn-AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysG 294
                  |||::: :::::::::|||::::::      |||      ||||||      :
Db      98307 CTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCAACAG 98251

Qy        294 lnTyrThrTrp 297
              ::|||||:::
Db      98250 AATATACATTC 98240
```

```
RESULT 33
US-10-928-446A-1
; Sequence 1, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825234)
;   OTHER INFORMATION: the 'n' at position 825234 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825270)
;   OTHER INFORMATION: the 'n' at position 825270 may be 'c' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825401)
;   OTHER INFORMATION: the 'n' at position 825401 may be 'c' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825428)
;   OTHER INFORMATION: the 'n' at position 825428 may be 'g' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825473)
;   OTHER INFORMATION: the 'n' at position 825473 may be 'g' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825765)
;   OTHER INFORMATION: a "c" may be deleted at this position
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825828)
;   OTHER INFORMATION: the 'n' at position 825828 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (826041)
;   OTHER INFORMATION: the 'n' at position 826041 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (826546)
;   OTHER INFORMATION: the 'n' at position 826546 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (826654)
;   OTHER INFORMATION: the 'n' at position 826654 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (826826)
;   OTHER INFORMATION: the 'n' at position 826826 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (826863)
;   OTHER INFORMATION: the 'n' at position 826863 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (827008)
;   OTHER INFORMATION: the 'n' at position 827008 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (843055)
;   OTHER INFORMATION: the 'n' at position 843055 may be 'g' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (843118)
;   OTHER INFORMATION: the 'n' at position 843118 may be 't' or 'c'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (871027)
;   OTHER INFORMATION: the 'n' at position 871027 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (871140)
;   OTHER INFORMATION: the 'n' at position 871140 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (871168)
;   OTHER INFORMATION: the 'n' at position 871168 may be 'c' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (872678)
;   OTHER INFORMATION: the 'n' at position 872678 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (872742)
;   OTHER INFORMATION: the 'n' at position 872742 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (925859)
;   OTHER INFORMATION: the 'n' at position 925859 may be 'c' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (993220)
;   OTHER INFORMATION: the 'n' at position 993220 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (993254)
;   OTHER INFORMATION: the 'n' at position 993254 may be 'g' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1006462)
;   OTHER INFORMATION: the 'n' at position 1006462 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1007820)..(1007823)
;   OTHER INFORMATION: "ttct" may be deleted at this position
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1018038)
;   OTHER INFORMATION: the 'n' at position 1018038 may be 'a' or 'g'
```

```
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1018704)
;    OTHER INFORMATION: the 'n' at position 1018704 may be 'c' or 't'
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1018718)..(1018720)
;    OTHER INFORMATION: "gtt" may be deleted at this position
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1026786)
;    OTHER INFORMATION: the 'n' at position 1026786 may be 'c' or 'a'
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1047134)
;    OTHER INFORMATION: the 'n' at position 1047134 may be 'a' or 'g'
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1047159)
;    OTHER INFORMATION: the 'n' at position 1047159 may be 'a' or 'g'
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1047378)
;    OTHER INFORMATION: the 'n' at position 1047378 may be 'c' or 't'
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1047739)
;    OTHER INFORMATION: the 'n' at position 1047739 may be 'a' or 'g'
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1050133)..(1050137)
;    OTHER INFORMATION: "ttaaa" may be deleted at this position
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1050539)
;    OTHER INFORMATION: the 'n' at position 1050539 may be 'c' or 't'
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1062808)
;    OTHER INFORMATION: the 'n' at position 1062808 may be 'c' or 'g'
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1066392)
;    OTHER INFORMATION: the 'n' at position 1066392 may be 'a' or 'g'
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (1073711)
;    OTHER INFORMATION: the 'n' at position 1073711 may be 'c' or 't'
US-10-928-446A-1

Alignment Scores:
Pred. No.:                1.23e+04        Length:          1080000
Score:                    99.00           Matches:         54
Percent Similarity:       43.6%           Conservative:    44
Best Local Similarity:    24.0%           Mismatches:      93
Query Match:              5.2%            Indels:          35
DB:                       8               Gaps:            9

US-10-757-745-2 (1-362) x US-10-928-446A-1 (1-1080000)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                 :::
Db     399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||     ::::::   |||   |||::::::|||      ||||||:::
Db     399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   :::::::::           |||   |||       :::
Db     399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::::|||       |||   ::::::
Db     399292 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                     ||||||:::   |||
Db     399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||       ||||||:::   |||:::
Db     399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                 ::::::   |||   |||:::           ||||||   |||:::      :::
Db     399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||           |||   |||           :::|||   :::|||
Db     399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                 :::|||         :::|||   |||:::|||      ||||||   |||
Db     399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db     399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          ::::::::::|||:::::      |||      ||||||
Db     399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy        293 CysGlnTyrThrTrp 297
                 :::||||||:::
Db     399733 ACAGAATATACATTT 399747

RESULT 34
US-10-928-446A-181
; Sequence 181, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 181
;    LENGTH: 1080000
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (826985)..(827008)
;    FEATURE:
;    OTHER INFORMATION: full exon 1 range is 826667-827008
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: (827008)..(827008)
;    OTHER INFORMATION: the 'n' at position 827008 may be 'a' or 'g'
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (843242)..(843315)
;    OTHER INFORMATION: exon
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (922549)..(922630)
;    OTHER INFORMATION: exon
;    FEATURE:
;    NAME/KEY: CDS
```

```
;   LOCATION: (926021)..(926059)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (929123)..(929176)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (993104)..(993154)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999547)..(999608)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full exon 30 range is 1073289-1075279
US-10-928-446A-181

Alignment Scores:
Pred. No.:                1.23e+04        Length:          1080000
Score:                    99.00           Matches:         54
Percent Similarity:       43.6%           Conservative:    44
Best Local Similarity:    24.0%           Mismatches:      93
Query Match:              5.2%            Indels:          35
DB:                       8               Gaps:            9

US-10-757-745-2 (1-362) x US-10-928-446A-181 (1-1080000)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                  :::
Db     399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||      ::::::   |||   |||::::::|||      |||||||:::
Db     399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   :::::::::             |||   |||      :::
Db     399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                 :::::::|||      |||   ::::::
Db     399292 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                |||||:::   |||
Db     399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||       ||||||:::   |||:::
Db     399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                 ::::::   |||   |||:::            ||||||   |||:::     :::
Db     399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513
```

```
Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                :::|||          |||  |||       :::|||  :::|||
Db     399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                :::|||       :::|||  |||:::|||     ||||||  |||
Db     399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||  :::
Db     399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                         :::::::::|||::::::     |||     ||||||
Db     399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy        293 CysGlnTyrThrTrp 297
                :::|||||:::
Db     399733 ACAGAATATACATTT 399747

RESULT 35
US-10-928-446A-183
; Sequence 183, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING       *
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 183
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (826985)..(827017)
;   FEATURE:
;   OTHER INFORMATION: full exon 1 range is 826667-827008
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (843242)..(843315)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (922549)..(922630)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (926021)..(926059)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (929123)..(929176)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (993104)..(993154)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999547)..(999608)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full exon 30 range is 1073289-1075279
US-10-928-446A-183

Alignment Scores:
Pred. No.: ..           1.23e+04        Length:         1080000
Score:                  99.00           Matches:        54
Percent Similarity:     43.6%           Conservative:   44
Best Local Similarity:  24.0%           Mismatches:     93
Query Match:            5.2%            Indels:         35
DB:                     8               Gaps:           9

US-10-757-745-2 (1-362) x US-10-928-446A-183 (1-1080000)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||  ||||||      :::  ||||||:::                 :::
Db     399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      ::::::   |||   |||::::::|||      ||||||:::
Db     399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   :::::::::             |||   |||      :::
Db     399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  ::::::|||      |||   ::::::
Db     399292 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||  |||      :::   |||                 ||||||:::   |||
Db     399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||      ||||||:::   |||:::
Db     399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::           ||||||   |||:::      :::
Db     399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||            |||   |||         :::|||   :::|||
Db     399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                 :::|||         :::|||   |||:::|||      ||||||   |||
Db     399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                  :::|||   :::
Db     399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db     399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy        293 CysGlnTyrThrTrp 297
                          :::||||||:::
Db     399733 ACAGAATATACATTT 399747

RESULT 36
US-10-928-446A-185
; Sequence 185, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 185
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (722487)..(722534)
;   FEATURE:
;   OTHER INFORMATION: full length exon 1 range is 722213-722534
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (843242)..(843315)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (922549)..(922630)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (926021)..(926059)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (929123)..(929176)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (993104)..(993154)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999547)..(999608)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
```

```
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-185
```

```
Alignment Scores:
Pred. No.:                1.23e+04        Length:          1080000
Score:                    99.00           Matches:         54
Percent Similarity:       43.6%           Conservative:    44
Best Local Similarity:    24.0%           Mismatches:      93
Query Match:              5.2%            Indels:          35
DB:                       8               Gaps:            9

US-10-757-745-2 (1-362) x US-10-928-446A-185 (1-1080000)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                   :::
Db     399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||       ::::::   |||   |||:::::|||       ||||||:::
Db     399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   :::::::::                |||   |||   :::
Db     399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||              ::::::|||       |||   ::::::
Db     399292 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                 ||||||:::   |||
Db     399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::       :::       |||       ||||||:::   |||:::
Db     399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::             ||||||   |||:::   :::
Db     399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                :::|||           |||   |||         :::|||   :::|||
Db     399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||       :::|||   |||:::|||       ||||||   |||
Db     399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db     399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::   |||       ||||||
Db     399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db     399733 ACAGAATATACATTT 399747

RESULT 37
US-10-928-446A-187
; Sequence 187, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 187
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-187

Alignment Scores:
Pred. No.:                1.23e+04        Length:          1080000
Score:                    99.00           Matches:         54
Percent Similarity:       43.6%           Conservative:    44
Best Local Similarity:    24.0%           Mismatches:      93
Query Match:              5.2%            Indels:          35
DB:                       8               Gaps:            9

US-10-757-745-2 (1-362) x US-10-928-446A-187 (1-1080000)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                 :::
Db     399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||       ::::::   |||   |||::::::|||       ||||||:::
Db     399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::.  :::::::::             |||   |||      :::
Db     399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||.|||                  ::::::|||      |||   ::::::
Db     399292 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                   ||||||:::   |||
Db     399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||       ||||||:::   |||:::
Db     399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::          ||||||   |||:::      :::
Db     399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                ::|||           |||   |||         :::|||   :::|||
Db     399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558
```

```
Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db     399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db     399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy        293 CysGlnTyrThrTrp 297
Db     399733 ACAGAATATACATTT 399747

RESULT 38
US-10-928-446A-189
; Sequence 189, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 189
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (723508)..(723516)
;   FEATURE:
;   OTHER INFORMATION: full length exon 1 range is 723483-723516
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-189

Alignment Scores:
Pred. No.:                1.23e+04        Length:          1080000
Score:                    99.00           Matches:         54
Percent Similarity:       43.6%           Conservative:    44
Best Local Similarity:    24.0%           Mismatches:      93
Query Match:              5.2%            Indels:          35
DB:                       8               Gaps:            9

US-10-757-745-2 (1-362) x US-10-928-446A-189 (1-1080000)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
```

```
Db      399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db      399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db      399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db      399292 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db      399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db      399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy         211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226

Db      399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db      399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db      399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272

Db      399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db      399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy         293 CysGlnTyrThrTrp 297

Db      399733 ACAGAATATACATTT 399747

RESULT 39
US-10-928-446A-191
; Sequence 191, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 191
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
```

```
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-191

Alignment Scores:
Pred. No.:              1.23e+04        Length:         1080000
Score:                  99.00           Matches:        54
Percent Similarity:     43.6%           Conservative:   44
Best Local Similarity:  24.0%           Mismatches:     93
Query Match:            5.2%            Indels:         35
DB:                     8               Gaps:           9

US-10-757-745-2 (1-362) x US-10-928-446A-191 (1-1080000)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::             :::
Db      399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||     ::::::   |||   |||::::::|||       ||||||:::
Db      399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   :::::::::           |||   |||      :::
Db      399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                :::::|||      |||   ::::::
Db      399292 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||       :::   |||                 ||||||:::   |||
Db      399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::         :::         |||       ||||||:::   |||:::
Db      399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                 ::::::   |||   |||:::           ||||||   |||:::      :::
Db      399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::|||           |||   |||          :::|||   :::|||
Db      399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                ::::|||           :::|||   |||:::|||       ||||||   |||
Db      399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy         266 -------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db      399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                           :::::::::|||::::::       |||       ||||||
Db      399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy         293 CysGlnTyrThrTrp 297
                  :::|||||:::
Db      399733 ACAGAATATACATTT 399747

RESULT 40
US-10-928-446A-193
; Sequence 193, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 193
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-193

Alignment Scores:
Pred. No.:              1.23e+04        Length:         1080000
Score:                  99.00           Matches:        54
Percent Similarity:     43.6%           Conservative:   44
Best Local Similarity:  24.0%           Mismatches:     93
Query Match:            5.2%            Indels:         35
DB:                     8               Gaps:           9

US-10-757-745-2 (1-362) x US-10-928-446A-193 (1-1080000)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                  :::
Db     399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||      ::::::   |||   |||::::::|||      ||||||:::
Db     399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   :::::::::          |||   |||      :::
Db     399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  ::::::|||      |||   ::::::
Db     399292 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                 ||||||:::   |||
Db     399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::       :::         |||      ||||||:::   |||:::
Db     399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::          ||||||   |||:::      :::
Db     399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||           |||   |||         :::|||   :::|||
Db     399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||          :::|||   |||:::|||      ||||||   |||
Db     399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db     399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          ::::::::::|||::::::      |||      ||||||
Db     399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db     399733 ACAGAATATACATTT 399747

RESULT 41
US-10-928-446A-195
; Sequence 195, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 195
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-195

Alignment Scores:
Pred. No.:                1.23e+04        Length:          1080000
Score:                    99.00           Matches:         54
Percent Similarity:       43.6%           Conservative:    44
Best Local Similarity:    24.0%           Mismatches:      93
Query Match:              5.2%            Indels:          35
DB:                       8               Gaps:            9

US-10-757-745-2 (1-362) x US-10-928-446A-195 (1-1080000)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::              :::
Db      399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||       ::::::   |||   |||:::::|||      ||||||:::
Db      399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   :::::::::             |||   |||     :::
Db      399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||             :::::::|||       |||   ::::::
Db      399292 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||       :::   |||             ||||||:::   |||
Db      399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::       :::             |||       ||||||:::   |||:::
Db      399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::          ||||||   |||:::     :::
Db      399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||             |||   |||       :::|||   :::|||
Db      399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||         :::|||   |||:::|||       ||||||   |||
Db      399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db      399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
               :::::::::|||::::::       |||       ||||||
Db      399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy         293 CysGlnTyrThrTrp 297
               :::|||||||:::
Db      399733 ACAGAATATACATTT 399747

RESULT 42
US-10-928-446A-197
; Sequence 197, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 197
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-197

Alignment Scores:
Pred. No.:              1.23e+04        Length:         1080000
Score:                  99.00           Matches:        54
Percent Similarity:     43.6%           Conservative:   44

Best Local Similarity:  24.0%           Mismatches:     93
Query Match:            5.2%            Indels:         35
DB:                     8               Gaps:           9

US-10-757-745-2 (1-362) x US-10-928-446A-197 (1-1080000)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||  |||  ||||||      :::  ||||||:::          :::
Db      399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      ::::::  |||   |||::::::|||      ||||||:::
Db      399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||  :::  :::::::::           |||  |||      :::
Db      399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||           ::::::|||      |||   ::::::
Db      399292 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||  |||      :::  |||                |||||:::  |||
Db      399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::      :::      |||      ||||||:::  |||:::
Db      399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::  |||  |||:::             ||||||  |||:::      :::
Db      399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::|||          |||  |||         :::|||  :::|||
Db      399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                  :::|||         :::|||  |||:::|||     ||||||  |||
Db      399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy         266 -------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||  :::
Db      399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                           :::::::::|||::::::      |||      ||||||
Db      399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy         293 CysGlnTyrThrTrp 297
                  :::||||||:::
Db      399733 ACAGAATATACATTT 399747

RESULT 43
US-10-928-446A-199
; Sequence 199, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 199
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-199

Alignment Scores:
Pred. No.:                1.23e+04        Length:          1080000
Score:                    99.00           Matches:         54
Percent Similarity:       43.6%           Conservative:    44
Best Local Similarity:    24.0%           Mismatches:      93
Query Match:              5.2%            Indels:          35
DB:                       8               Gaps:            9

US-10-757-745-2 (1-362) x US-10-928-446A-199 (1-1080000)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                  :::
Db     399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      ::::::   |||  |||::::::|||       ||||||:::
Db     399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   :::::::::             |||   |||      :::
Db     399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                 ::::::|||       |||   ::::::
Db     399292 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||  |||     :::   |||                    ||||||:::   |||
Db     399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::       :::        |||      ||||||:::   |||:::
Db     399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||  |||:::                ||||||   |||:::      :::
Db     399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||   |||         :::||||  :::|||
Db     399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||      :::|||   |||:::|||     ||||||   |||
Db     399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db     399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db     399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy        293 CysGlnTyrThrTrp 297
               :::||||||:::
Db     399733 ACAGAATATACATTT 399747

RESULT 44
US-10-928-446A-201
; Sequence 201, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
```

```
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 201
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (827008)..(827008)
;   OTHER INFORMATION: the 'n' at position 827008 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (872617)..(872860)
;   OTHER INFORMATION: Exon 2a
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (904952)..(905030)
;   OTHER INFORMATION: Exon 2d
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (922549)..(922630)
;   OTHER INFORMATION: Exon 3
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (926021)..(926059)
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (929123)..(929176)
;   OTHER INFORMATION: Exon 5
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (993104)..(993154)
;   OTHER INFORMATION: Exon 6
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   OTHER INFORMATION: Full exon range is 999547-999608
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1000354)..(1000456)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1002118)..(1002284)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1006117)..(1006249)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1007860)..(1008036)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1010940)..(1011014)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1018160)..(1018291)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1018800)..(1018919)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1020028)..(1020225)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1026659)..(1026736)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1028113)..(1028167)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1034316)..(1034374)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1041390)..(1041455)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1043121)..(1043350)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1044868)..(1044989)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1047519)..(1047589)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1050296)..(1050391)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1060368)..(1060441)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1062648)..(1062708)
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1064561)..(1064620)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1066207)..(1066314)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1067768)..(1067864)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1068609)..(1068681)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1073289)..(1075279)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
US-10-928-446A-201

Alignment Scores:
Pred. No.:                1.23e+04        Length:          1080000
Score:                    99.00           Matches:         54
Percent Similarity:       43.6%           Conservative:    44
Best Local Similarity:    24.0%           Mismatches:      93
Query Match:              5.2%            Indels:          35
DB:                       8               Gaps:            9

US-10-757-745-2 (1-362) x US-10-928-446A-201 (1-1080000)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
                |||   |||    ||||||       :::   ||||||:::                 :::
Db     399122 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 399180

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||       ::::::   |||   |||:::::::|||       ||||||:::
Db     399181 AATTCACACATAACAATATTCACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 399234

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   :::::::::              |||   |||      :::
Db     399235 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 399291

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                    ::::::|||       |||   ::::::
Db     399292 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACACAGGCTCAAA 399333

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                 |||||:::   |||
Db     399334 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 399393

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::   :::          |||       ||||||:::   |||:::
Db     399394 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 399453

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                  ::::::   |||   |||:::           |||||   |||:::      :::
Db     399454 CATTACATAATGGTAAAGGGATCAATTCGACAAGAAGAGCTAACTCTCCTAAATATATAT 399513

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
```

```
                  :::|||           |||   |||       :::|||   :::|||
Db     399514 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAATCCTGAGT 399558

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                ::::|||       :::|||   |||:::|||       ||||||   |||
Db     399559 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 399615

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db     399616 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 399675

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          ::::::::::|||::::::       |||       ||||||
Db     399676 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 399732

Qy        293 CysGlnTyrThrTrp 297
                ::::||||||:::
Db     399733 ACAGAATATACATTT 399747

RESULT 45
US-11-072-512-1677
; Sequence 1677, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1677
;   LENGTH: 2257
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-072-512-1677

Alignment Scores:
Pred. No.:                5.12            Length:          2257
Score:                    98.50           Matches:         77
Percent Similarity:       36.6%           Conservative:    60
Best Local Similarity:    20.6%           Mismatches:      150
Query Match:              5.2%            Indels:          87
DB:                       9               Gaps:            17

US-10-757-745-2 (1-362) x US-11-072-512-1677 (1-2257)

Qy         10 GlyArgGluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCys 29
                ||||:::||||||      :::                 ::::::|||
Db       1273 GGTAAAGAAGCAGAGACTAAAAGTCCTTTAAATGTTTTGAGAAAA--------------- 1317

Qy         30 ValGluPheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGlu 49
```

```
Db      1318 GTAAGCCATAATACAGTCTCTTTGATGGATCATTTATTAAGTGTCCCAGAAACAGTAGAA 1377

Qy        50 AsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSer 69

Db      1378 AAAGAAACAAGTTCTGAACATCATGTAAATGCTGTGTTTCAGAAAACCATT---GAACCA 1434

Qy        70 AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsn 89

Db      1435 CTTTTGAAAGAAGAAACAGAGAATGCTTCAGAGCCGTTAGGTTATGAAAGCATGGCATCG 1494

Qy        90 GluGluThrThrAspSerThrThrSerLysIle------SerProSerGluAspThrGln 107

Db      1495 AAAGAGGATTTTAAATCGATGAAAAGCTTCATAGGGAAATCACCTAATGAGTACCATATT 1554

Qy       108 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn 127

Db      1555 GAAAGGAGATCTTCACGG----------------------------GAAGACTTAAGA 1584

Qy       128 AsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspVal 147

Db      1585 AGTGCATCTGAAGAATTGAAGTTAAGCTGT---------------------------- 1614

Qy       148 IlePheLeuGlnGluValIlePro--------------------ProTyrTyrSerTyr 160

Db      1615 ---------CAGAGAACAATACCTATGACTGGTAAAAGAACTTGGCCCTATTATTCATGT 1665

Qy       161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db      1666 GCTAGAATATCTGCCTGGTGTTGGAAAAAGGCTTCCTTGCCAGAATCAAGTTACTTT--- 1722

Qy       181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db      1723 ---------CTTCGTGGGTCTCAGGAAAGTTGTAGGCAAGTTGATGTC------------ 1761

Qy       201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220

Db      1762 ---CCTAAACACCAAACAAAC------------CAGACCCATTTAACTGACTCCAAATTA 1806

Qy       221 CysLeuMetThrSerHisLeuGlu---SerThrArgGlyHisAlaAlaGluArgMetAsn 239

Db      1807 TTATTACAAAGTTCCTTAACAGAAACAAACACTGAATCTTCAAGTAAAGAAAAATTAGAT 1866

Qy       240 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal------ 257

Db      1867 TCTAATTCTAATTGTTTGTCTTCAGTTTCTGCAGTAGAACCTACTTTAATGGTTATAAAG 1926

Qy       258 ------IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275

Db      1927 GAACCTATAATCAAGGATGATAAAAAGATAAAATCAGAGGAACTGAGCAGAAGAGGGTCA 1986

Qy       276 ------LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCys 293

Db      1987 GAGGTAATTTCTAACACTACTGAAGATACT---------------------------- 2016

Qy       294 GlnTyrThrTrpAspThrGln---MetAsnSerAsn----------------------- 304

Db      2017 CAATTAACCAGTGAGACTCAAAGCTTAACGGGAAATAAAAAAAAAGCTAGAGGAAATTTA 2076

Qy       305 ------LeuGlyIleThrAlaAlaCysLys-----------------LeuArgPheAsp 316

Db      2077 ACGAAACTTAATTTGACAGCGACTTCCAAAGATGGTCAGGAAGCAAATAACTCTGCAGGC 2136

Qy       317 ArgIlePhePheArgAlaAla------AlaGluGluGlyHisIleIleProArgSerLeu 334

Db      2137 AAAACTATTCATCGAAAAGCATGCATTGCTCAACAAACATTTATAGTTCCAGACTTGGTT 2196

Qy       335 AspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348

Db      2197 AAAATATTGAACACAGGACGGCTGACTAATTTTAAAATTCCT 2238

RESULT 46
US-11-122-329-100
```

```
; Sequence 100, Application US/11122329
; Publication No. US20060019272A1
; GENERAL INFORMATION:
;  APPLICANT: Geraci, Mark
;  APPLICANT:  Bull, Todd
;  APPLICANT:  Voelkel, Norbert
;  APPLICANT:  Coldren, Chris
;  TITLE OF INVENTION: Diagnosis of Disease and Monitoring of Therapy Using Gene
;  TITLE OF INVENTION:  Expression Analysis of Peripheral Blood Cells
;  FILE REFERENCE: 2848-54
;  CURRENT APPLICATION NUMBER: US/11/122,329
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: 60/568,129
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 100
;   LENGTH: 11185
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-122-329-100

Alignment Scores:
Pred. No.:                39.8          Length:          11185
Score:                    98.50         Matches:         75
Percent Similarity:       37.2%         Conservative:    53
Best Local Similarity:    21.8%         Mismatches:      133
Query Match:              5.2%          Indels:          83
DB:                       12            Gaps:            17

US-10-757-745-2 (1-362) x US-11-122-329-100 (1-11185)

Qy          7 LeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluProGluValLysLysArgArg 26

Db       6462 GTGGAAGGTACGAAAGCTCCAGTAGAGAAGGAG---------GAAGTAAAGGTCAGTGGC 6512

Qy         27 LeuLeuCysValGluPhe-----------------AlaSerValAlaSerCysAspAla 40

Db       6513 ACAGTTTCAACAAACTTTCCCCAAACTATAGAGCCAGCCAAATTATGGTCTAGGCAAGAA 6572

Qy         41 Ala------ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeu 58

Db       6573 GTCAACCCTGTAAGACAAGAAATTGAAAGTGAAACAACATCAGAGGAACAAATTCAAGAA 6632

Qy         59 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle 78

Db       6633 GAAAAGTCATTTGAATCCCCTCAAAACTCTCCTGCAACAGAA---------CAAACAATC 6683

Qy         79 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp------------ 94

Db       6684 TTTGATTCACAGACATTTACTGAA---ACTGAACTCAAAACCACAGATTATTCTGTACTA 6740

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114

Db       6741 ACAACAAAGAAAACTTACAGTGATGATAAAGAAATGAAGGAGGAAGACACTTCTTTAGTT 6800

Qy        115 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134

Db       6801 AACATGTCTACTCCAGATCCAGAT----------------------------GCAAAT 6830

Qy        135 GlyValCysSerTyrLeuAlaLeuTyrSerProAsp------------------------ 146

Db       6831 GGCTTGGAATCTTACACAACTCTC------CCTGAAGCTACTGAAAAGTCACATTTTTTC 6884

Qy        147 ------ValIlePheLeuGlnGluValIleProProTyrTyr------------SerTyr 160

Db       6885 TTAGCTACTGCATTAGTAACTGAATCTATACCAGCTGAACATGTAGTCACAGATTCACCA 6944

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db       6945 ATCAAAAAGGAAGAAAGTACAAAACATTTTCCGAAAGGCATGAGACCA------------ 6992

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeu-----------------LysSerGln 194
```

```
Db       6993 ------ACAATTCAAGAGTCAGATACTGAGCTCTTATTCTCTGGACTGGGATCAGGAGAA 7046

Qy        195 GluIleIlePro----------------------ProSerThrLysMetMetArg 206

Db       7047 GAAGTTTTACCTACTCTACCAACAGAGTCAGTGAATTTTACTGAAGTGGAACAAATCAAT 7106

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 226

Db       7107 AACACATTATATCCCCAC------------------------------ACTTCTCAA 7133

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db       7134 GTGGAAAGTACCTCAAGTGACAAAATTGAAGACTTTAACAGAATGGAAAATGTGGCAAAA 7193

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 266

Db       7194 GAAGTTGGACCACTC---GTATCTCAAACAGACATCTTTGAAGGTAGTGGGTCAGTAACC 7250

Qy        267 AspArg---------GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 283

Db       7251 AGCACAACATTAATAGAAATTTTAAGTGACACTGGAGCAGAAGGACCCACGGTGGCACCT 7310

Qy        284 TrpGluPhe---------LeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db       7311 CTCCCTTTCTCCACGGACATCGGACATCCTCAAAATCAGACTGTCAGGTGGGCAGAAGAA 7370

Qy        301 MetAsnSerAsn 304

Db       7371 ATCCAGACTAGT 7382
```

RESULT 47
US-11-121-086-76/c
; Sequence 76, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76
; LENGTH: 98862
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-76

Alignment Scores:

Pred. No.:	650	Length:	98862
Score:	98.50	Matches:	53
Percent Similarity:	44.9%	Conservative:	43
Best Local Similarity:	24.8%	Mismatches:	86
Query Match:	5.2%	Indels:	34
DB:	12	Gaps:	8

US-10-757-745-2 (1-362) x US-11-121-086-76 (1-98862)

```
Qy         68 GluSerAlaLeuGluArgArgProGluThrIleSerGlu---------------------- 80

Db      35672 CAGGCCTGCCTTAAAAGAGCTCCTGAAGGAAGCACTAAACATGGAAAGGAACAACCGGTA 35613

Qy         81 ProLysThrTyrValAspLeuThrAsn-----------------GluGluThrThrAsp 94

Db      35612 CCAGCCACTGC-AAAAACATGACAAATTGTAAAGATCATTGAGGCTAGGAAGATACTGCA 35554

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114

Db      35553 TCAACTAATGAGCAAAATAA-CCAGCTAACATCATAATGACAAGATCAAATTCACACATA 35495

Qy        115 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134

Db      35494 ACAATATTAACCTTAAATGTAAATGGGCTA---------AATGCTACAACTAAAAGACAC 35444

Qy        135 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 154

Db      35443 AGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA------ 35390

Qy        155 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 174

Db      35389 -----------ACCCATCTCATGTGCAGAGACATACATAGGCTCGAAATAAAGGGATGG 35342

Qy        175 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194

Db      35341 AGGAAGATCTACCAAGCAAACGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTTTCT 35282

Qy        195 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 214

Db      35281 GATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAAGCCATTACATAGTG 35222

Qy        215 ValSerGlyAsn------------GluLeuCysLeuMetThrSerHisLeuGluSerThr 230

Db      35221 GTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATATGCACCCAATACA 35162

Qy        231 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 250

Db      35161 AGA---------GCACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGA 35117

Qy        251 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 264

Db      35116 GACTTAGACTCCCATACAATAATAATG---GGAGACTTTAAC 35078
```

RESULT 48
US-09-925-065A-124743
; Sequence 124743, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 124743
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-124743

Alignment Scores:

Pred. No.:	0.974	Length:	561
Score:	98.00	Matches:	36
Percent Similarity:	47.1%	Conservative:	30
Best Local Similarity:	25.7%	Mismatches:	50
Query Match:	5.2%	Indels:	24
DB:	6	Gaps:	6

US-10-757-745-2 (1-362) x US-09-925-065A-124743 (1-561)

```
Qy         73 ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 92
```

```
Db        168 AGGGCCTATAAAACAATAACACAATGAAAAACATACCAAAGTATT---------CAAGTA 218

Qy         93 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 112

Db        219 ACAACTAGCATAATGAACAGAATAGCACCTCAT-------------------------- 251

Qy        113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132

Db        252 ---ATCTCAATACTAACGGTGAACGTAAATGGCCTA---------AATGCTCCACTTAAA 299

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152

Db        300 AGACACAGAATGGCAGAATGGATAAGAATTCACCAACCAAGTATCTGTTGTCTTCAAGAA 359

Qy        153 Val---IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171

Db        360 ACTCACCTACCATGTAAGGTCTCACATAAACTTAAAGTAAAGGGATGGAAAGAGATATTT 419

Qy        172 ------ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 189

Db        420 TATGCAAATGGACACCAAAAGCAAGCAAGAGTAGCTATATTCTTATATCAGACAAAACAC 479

Qy        190 LysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu 209

Db        480 ACTTTAAAGCAGCAACARTTA--------------AAAAAGACAAAAAGGGACCTTATA 524

RESULT 49
US-09-925-065A-911490
; Sequence 911490, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 911490
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-911490

Alignment Scores:
Pred. No.:              1.14          Length:         633
Score:                  98.00         Matches:        36
Percent Similarity:     47.1%         Conservative:   30
Best Local Similarity:  25.7%         Mismatches:     50
Query Match:            5.2%          Indels:         24
DB:                     6             Gaps:           6

US-10-757-745-2 (1-362) x US-09-925-065A-911490 (1-633)

Qy         73 ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 92

Db        209 AGGGCCTATAAAACAATAACACAATGAAAAACATACCAAAGTATT---------CAAGTA 259

Qy         93 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 112

Db        260 ACAACTAGCATAATGAACAGAATAGCACCTCAT-------------------------- 292

Qy        113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132

Db        293 ---ATCTCAATACTAACGGTGAACGTAAATGGCCTA---------AATGCTCCACTTAAA 340

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152

Db        341 AGACACAGAATGGCAGAATGGATAAGAATTCACCAACCAAGTATCTGTTGTCTTCAAGAA 400

Qy        153 Val---IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171

Db        401 ACTCACCTACCATGTAAGGTCTCACATAAACTTAAAGTAAAGGGATGGAAAGAGATATTT 460

Qy        172 ------ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 189

Db        461 TATGCAAATGGACACCAAAAGCAAGCAAGAGTAGCTATATTCTTATATCAGACAAAACAC 520

Qy        190 LysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu 209

Db        521 ACTTTAAAGCAGCAACAGTTA--------------AAAAAGACAAAAAGGGACCTTATA 565

RESULT 50
US-09-925-065A-911273
; Sequence 911273, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 911273
;   LENGTH: 639
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-911273

Alignment Scores:
Pred. No.:              1.15          Length:         639
Score:                  98.00         Matches:        36
Percent Similarity:     47.1%         Conservative:   30
Best Local Similarity:  25.7%         Mismatches:     50
Query Match:            5.2%          Indels:         24
DB:                     6             Gaps:           6

US-10-757-745-2 (1-362) x US-09-925-065A-911273 (1-639)

Qy         73 ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 92

Db        215 AGGGCCTATAAAACAATAACACAATGAAAAACATACCAAAGTATT---------CAAGTA 265

Qy         93 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 112

Db        266 ACAACTAGCATAATGAACAGAATAGCACCTCAT-------------------------- 298

Qy        113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132

Db        299 ---ATCTCAATACTAACGGTGAACGTAAATGGCCTA---------AATGCTCCACTTAAA 346

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
```

```
Db        347 AGACACAGAATGGCAGAATGGATAAGAATTCACCAACCAAGTATCTGTTGTCTTCAAGAA 406

Qy        153 Val---IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171

Db        407 ACTCACCTACCATGTAAGGTCTCACATAAACTTAAAGTAAAGGGATGGAAAGAGATATTT 466

Qy        172 ------ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 189

Db        467 TATGCAAATGGACACCAAAAGCAAGCAAGAGTAGCTATATTCTTATATCAGACAAAACAC 526

Qy        190 LysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu 209

Db        527 ACTTTAAAGCAGCAACAGTTA--------------AAAAAGACAAAAAGGGACCTTATA 571

RESULT 51
US-10-821-234-746
; Sequence 746, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
;  APPLICANT: Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Andarmani, Susan
;  APPLICANT:  Tang, Y. Tom
;  TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;  FILE REFERENCE: 821A
;  CURRENT APPLICATION NUMBER: US/10/821,234
;  CURRENT FILING DATE:  2004-04-07
;  PRIOR APPLICATION NUMBER: US 60/462,047
;  PRIOR FILING DATE: 2003-04-07
;  NUMBER OF SEQ ID NOS: 1704
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 746
;   LENGTH: 3828
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-821-234-746

Alignment Scores:
Pred. No.:              11.4          Length:          3828
Score:                  98.00         Matches:         49
Percent Similarity:     44.2%         Conservative:    42
Best Local Similarity:  23.8%         Mismatches:      81
Query Match:            5.2%          Indels:          34
DB:                     8             Gaps:            9

US-10-757-745-2 (1-362) x US-10-821-234-746 (1-3828)

Qy        110 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 129

Db         10 TCAAATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCA 63

Qy        130 SerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePhe 149

Db         64 ATTAAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGT 120

Qy        150 LeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGlu 169

Db        121 ATTCAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTC 162

Qy        170 IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 189

Db        163 AAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTT 222

Qy        190 LysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu 209

Db        223 GCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAA 282

Qy        210 CysValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSer 225

Db        283 GGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATT 342

Qy        226 HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeu 245

Db        343 TATGCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTC 387

Qy        246 LysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu 265

Db        388 AGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACC 444

Qy        266 ----------------------------------------ArgAspArgGluValThr 271

Db        445 CCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTG 504

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291

Db        505 AACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAA 561

Qy        292 HisCysGlnTyrThrTrp 297

Db        562 TCAACAGAATATACCTTT 579

RESULT 52
US-11-121-086-37/c
; Sequence 37, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 37
;   LENGTH: 184000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-37

Alignment Scores:
Pred. No.:              1.63e+03      Length:          184000
Score:                  98.00         Matches:         60
Percent Similarity:     43.2%         Conservative:    51
Best Local Similarity:  23.3%         Mismatches:      99
Query Match:            5.2%          Indels:          49
DB:                     12            Gaps:            10

US-10-757-745-2 (1-362) x US-11-121-086-37 (1-184000)

Qy         72 GluArgArgProGluThrIleSerGlu----------ProLysThrTyr 84

Db      66941 AAAAGAGCTCCTGAAGGAAGCGCTAAACATGGAAACGAACAACCGGTACCAGCCACTGCA 66882

Qy         85 ValAspLeuThrAsnGluGluThrThrAspSer----------ThrThrSer 98

Db      66881 AAATCATGCCAAAATGTAAAGACCATCGAGACTAGGAAGAAACTGCATCAACTAACAAGC 66822

Qy         99 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 118

Db      66821 AAGATAA--CCAGCTAACATCATAATGACAGGATCAAATTCACACATAACAATATTAACT 66764

Qy        119 TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 138

Db      66763 TTCAATGTAAATGGACTA---------AATGCTTCAATTAAAAGACACAGACTGGCAAAT 66713

Qy        139 TyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 158

Db      66712 TGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA---------- 66671

Qy        159 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178

Db      66670 ACCCATCTCACGTGCAGAGACACACATAGGCTCAAAATCAAAGGATA-GAGGAAGATCTA 66612
```

```
Qy        179 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198
Db      66611 CCAAGCAAATGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGCCTCTGATAAAACAGAC 66552

Qy        199 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 218
Db      66551 TTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCA 66492

Qy        219 ------------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 234
Db      66491 ATTCAACAAGAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGA 66441

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
Db      66440 GCATCCAGATTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCC 66387

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeu--------------------------- 265
Db      66386 CACACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACA 66330

Qy        266 --------------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
Db      66329 AGACAGAAAGTCAACAAGGATAACCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTA 66270

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
Db      66269 ATAGACATCTACGGAACTCTC---CACCCCAAATCAACAGAATATACATTT 66222

RESULT 53
US-11-121-086-13/c
; Sequence 13, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
;   LENGTH: 191797
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-13

Alignment Scores:
Pred. No.:              1.72e+03        Length:         191797
Score:                  98.00           Matches:        55
Percent Similarity:     44.0%           Conservative:   44
Best Local Similarity:  24.4%           Mismatches:     92
Query Match:            5.2%            Indels:         35
DB:                     12              Gaps:           9

US-10-757-745-2 (1-362) x US-11-121-086-13 (1-191797)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db      80922 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 80864

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db      80863 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTCCAATT 80810

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      80809 AAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGATCCATCAGTGTGCTGTATT 80753

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db      80752 CAGGAA----------------ACCCATCTCACATGCAGAGACACACATAGGCTCAAA 80711

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db      80710 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 80651

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db      80650 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 80591

Qy        211 ValHisValAsnValSerGlyAsn----------GluLeuCysLeuMetThrSerHis 226
Db      80590 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATAT 80531

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db      80530 GCACCCAATACG---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 80486

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db      80485 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 80429

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db      80428 CTGTCAACATTAGACAGATCAACAAGACAGAAAGTCAACAAGGATATCCAGGAATTGAAC 80369

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db      80368 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACACTACTCTC---CACCCCAAATCA 80312

Qy        293 CysGlnTyrThrTrp 297
Db      80311 ACAGAATATACATTT 80297

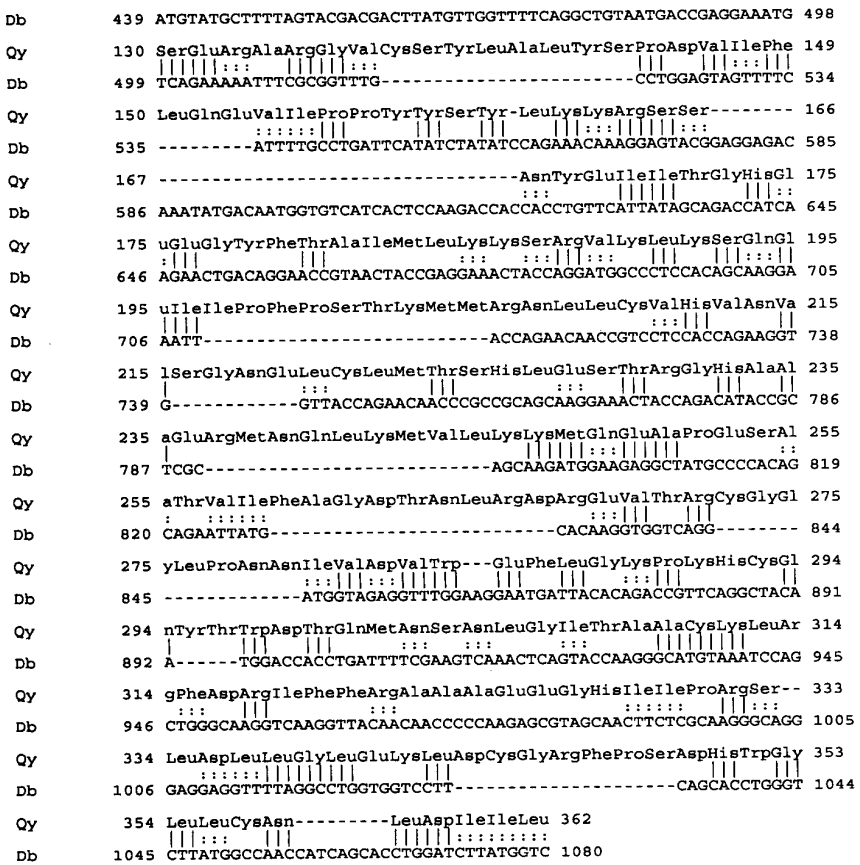
RESULT 54
US-11-096-568A-22794
; Sequence 22794, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 22794
;   LENGTH: 1466
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1466)
;   OTHER INFORMATION: Ceres Seq. ID no. 12410153
US-11-096-568A-22794

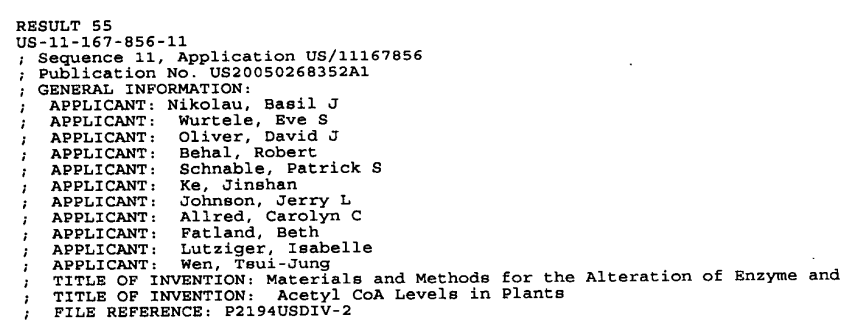
Alignment Scores:
Pred. No.:              3.78            Length:         1466
Score:                  97.50           Matches:        67
Percent Similarity:     38.4%           Conservative:   46
Best Local Similarity:  22.8%           Mismatches:     101
Query Match:            5.1%            Indels:         80
DB:                     9               Gaps:           13

US-10-757-745-2 (1-362) x US-11-096-568A-22794 (1-1466)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db        379 GAGACATACCTCCAGACTCTCGCCAAGGTCGTCGGGAGTTATGAGGAGGCCAAGAAGAGG 438

Qy        111 GlySerMetPheSerLeuIleThrTrp---AsnIleAspGlyLeuAspLeuAsnAsnLeu 129
```



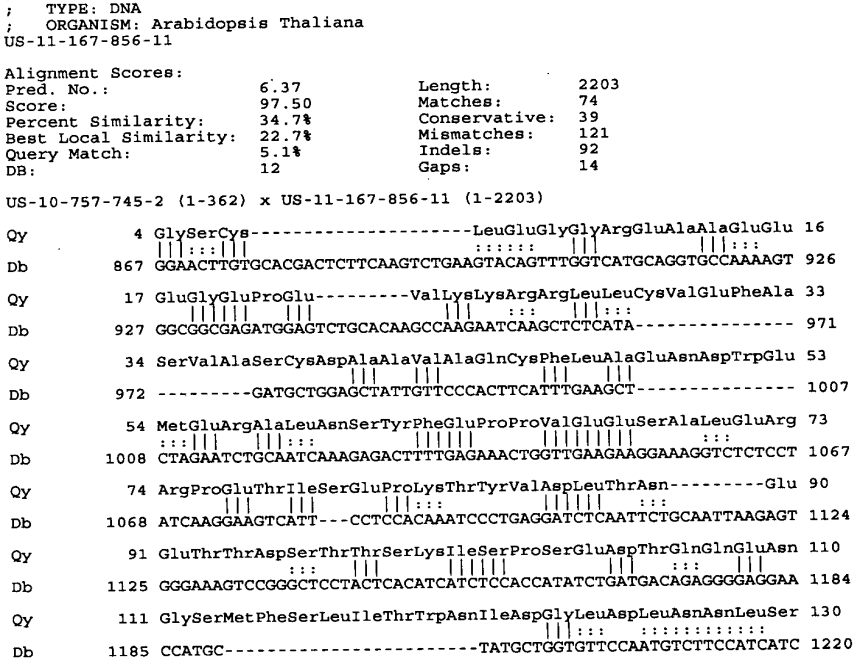


RESULT 55
US-11-167-856-11
; Sequence 11, Application US/11167856
; Publication No. US20050268352A1
; GENERAL INFORMATION:
; APPLICANT: Nikolau, Basil J
; APPLICANT: Wurtele, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
; APPLICANT: Ke, Jinshan
; APPLICANT: Johnson, Jerry L
; APPLICANT: Allred, Carolyn C
; APPLICANT: Fatland, Beth
; APPLICANT: Lutziger, Isabelle
; APPLICANT: Wen, Tsui-Jung
; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
; TITLE OF INVENTION: Acetyl CoA Levels in Plants
; FILE REFERENCE: P2194USDIV-2
; CURRENT APPLICATION NUMBER: US/11/167,856
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: US 10/293,865
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/344,882
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,717
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 11
; LENGTH: 2203
; TYPE: DNA
; ORGANISM: Arabidopsis Thaliana
US-11-167-856-11

Alignment Scores:

Pred. No.:	6.37	Length:	2203
Score:	97.50	Matches:	74
Percent Similarity:	34.7%	Conservative:	39
Best Local Similarity:	22.7%	Mismatches:	121
Query Match:	5.1%	Indels:	92
DB:	12	Gaps:	14

US-10-757-745-2 (1-362) x US-11-167-856-11 (1-2203)



```
Qy        219 uLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAs 239
Db       1489 ------------ATCTCACACCTT----------------------------ATGAA 1505

Qy        239 nGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePh 259
Db       1506 TTTGTTGAGGGAATGAAGAAAAAGGGAATCCGAGTCCCCGGGATTGGACACAGGATCAAG 1565

Qy        259 eAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAs 279
Db       1566 AGCAGAGACAACAGAGACAAAAGAG---------------------------------- 1590

Qy        279 nIleValAspValTrpGluPhe------------LeuGlyLysProLysHisCysGlnTy 295
Db       1591 ------------TGGAGCTTCTTCAGAAATTTGCTCGGTCCAACTTCCCATCAGTGAAG 1637

Qy        295 rThrTrpAspThrGln 300
Db       1638 TACATGGAGTACGCAG 1653

RESULT 56
US-10-932-182A-76957
; Sequence 76957, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76957
;   LENGTH: 2991
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-76957

Alignment Scores:
Pred. No.:              9.42          Length:          2991
Score:                  97.50         Matches:         70
Percent Similarity:     40.1%         Conservative:    63
Best Local Similarity:  21.1%         Mismatches:      114
Query Match:            5.1%          Indels:          85
DB:                     7             Gaps:            16

US-10-757-745-2 (1-362) x US-10-932-182A-76957 (1-2991)

Qy         23 LysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaVal 42
Db        835 AAGAAGGACAAAATAATCATAATTGAT-----------TCTGGTATGGAAGTTTGGGTC 882

Qy         43 AlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPhe 62
Db        883 TTGAAGGCTACAACTAAGGAAAATTGGCAGTCGTGGGTCGATGCGTTACAAACTTGTTTT 942

Qy         63 GluProProValGluGlu---------SerAlaLeuGluArgArgProGluThrIleSer 79
Db        943 GACGATCAGTTCGAAGATAAGGACACGTCCACTTTGGAAGAAAATCCCGACATTCTTGAC 1002

Qy         80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp--------------- 94
Db       1003 GATGACAAGGAAGTAATTAATAAAAGCTCACCTCAAGACCACGACCACCTCACGCCCACG 1062

Qy         95 ---SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 113
Db       1063 GCTACGACCAAAAGTGCACTATCACATAGACAGCATACTCAGAAAGAC------------ 1110

Qy        114 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 133
Db       1111 ---------------------ATGGATGACATATATGTCCCATTGCCCAGCGAA------ 1143

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
Db       1144 ------------TCTTAC---GCTACTTTCTCCATGAATCTACGTTTGATTCAACAACGA 1188

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
Db       1189 CTG------------GAGCAGTGTAAAAAAGACTCGTTATCTTATAAACCAACTACTTTA 1236

Qy        174 HisGlu------GluGly------------TyrPheThrAlaIleMetLeuLysLysSer 187
Db       1237 CATCAAAGATCGGAGGGGCTAAATGGAACACATTCGTCATCTTCCGTCTTTACCAATAAT 1296

Qy        188 ArgVal-----------------------------------------LysLeuLysSer 193
Db       1297 AGAGTGTCTTCATTCAATCATTCTTCTTCTGGTATGACGTCATCTGATTCTTTAGCCTCT 1356

Qy        194 GlnGluIleIleProPheProSerThrLys-----------------MetMetArgAsn 207
Db       1357 GAAGAGGTT---------CCTTCCAACAAAACATATATTGAGCATGCTTTATATAACCAA 1407

Qy        208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
Db       1408 TTGGCAGATCTTGAAGTATTTGTTAGC----------------------------CGG 1437

Qy        228 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 247
Db       1438 TTCGTTACACAGGGAGAGGTTTTATTCAAGGACCACCAGATC------CTGTGCAAGAAA 1491

Qy        248 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 267
Db       1492 GCAAAAGACACGAGAGTTTCTTTAACTTCATATCTTAGCGAGAATGATGAATTTTTTGAT 1551

Qy        268 ArgGluValThrArgCysGlyGly---------LeuPro-----------------Asn 278
Db       1552 GCAGAAGAGGAAATCAGTCGGGGAGTCATTATATTACCTGACACAGAAGATGATATTAAT 1611

Qy        279 AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAsp 298
Db       1612 AATATAGTCGAGGAAACTCCCCTTCTTGGTAAAAGTGACCAAAATGAGTTCACAACAGAG 1671

Qy        299 ThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 310
Db       1672 GTCCAATTGTCGGGATCCGAACAGATAGCTTCATCA 1707

RESULT 57
US-10-932-182A-76957
; Sequence 76957, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76957
;   LENGTH: 2991
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-76957

Alignment Scores:
Pred. No.:              9.42          Length:          2991
Score:                  97.50         Matches:         70
```

Percent Similarity:	40.1%	Conservative:	63
Best Local Similarity:	21.1%	Mismatches:	114
Query Match:	5.1%	Indels:	85
DB:	7	Gaps:	16

US-10-757-745-2 (1-362) x US-10-932-182A-76957 (1-2991)

```
Qy        23 LysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaVal 42
Db       835 AAGAAGGACAAAATAATCATAATTGAT-----------TCTGGTATGGAAGTTTGGGTC 882

Qy        43 AlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPhe 62
Db       883 TTGAAGGCTACAACTAAGGAAAATTGGCAGTCGTGGGTCGATGCGTTACAAACTTGTTTT 942

Qy        63 GluProProValGluGlu---------SerAlaLeuGluArgArgProGluThrIleSer 79
Db       943 GACGATCAGTTCGAAGATAAGGACACGTCCACTTTGGAAGAAAATCCCGACATTCTTGAC 1002

Qy        80 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp--------------- 94
Db      1003 GATGACAAGGAAGTAATTAATAAAAGCTCACCTCAAGACCACGACCACCTCACGCCCACG 1062

Qy        95 ---SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 113
Db      1063 GCTACGACCAAAAGTGCACTATCACATAGACAGCATACTCAGAAAGAC------------ 1110

Qy       114 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 133
Db      1111 --------------------ATGGATGACATATATGTCCCATTGCCCAGCGAA------ 1143

Qy       134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
Db      1144 ------------TCTTAC---GCTACTTTCTCCATGAATCTACGTTTGATTCAACAACGA 1188

Qy       154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
Db      1189 CTG------------GAGCAGTGTAAAAAAGACTCGTTATCTTATAAACCAACTACTTTA 1236

Qy       174 HisGlu------GluGly------------TyrPheThrAlaIleMetLeuLysLysSer 187
Db      1237 CATCAAAGATCGGAGGGGCTAAATGGAACACATTCGTCATCTTCCGTCTTTACCAATAAT 1296

Qy       188 ArgVal-----------------------------------------LysLeuLysSer 193
Db      1297 AGAGTGTCTTCATTCAATCATTCTTCTTCTGGTATGACGTCATCTGATTCTTTAGCCTCT 1356

Qy       194 GlnGluIleIleProPheProSerThrLys-----------------MetMetArgAsn 207
Db      1357 GAAGAGGTT---------CCTTCCAACAAAACATATATTGAGCATGCTTTATATAACCAA 1407

Qy       208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
Db      1408 TTGGCAGATCTTGAAGTATTTGTTAGC---------------------------CGG 1437

Qy       228 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 247
Db      1438 TTCGTTACACAGGGAGAGGTTTTATTCAAGGACCACCAGATC------CTGTGCAAGAAA 1491

Qy       248 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 267
Db      1492 GCAAAAGACACGAGAGTTTCTTTAACTTCATATCTTAGCGAGAATGATGAATTTTTTGAT 1551

Qy       268 ArgGluValThrArgCysGlyGly---------LeuPro-----------------Asn 278
Db      1552 GCAGAAGAGGAAATCAGTCGGGGAGTCATTATATTACCTGACACAGAAGATGATATTAAT 1611

Qy       279 AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAsp 298
Db      1612 AATATAGTCGAGGAAACTCCCCTTCTTGGTAAAAGTGACCAAAATGAGTTCACAACAGAG 1671

Qy       299 ThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 310
```

```
Db      1672 GTCCAATTGTCGGGATCCGAACAGATAGCTTCATCA 1707
```

RESULT 58
US-09-925-065A-918150
; Sequence 918150, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 918150
; LENGTH: 791
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-918150

Alignment Scores:			
Pred. No.:	1.94	Length:	791
Score:	97.00	Matches:	50
Percent Similarity:	41.3%	Conservative:	43
Best Local Similarity:	22.2%	Mismatches:	98
Query Match:	5.1%	Indels:	35
DB:	6	Gaps:	7

US-10-757-745-2 (1-362) x US-09-925-065A-918150 (1-791)

```
Qy        91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db       127 GAAGAAACTGCATCAACTAACRAGCAAAA-CAACCAGCTAACATCATAATGACAGGATCA 185

Qy       111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db       186 AATTCACACATAACAATATTAACTTTAAATGTAAATGGGCTA---------AATGCTCCA 236

Qy       131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db       237 ATTAAAAGACACAGACTGGCAAATTGGATAAGGAGTCAAGACCCATCAGTGTGCTGTATT 296

Qy       151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db       297 CAGGAA-----------------ACCCATCTCAAGTGCAGAGACACACATAGGCTCAAA 338

Qy       171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db       339 ATAAAGGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 398

Qy       191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db       399 ATCCTACTCTCTGATAGAACAGATTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 458

Qy       211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db       459 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATAT 518

Qy       227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db       519 GCACCCAATACA--------------GGAGAACCCAGATTCATAAAGCAAGTCCTGAGT 563
```

```
Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db        564 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCCCA 620

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db        621 CTGTCAACATTAGATAGATCAATGAGACAGAAAGTTAATAAGGATACCCAGGAATTGAAC 680

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db        681 TCAGCTCTGCACCAAGCGGAACTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 737

Qy        293 CysGlnTyrThrTrp 297
Db        738 ACAGAATATACATTT 752

RESULT 59
US-09-925-065A-944994
; Sequence 944994, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 944994
;   LENGTH: 791
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-944994

Alignment Scores:
Pred. No.:              1.94          Length:          791
Score:                  97.00         Matches:         50
Percent Similarity:     41.3%         Conservative:    43
Best Local Similarity:  22.2%         Mismatches:      98
Query Match:            5.1%          Indels:          35
DB:                     6             Gaps:            7

US-10-757-745-2 (1-362) x US-09-925-065A-944994 (1-791)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db        127 GAAGAAACTGCATCAACTAACGAGCAAAA-CAACCAGCTAACATCATAATGACAGGATCA 185

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db        186 AATTCACACATAACAATATTAACTTTAAATGTAAATGGGCTA---------AATGCTCCA 236

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db        237 ATTAAAAGACACAGACTGGCAAATTGGATAAGGAGTCAAGACCCATCAGTGTGCTGTATT 296

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db        297 CAGGAA-----------------ACCCATCTCAAGTGCAGAGACACACATAGGCTCAAA 338

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db        339 ATAAAGGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 398

Qy        191 LeuLysSerGlnGluIleIleProProSerThrLysMetMetArgAsnLeuLeuCys 210
Db        399 ATCCTACTCTCTGATAGAACAGATTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 458

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db        459 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATAT 518

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db        519 GCACCCAATACA--------------GGAGAACCCAGATTCATAAAGCAAGTCCTGAGT 563

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db        564 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCCCA 620

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db        621 CTGTCAACATTAGATAGATCAATGAGACAGAAAGTTAATAAGGATACCCAGGAATTGAAC 680

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db        681 TCAGCTCTGCACCAAGCGGAACTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 737

Qy        293 CysGlnTyrThrTrp 297
Db        738 ACAGAATATACATTT 752

RESULT 60
US-10-330-773-114/c
; Sequence 114, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114
;   LENGTH: 440684
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(440684)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-114

Alignment Scores:
Pred. No.:              6.4e+03       Length:          440684
Score:                  97.00         Matches:         46
Percent Similarity:     42.7%         Conservative:    42
Best Local Similarity:  22.3%         Mismatches:      84
Query Match:            5.1%          Indels:          34
DB:                     7             Gaps:            8

US-10-757-745-2 (1-362) x US-10-330-773-114 (1-440684)

Qy        110 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 129
Db     253576 TCAAATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCT 253526

Qy        130 SerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePhe 149
Db     253525 CCAATTAAAAGACACAGACTGGCAAATTGGATAAACGGTCAAGACCCATCAGTGTGCTGT 253466

Qy        150 LeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGlu 169
```

```
Db     253465 ATTCAGGAA-----------------ACCCATCTCACGTGCAGAGACACAAATAGGCTC 253424
Qy        170 IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 189
Db     253423 AAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTT 253364
Qy        190 LysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg--------- 206
Db     253363 GCAATCCTAGTGTCTGATAAAACAGACTTTAAACCAACAAAGGTCAAAAGAGAGAAAGAA 253304
Qy        207 ---AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSer 225
Db     253303 GGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATA 253244
Qy        226 HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeu 245
Db     253243 TATGCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTG 253199
Qy        246 LysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu 265
Db     253198 AGGGACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACC 253142
Qy        266 ----------------------------------------ArgAspArgGluValThr 271
Db     253141 CCACTGTCAACATTAGACAGATCAACGAAACAGAAAGTTAACAAGGATATCCAGGAATTG 253082
Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
Db     253081 AACTCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTG---CACCCCAAA 253025
Qy        292 HisCysGlnTyrThrTrp 297
Db     253024 TCAACAGAATATACATTT 253007
```

RESULT 61
US-11-121-086-101/c
; Sequence 101, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 101
; LENGTH: 185393
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-101

Alignment Scores:

Pred. No.:	2.39e+03	Length:	185393
Score:	96.50	Matches:	53
Percent Similarity:	42.7%	Conservative:	44
Best Local Similarity:	23.3%	Mismatches:	96
Query Match:	5.1%	Indels:	35
DB:	12	Gaps:	8

US-10-757-745-2 (1-362) x US-11-121-086-101 (1-185393)

```
Qy         89 AsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGln 108
Db      18726 GACGAAGAAGAAACTGCATCAACTAACGAGCAAAATAA-CCAGCTAACATCATAATGACA 18668
Qy        109 GluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsn 128
Db      18667 GGATCAAATTCACACATAACAATATTAACTTTAAATATAAATGGACTA---------AAT 18617
```

```
Qy        129 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 148
Db      18616 GCTCCAATTAAAAGACACAGACTGGAAAATTGGATAAAAAGTCAAGACCCATCAGTGTGC 18557
Qy        149 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 168
Db      18556 TGTATTCAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGG 18515
Qy        169 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 188
Db      18514 CTCAAAATAAAAGGATGGAGGAAGATCTACCCAGCAAATGGAAAACAAAAAAAGGCAGGG 18455
Qy        189 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 208
Db      18454 GTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAA 18395
Qy        209 LeuCysValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThr 224
Db      18394 GAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAAT 18335
Qy        225 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 244
Db      18334 ATATATGCAGCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTC 18290
Qy        245 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 264
Db      18289 CTGAGTGACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAAC 18233
Qy        265 Leu----------------------------------------ArgAspArgGluVal 270
Db      18232 ACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTTAACAAGGATACCCAGGAA 18173
Qy        271 ThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 290
Db      18172 TTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCC 18116
Qy        291 LysHisCysGlnTyrThrTrp 297
Db      18115 AAATCAACAGAATATACATTT 18095
```

RESULT 62
US-10-995-561-13233/c
; Sequence 13233, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13233
; LENGTH: 199130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(199130)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13233

Alignment Scores:

Pred. No.:	2.62e+03	Length:	199130
Score:	96.50	Matches:	55
Percent Similarity:	46.3%	Conservative:	50
Best Local Similarity:	24.2%	Mismatches:	86
Query Match:	5.1%	Indels:	37
DB:	8	Gaps:	8

US-10-757-745-2 (1-362) x US-10-995-561-13233 (1-199130)

```
Qy        91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db     87629 GAAGAAACTGAATCAACTAACGAGCAAAATAA-CCAGCTAACATCAAAATGACAAGATCA 87571

Qy       111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db     87570 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTCCA 87520

Qy       131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db     87519 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 87460

Qy       151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db     87459 CAGGAA-----------------ACCCATCTCACGTGCAGAAACACACATAGCCTCAAA 87418

Qy       171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db     87417 ATAAAAGGATGGAGGAAGATCTACCAAGCAAGTGGAAAACAAAAAAAGGCAGGGGTTGCA 87358

Qy       191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db     87357 ATCCTAGTCTCTGATAGAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 87298

Qy       211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db     87297 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATAT 87238

Qy       227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     87237 GCACCCAATACAGGAGCACCCAGTTTC--------------ATAAAGCAAGTCCTGAGT 87193

Qy       247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp------------ 262
Db     87192 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCCCA 87136

Qy       263 ------------ThrAsnLeuArgAspArgGluValThrArgCys-GlyGlyLeu---- 276
Db     87135 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTTAACAAGGATACTCAGGAATTGAAC 87076

Qy       277 -----------ProAsn-AsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     87075 TCAGCTCTGCACCAAGCGGACCTAATAGACATCCACAGAACTCTC---CACCCCAAATCA 87019

Qy       293 CysGlnTyrThrTrp 297
Db     87018 ACAGAATATACATTT 87004

RESULT 63
US-10-330-773-398/c
; Sequence 398, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 398
;   LENGTH: 228835
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-398

Alignment Scores:
Pred. No.:              3.55e+03        Length:         228835
Score:                  96.00           Matches:        48
Percent Similarity:     42.2%           Conservative:   49
Best Local Similarity:  20.9%           Mismatches:     91

Query Match:            5.0%            Indels:         42
DB:                     7               Gaps:           9

US-10-757-745-2 (1-362) x US-10-330-773-398 (1-228835)

Qy       110 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 129
Db     91266 TCAAATTCACACATAACAATATTAATCTTAAGTGAAAATGGGCTA---------AATGCC 91216

Qy       130 SerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePhe 149
Db     91215 CCAGTTAAAAGTCACGGACTGGCAAATTGGATAAAGAGTCAAGACCCATGGGTGTGCTGT 91156

Qy       150 LeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGlu 169
Db     91155 ATTCAGGAG-----------------ACCCATTTCATGTGCAAAGACGCACATAGGCTC 91114

Qy       170 IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 189
Db     91113 AAAATAAAGGGATGGAGGAAGATCTACCAAACAAATGGAAAGCAAAAAAAAGCAGGGGTT 91054

Qy       190 LysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu 209
Db     91053 ACAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAA 90994

Qy       210 CysValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSer 225
Db     90993 GGCCACCACATAATGGTAAGGGGATCAATTCACCAAGAAGAGCTAACTATCCTAAATATA 90934

Qy       226 HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeu 245
Db     90933 TATGCACCCAATACA---------GGAGCAACCAGATTC------ATAAAGCAAGTCCTT 90889

Qy       246 LysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly------------ 261
Db     90888 AGAGATCTACAAAGAGACTTAGACTCCCACACAATAATAATGGGAGACTTTAAAACCCCA 90829

Qy       262 ------------AspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro 277
Db     90828 CTGTCAGTATTAGACAGATCAATGAGACAGAAGGTTAACAAGGATATCGAGGGCTTGAAC 90769

Qy       278 Asn--------------AsnIleValAspValTrpGluPheLeuGlyLysProLys--- 291
Db     90768 TCAGCTCTGCACCAAGTAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 90712

Qy       292 --------------------------HisCysGlnTyrThrTrpAspThrGlnMetAsn 302
Db     90711 ACAGAATATACATTCTTCTCAGCACCACATTGCACTTATTCTAAAATTGACCACTTAGTT 90652

Qy       303 SerAsnLeuGlyIleThrAlaAlaCysLys 312
Db     90651 GGAAGTAAAGCACTCCTCAGCAAATGTAAA 90622

RESULT 64
US-11-096-568A-29182
; Sequence 29182, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29182
;   LENGTH: 981
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(981)
;   OTHER INFORMATION: Ceres Seq. ID no. 4362003
```

US-11-096-568A-29182

Alignment Scores:

Pred. No.:	3.71	Length:	981
Score:	95.50	Matches:	67
Percent Similarity:	37.7%	Conservative:	54
Best Local Similarity:	20.9%	Mismatches:	105
Query Match:	5.0%	Indels:	95
DB:	9	Gaps:	17

US-10-757-745-2 (1-362) x US-11-096-568A-29182 (1-981)

```
Qy        100 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 119
Db         55 ATTGTCCCTAGTCACAAGCCTCATCGTAGTAACGGC---AGATTTCGAGTCGTCTCTTAC 111

Qy        120 AsnIleAspGlyLeuAspLeuAsnAsnLeuSer---------------------------- 130
Db        112 AACACCCTAGCTCATCGGATGGCGTCGTCGTCTCTCATGTATGTCTCTCCTCCAGAGTCT 171

Qy        131 ------------GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAsp 146
Db        172 ATCCTGTGGGAGAAACGTTCGAAGGCGATTCTTGATAATATAAAGAATTTTGAAGCTGAT 231

Qy        147 ValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys------Arg 164
Db        232 TTCATCTGTCTACAAGAAGTA---GACGAGTACCATAGTTTTTTCGATAGAAACATGGAA 288

Qy        165 SerSerAsnTyrGlu---IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 183
Db        289 GCTCAAGGATACTCTGGAATTCCTATTGAAAATAAAGAAGGCTATGAATGCGCAATCTTT 348

Qy        184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 203
Db        349 TTCAAGCCC---------------AAATTTGCAGAGTTCATTACATACCAAACAACCAGA 393

Qy        204 MetMet------------ArgAsnLeuLeuCysValHisValAsn--------------- 214
Db        394 ATCCAAGGATACACCAAATATGAAAACCTTTGCGTGGCGCCGAGTAGTAGTACCGTAAGT 453

Qy        215 ------------------ValSerGlyAsnGluLeuCysLeuMet---------------- 223
Db        454 AGTGAGTCTAGTGACGTTGTTAACGCTGAAGAGCTTTCTGTGGTAATGGTAGCGTTCAAG 513

Qy        224 -----------------------------------ThrSerHisLeuGluSerThrArg 231
Db        514 ATTTTAAAACCATTTAATCATGTCGTTATTATTGCAAGCTCTCATTTAAAATCCGGGAAA 573

Qy        232 GlyHisAlaAlaGlu------ArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 249
Db        574 CCTGATCGATGGGATGATTTAAAGCTAGCTCAAGTCAAGACTCTAATGACTGAATTAGCC 633

Qy        250 GluAlaProGlu----------------------SerAlaThrValIlePheAlaGly 261
Db        634 TCGTTCAAGGAGATTATATCAGCGCTAACAAATTGCAGTCCATCGGTGATCCTTGCTGGT 693

Qy        262 AspThrAsnLeuArgAsp------ArgGluValThrArgCysGlyGlyLeuProAsnAsn 279
Db        694 GACTTCAATTCGAAACCCTACGTGCATAAATACATAAATTCAGACAATATACCGTCTGAT 753

Qy        280 Ile------ValAspValTrpGluPheLeu---GlyLysProLysHisCysGln------ 294
Db        754 ATTGATCTACGAAGTGTGTATGAATTTACGAAAGGAGAGCCAAGGTTCACAAATAATGTG 813

Qy        295 ------------TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 310
Db        814 CCAGGGTTTGCTGAAACACTCGATTACATGTTCTACACGCATTCTGAAATTATTTCTCCT 873

Qy        311 CysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 330
Db        874 GTGAAGCTC--------------------CTGGATTCGCCGGACGAAGTCGATTTCTTA 912

Qy        331 ProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAsp 350
```

```
Db        913 CCCAATGAAATT----------------------------------CATCCGAGTGAT 936

Qy        351 His 351
Db        937 CAT 939
```

RESULT 65
US-10-330-773-969/c
; Sequence 969, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 969
; LENGTH: 74828
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(74828)
; OTHER INFORMATION: n = A,T,C or G
US-10-330-773-969

Alignment Scores:

Pred. No.:	959	Length:	74828
Score:	95.50	Matches:	52
Percent Similarity:	41.8%	Conservative:	42
Best Local Similarity:	23.1%	Mismatches:	96
Query Match:	5.0%	Indels:	36
DB:	7	Gaps:	8

US-10-757-745-2 (1-362) x US-10-330-773-969 (1-74828)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db      30631 GAAGAAACTGCATCAACTAACGAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 30573

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db      30572 AATTCACACATAACAATAATAACCTTAAATGTAAATGGGCTA---------AATGCTCCA 30522

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      30521 ATTAAAAGGCACAGACTGGCAAACTGGATGAAGAGTCAAGACCCATCAGTGTGCTGTATT 30462

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db      30461 CAGGAA-----------------ACACATCTCACGTGCAGAGACACATATAGGCTCAAA 30420

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db      30419 ATAAAGGGATGGAGGAAGATCTACTAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 30360

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db      30359 ATCCTAGTCTCGGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 30300

Qy        211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
Db      30299 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAACTAACTATCCTAAATATACAT 30240

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db      30239 GCACCCAATACA---------GGAGCACCCAGACTC------ATAAAGCAAGTCCTTAGT 30195

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
```

```
Db      30194 GACCTACAAAGAGACTTACACTCCCACACAATAATG------GGAGACTTTAACACCCCA 30141

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db      30140 CTGTCAACATTAGACAGATCAACAAGACAGAAAGTTAACAAGGATATCCAGGAATTGAAC 30081

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db      30080 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 30024

Qy        293 CysGlnTyrThrTrp 297

Db      30023 ACAGAATATACATTC 30009

RESULT 66
US-11-121-086-97
; Sequence 97, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 97
;   LENGTH: 170837
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-97

Alignment Scores:
Pred. No.:              2.76e+03        Length:         170837
Score:                  95.50           Matches:        52
Percent Similarity:     41.8%           Conservative:   42
Best Local Similarity:  23.1%           Mismatches:     96
Query Match:            5.0%            Indels:         36
DB:                     12              Gaps:           8

US-10-757-745-2 (1-362) x US-11-121-086-97 (1-170837)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110

Db     112896 GAAGAAACTGCATCAACTAACGAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 112954

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db     112955 AATTCACACATAACAATAATAACCTTAAATGTAAATGGGCTA---------AATGCTCCA 113005

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db     113006 ATTAAAAGGCACAGACTGGCAAACTGGATGAAGAGTCAAGACCCATCAGTGTGCTGTATT 113065

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db     113066 CAGGAA-----------------ACACATCTCACGTGCAGAGACACATATAGGCTCAAA 113107

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db     113108 ATAAAGGGATGGAGGAAGATCTACTAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 113167

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db     113168 ATCCTAGTCTCGGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 113227

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226

Db     113228 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAACTAACTATCCTAAATATACAT 113287

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db     113288 GCACCCAATACA---------GGAGCACCCAGACTC------ATAAAGCAAGTCCTTAGT 113332

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db     113333 GACCTACAAAGAGACTTACACTCCCACACAATAATG------GGAGACTTTAACACCCCA 113386

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db     113387 CTGTCAACATTAGACAGATCAACAAGACAGAAAGTTAACAAGGATATCCAGGAATTGAAC 113446

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db     113447 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 113503

Qy        293 CysGlnTyrThrTrp 297

Db     113504 ACAGAATATACATTC 113518

RESULT 67
US-11-112-908-18
; Sequence 18, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
;  APPLICANT: Harris, Cole
;  APPLICANT:  Davis, Lisa M.
;  TITLE OF INVENTION: Breast Cancer Biomarkers
;  FILE REFERENCE: 04-164-US
;  CURRENT APPLICATION NUMBER: US/11/112,908
;  CURRENT FILING DATE:  2005-04-22
;  PRIOR APPLICATION NUMBER: US 60/564,758
;  PRIOR FILING DATE: 2004-04-23
;  PRIOR APPLICATION NUMBER: US 60/575,978
;  PRIOR FILING DATE: 2004-06-01
;  PRIOR APPLICATION NUMBER: US 60/631,702
;  PRIOR FILING DATE: 2004-11-30
;  PRIOR APPLICATION NUMBER: US 60/633,826
;  PRIOR FILING DATE: 2004-12-07
;  NUMBER OF SEQ ID NOS: 511
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18
;   LENGTH: 165883
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-112-908-18

Alignment Scores:
Pred. No.:              3.01e+03        Length:         165883
Score:                  95.00           Matches:        53
Percent Similarity:     42.7%           Conservative:   43
Best Local Similarity:  23.6%           Mismatches:     95
Query Match:            5.0%            Indels:         35
DB:                     12              Gaps:           9

US-10-757-745-2 (1-362) x US-11-112-908-18 (1-165883)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110

Db      93257 GAAGAAACTGCATCAACTAACGAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 93315

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db      93316 AATTCACACATAACAATATTAACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 93369

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db      93370 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 93426

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db      93427 CAGGAA-----------------ACCCATCTCACGTGCAGAGACATACATAGGCTCAAA 93468
```

```
Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db      93469 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 93528

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg------------ 206
Db      93529 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGAAAAAGAAGGC 93588

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 226
Db      93589 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATAT 93648

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db      93649 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 93693

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db      93694 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 93750

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db      93751 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 93810

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db      93811 TCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 93867

Qy        293 CysGlnTyrThrTrp 297
Db      93868 ACAGAATATACATTT 93882

RESULT 68
US-10-995-561-13293/c
; Sequence 13293, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13293
;   LENGTH: 645179
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-13293

Alignment Scores:
Pred. No.:              1.72e+04        Length:         645179
Score:                  95.00           Matches:        64
Percent Similarity:     41.5%           Conservative:   48
Best Local Similarity:  23.7%           Mismatches:     100
Query Match:            5.0%            Indels:         61
DB:                     8               Gaps:           11

US-10-757-745-2 (1-362) x US-10-995-561-13293 (1-645179)

Qy         54 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAla--------- 70
Db      32613 CTGAAGGAAGCACTAAACAYGGAAAG-GAACAACCAGTACAAGCCACTGCAAAATCATGC 32555

Qy         71 ---------------LeuGluArgArgProGluThrIleSerGluProLysThrTyrVal 85
Db      32554 CAAAATGTAAAGACCATCGAGACTAGGAA-GAAACTGCATCA----------------- 32514

Qy         86 AspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 105

Db      32513 ------ACTAACGAGCAAAATAACCAGCTAACATCATCA---------------------- 32481

Qy        106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
Db      32480 ---TGACAGGATCA-AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA--- 32428

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
Db      32427 ------AATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCA 32374

Qy        146 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 165
Db      32373 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 32332

Qy        166 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 185
Db      32331 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 32272

Qy        186 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 205
Db      32271 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGATTTTAAACCAACAAAGATCAAA 32212

Qy        206 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 221
Db      32211 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACT 32152

Qy        222 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 241
Db      32151 AWCCTAAATATATACGCACCCAACACA---------GGAGCACCCAGATTC------ATA 32107

Qy        242 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 261
Db      32106 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 32050

Qy        262 AspThrAsnLeu------------------------------------------ArgAsp 267
Db      32049 GACTTTAATACCCCACTGTCAACACTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 31990

Qy        268 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
Db      31989 ACCCAGGAATTGAACTCAGCTCKGCACCAAGCGGACCTGATAGACATCTACAGAACTCTC 31930

Qy        288 GlyLysProLysHisCysGlnTyrThrTrp 297
Db      31929 ---CACCCCAAATCAACAGAATAYACATTT 31903

RESULT 69
US-10-932-182A-3111
; Sequence 3111, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3111
;   LENGTH: 4251
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-3111

Alignment Scores:
Pred. No.:              40              Length:         4251
Score:                  93.50           Matches:        87
Percent Similarity:     33.3%           Conservative:   56
Best Local Similarity:  20.3%           Mismatches:     157
```

```
Query Match:          4.9%          Indels:        129
DB:                   7             Gaps:          18

US-10-757-745-2 (1-362) x US-10-932-182A-3111 (1-4251)

Qy          7 LeuGluGlyGlyArgGluAlaAla-----------GluGluGluGlyGluProGluVal 22
              ||||||   |||:::|||   |||           |||:::|||   :::      :::
Db       1111 TTGGAATCAGGAAAGGAAACTGCATTGGAATTTTTGGAAAAGGAGAAGCAATTAACGCTT 1170

Qy         23 LysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaVal 42
                |||   :::|||          :::|||      :::   |||               |||
Db       1171 TTGAAATCAAAATTGTTT------CAATTTAAATTATTGCAAAGTAACTCTAAATTAGTT 1224

Qy         43 AlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPhe 62
                           |||      |||         :::||||||      |||
Db       1225 --------------------AATACGTTAGAGAAAGCTTCCTCTTTGAATAAAGATTTT 1263

Qy         63 Glu-----------------------------ProProValGluGluSerAlaLeuGlu 72
              |||                                |||   |||          :::
Db       1264 GAATCCGAAAAGATAAAATTTCAGGAATCCTTGAAAAAGGTTAATGAAGTTGAAAGTCAA 1323

Qy         73 ArgArg-----------------------------------ProGluThrIleSerGlu 80
              |||:::                                   :::|||::::::
Db       1324 CGTAAAGAAATCAAAGATCGAATATCATCTTGCGCTTCTCAAGAAAAGACGCTGGCTCTG 1383

Qy         81 ProLysThrTyrValAspLeuThrAsn---------GluGluThrThrAspSerThrThr 97
                 |||      ::::::   |||            ||||||   |||   :::
Db       1384 GAAAAGAGAGAACTAGAAAGCACCAGAGTTTCTTTAGAAGAAAGAACAAAGAATCTGGTT 1443

Qy         98 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 117
              :::|||:::      :::|||         :::                  |||
Db       1444 AACAAAATGGGGAAAGCAGAAAAGATTTTGAAGTCGACCAACCAATCAATATCTGAAGCT 1503

Qy        118 ThrTrpAsnIleAspGlyLeu-----------------------------------Asp 125
                    ::::::   |||                                      |||
Db       1504 GAACATCTTCTCGAAGAACTTCATGGGGAGCAGAATGAACACGAAACAGAGGTTAAAGAT 1563

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
              ||||||         |||      |||   :::         :::   |||
Db       1564 TTGAACCAATCGCTGGAAGAGGAACGTCGAATACTCGATAATATCAAATTATCTTTAAAG 1623

Qy        146 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 165
              |||                                                  |||   :::
Db       1624 GAT-----------------------------------------------AAAACTAAGGAC 1638

Qy        166 SerAsnTyrGluIleIleThrGlyHisGlu---------GluGlyTyrPheThrAlaIle 182
                 :::   |||||||||       ||||||          |||   :::         :::
Db       1639 ATTTCTGCAGAAATTATTCAA---CATGAAAAGGATTTAGAACCTTGGGATCTCCAACTT 1695

Qy        183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 202
                 |||||||||:::::::::|||       |||:::                  |||
Db       1696 CAAGAGAAGAAATCGCAAATACAATTAGCAGAATCAGAACTATCTCTATTAGAAGAAACT 1755

Qy        203 ------LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu--- 219
                    |||:::   :::|||:::       :::      |||:::|||       :::
Db       1756 CAAGTCAAGCTGAAAAAAAATGTTGAATCTTTGGAAGAGAATATTTCCGCAAAGAAAGCA 1815

Qy        220 ----------------------LeuCysLeuMetThrSerHisLeuGluSerThr--- 230
                                    :::      :::      :::|||:::|||
Db       1816 CGCAAGCAAGGGCTACAAGGGTTTATTCTCGATCTAAAAAAGAACCTAAAGTCACTCATA 1875

Qy        231 ------ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
                    |||          :::      |||         :::   ||||||:::|||
Db       1876 GATGGAAGATCGCAAGGTGAAAAAGATTTTAATACTGCCAACTTAAAGTTAAAAGAGATG 1935

Qy        249 Gln-----------------------------------------------------Glu 250
              |||                                                      :::
Db       1936 CAAACGATTTTGAACGCCCATCGTCAACGTGCCATGGAAGCTCGCTCTTCCCTATCAAAA 1995

Qy        251 AlaProGluSerAlaThrValIlePheAla---------------------------- 260
              |||         :::   |||:::   |||
Db       1996 GCTCAGAATAAAAGTAAAGTTTTAACAGCTCTATCAAGGTTACAAAAGTCCGGACGTATC 2055

Qy        261 ----------------------GlyAspThrAsnLeuArgAspArgGluVal--------- 270
                                     ||||||      :::   |||
Db       2056 AGTGGGTTCCATGGACGCCTGGGGGATCTCGGTGTTATAGACGACAGCTTTGATGTAGCT 2115

Qy        271 ------ThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
                   |||   |||       |||   :::   :::||||||      |||
Db       2116 ATTTCTACAGCCTGTCCAAGGTTGGATGATGTTGTTGTTGATACTGTGGAATGT------ 2169

Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
                   :::||||||          |||             :::   ||||||
Db       2170 ---GCTCAGCACTGC---------ATAGACTATTTGAGAAAGAATAAACTAGGTTATGCA 2217

Qy        309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328
                        |||      |||   |||             :::
Db       2218 AGGTTTATTTTATTGGACAGACTACGCCAATTTAATTTGCAACCAGTTGATACACCAGAA 2277

Qy        329 IleIleProArgSerLeuAspLeuLeu 337
                 :::||||||      ||||||:::
Db       2278 AATGTACCAAGGCTGTTTGACCTAGTT 2304

RESULT 70
US-10-932-182A-3111
; Sequence 3111, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3111
;   LENGTH: 4251
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-3111

Alignment Scores:
Pred. No.:              40              Length:         4251
Score:                  93.50           Matches:        87
Percent Similarity:     33.3%           Conservative:   56
Best Local Similarity:  20.3%           Mismatches:     157
Query Match:            4.9%            Indels:         129
DB:                     7               Gaps:           18

US-10-757-745-2 (1-362) x US-10-932-182A-3111 (1-4251)

Qy          7 LeuGluGlyGlyArgGluAlaAla-----------GluGluGluGlyGluProGluVal 22
              ||||||   |||:::|||   |||           |||:::|||   :::      :::
Db       1111 TTGGAATCAGGAAAGGAAACTGCATTGGAATTTTTGGAAAAGGAGAAGCAATTAACGCTT 1170

Qy         23 LysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAlaVal 42
                |||   :::|||          :::|||      :::   |||               |||
Db       1171 TTGAAATCAAAATTGTTT------CAATTTAAATTATTGCAAAGTAACTCTAAATTAGTT 1224

Qy         43 AlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyrPhe 62
                           |||      |||         :::||||||      |||
Db       1225 --------------------AATACGTTAGAGAAAGCTTCCTCTTTGAATAAAGATTTT 1263

Qy         63 Glu-----------------------------ProProValGluGluSerAlaLeuGlu 72
              |||                                |||   |||          :::
Db       1264 GAATCCGAAAAGATAAAATTTCAGGAATCCTTGAAAAAGGTTAATGAAGTTGAAAGTCAA 1323

Qy         73 ArgArg-----------------------------------ProGluThrIleSerGlu 80
              |||:::                                   :::|||::::::
Db       1324 CGTAAAGAAATCAAAGATCGAATATCATCTTGCGCTTCTCAAGAAAAGACGCTGGCTCTG 1383

Qy         81 ProLysThrTyrValAspLeuThrAsn---------GluGluThrThrAspSerThrThr 97
```

```
Db       1384 GAAAAGAGAGAACTAGAAAGCACCAGAGTTTCTTTAGAAGAAAGAACAAAGAATCTGGTT 1443

Qy         98 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 117

Db       1444 AACAAAATGGGGAAAGCAGAAAAGATTTGAAGTCGACCAACCAATCAATATCTGAAGCT 1503

Qy        118 ThrTrpAsnIleAspGlyLeu-----------------------------------Asp 125

Db       1504 GAACATCTTCTCGAAGAACTTCATGGGGAGCAGAATGAACACGAAACAGAGGTTAAAGAT 1563

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145

Db       1564 TTGAACCAATCGCTGGAAGAGGAACGTCGAATACTCGATAATATCAAATTATCTTTAAAG 1623

Qy        146 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 165

Db       1624 GAT-----------------------------------------------AAAACTAAGGAC 1638

Qy        166 SerAsnTyrGluIleIleThrGlyHisGlu---------GluGlyTyrPheThrAlaIle 182

Db       1639 ATTTCTGCAGAAATTATTCAA---CATGAAAAGGATTTAGAACCTTGGGATCTCCAACTT 1695

Qy        183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 202

Db       1696 CAAGAGAAGAAATCGCAAATACAATTAGCAGAATCAGAACTATCTCTATTAGAAGAAACT 1755

Qy        203 ------LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu--- 219

Db       1756 CAAGTCAAGCTGAAAAAAAAATGTTGAATCTTTGGAAGAGAATATTTCCGCAAAGAAAGCA 1815

Qy        220 ------------------------LeuCysLeuMetThrSerHisLeuGluSerThr--- 230

Db       1816 CGCAAGCAAGGGCTACAAGGGTTTATTCTCGATCTAAAAAAGAACCTAAAGTCACTCATA 1875

Qy        231 ------ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248

Db       1876 GATGGAAGATCGCAAGGTGAAAAAGATTTTAATACTGCCAACTTAAAGTTAAAAGAGATG 1935

Qy        249 Gln-----------------------------------------------------Glu 250

Db       1936 CAAACGATTTTGAACGCCCATCGTCAACGTGCCATGGAAGCTCGCTCTTCCCTATCAAAA 1995

Qy        251 AlaProGluSerAlaThrValIlePheAla---------------------------- 260

Db       1996 GCTCAGAATAAAAGTAAAGTTTTAACAGCTCTATCAAGGTTACAAAAGTCCGGACGTATC 2055

Qy        261 --------------------GlyAspThrAsnLeuArgAspArgGluVal--------- 270

Db       2056 AGTGGGTTCCATGGACGCCTGGGGGATCTCGGTGTTATAGACGACAGCTTTGATGTAGCT 2115

Qy        271 ------ThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288

Db       2116 ATTTCTACAGCCTGTCCAAGGTTGGATGATGTTGTTGTTGATACTGTGGAATGT------ 2169

Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308

Db       2170 ---GCTCAGCACTGC---------ATAGACTATTTGAGAAAGAATAAACTAGGTTATGCA 2217

Qy        309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328

Db       2218 AGGTTTATTTTATTGGACAGACTACGCCAATTTAATTTGCAACCAGTTGATACACCAGAA 2277

Qy        329 IleIleProArgSerLeuAspLeuLeu 337

Db       2278 AATGTACCAAGGCTGTTTGACCTAGTT 2304
```

RESULT 71
US-11-121-086-85
; Sequence 85, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 85
; LENGTH: 171423
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-85

Alignment Scores:

Pred. No.:	4.56e+03	Length:	171423
Score:	93.50	Matches:	57
Percent Similarity:	42.4%	Conservative:	46
Best Local Similarity:	23.5%	Mismatches:	96
Query Match:	4.9%	Indels:	44
DB:	12	Gaps:	9

US-10-757-745-2 (1-362) x US-11-121-086-85 (1-171423)

```
Qy        102 ProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle 121

Db     149784 CCAGCTAACATCATAATGACAGGATCAAATTCACACATAACAATATTAACCTTAAATGTA 149843

Qy        122 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 141

Db     149844 AATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 149894

Qy        142 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 161

Db     149895 AGTTAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCGTCTC 149936

Qy        162 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 181

Db     149937 ACGTGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCAACCAAGCAAAT 149996

Qy        182 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 201

Db     149997 GGAAAACAAAAAAAGGCAGGGGTTGCAATCTTAGGCTCTCATAAAACAGACTTTAAACCA 150056

Qy        202 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 218

Db     150057 AGAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 150116

Qy        219 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 237

Db     150117 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 150167

Qy        238 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257

Db     150168 TTCATAAACATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACAATA 150227

Qy        258 IlePheAlaGlyAsp------------------------ThrAsnLeuArgAspArgGl 269

Db     150228 ATAATG---GGAGATTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 150284

Qy        269 uValThrArgCysGly-GlyLeu---------------ProAsn-AsnIleValAspVal 283

Db     150285 GTTAACAAGGATATCCAGGAATTGAACTCAGCTCTGCACCAAGCAAACCTAATAGACGTC 150344

Qy        284 TrpGluPheLeuGlyLysProLys----------------------------HisCys 293

Db     150345 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTCTTCTCAGCACCACACTGC 150401

Qy        294 GlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeu 313

Db     150402 ACTTATTCCAAAATTGACCACATAGTTGGAAGTAAAGCACTCCTCAGCAAATGTAAAAGA 150461
```

```
Qy        314 Arg 314
              |||
Db     150462 AGA 150464

RESULT 72
US-11-121-086-77
; Sequence 77, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77
;   LENGTH: 215308
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-77

Alignment Scores:
Pred. No.:              6.11e+03      Length:          215308
Score:                  93.50         Matches:         53
Percent Similarity:     44.2%         Conservative:    50
Best Local Similarity:  22.7%         Mismatches:      84
Query Match:            4.9%          Indels:          46
DB:                     12            Gaps:            9

US-10-757-745-2 (1-362) x US-11-121-086-77 (1-215308)

Qy        110 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 129
              :::   |||      :::::::::|||   |||::::::|||||||         |||
Db      18305 TCAAATTCACACACAACAATTTTAACCTTAAACGTAAATGGGCTA---------AATGCG 18355

Qy        130 SerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePhe 149
                  :::      :::   :::::::::            |||   |||
Db      18356 CCAATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATTAGTGTGCTGT 18415

Qy        150 LeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGlu 169
              :::||||||                . ::::::|||     :::   ::::::
Db      18416 ATTCAGGAG-----------------ACCCATCTCATGTGCAAAGATACACATAGGCTC 18457

Qy        170 IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 189
                |||  |||              |||                 ||||||::::::|||
Db      18458 AAAATAAAGGGATGGAGGAATATTTACCAAGCAAATGGAAAGCAAAAAAAAGCACAAGTT 18517

Qy        190 LysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu 209
                 :::      :::       |||     ||||||:::   |||:::
Db      18518 GCAATCCTAATCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAA 18577

Qy        210 CysValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSer 225
                ::::::   |||   |||:::           |||:::   ::::::
Db      18578 GGGCATTACATAATGGTAAAGGGATCAATGCAGCAAGAAGAGGTAACTATCCTAAATATA 18637

Qy        226 HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeu 245
              :::      :::|||                |||           :::|||   ||||||
Db      18638 TATGCACCCAATACA--------------GAAGCACCCAGATTCATAAAGCAAGTTCTT 18682

Qy        246 LysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn--- 264
              :::   :::|||        :::|||   |||::::::      ||||||   |||
Db      18683 AGAGACCTACAAAGAGACTTAGACTCCCACACAATAGTAATG---GGAGACTTTAACACC 18739

Qy        265 --------------------Leu-ArgAspArgGluValThrArgCys-GlyGlyLeu- 276
                              ||| |||||||||:::::||||||      ||||||
Db      18740 CTACTGTCAATATTAGGCAGGTTAACGAGACAGAAAATTAACAAGGATATTCAGGACTAG 18799

Qy        277 -------------ProAsn-AsnIleValAspValTrpGluPheLeuGlyLysProLys 291
                          |||::: :::::::::|||::::::     |||      ||||||
```

```
Db      18800 AACTCAACTCTGGACCAAGCGGACCTATTAGACATCTATAGAACTCTC---CACCCCAAA 18856

Qy        292 ----------------------------HisCysGlnTyrThrTrpAspThrGlnMet 301
                                          ||||||   |||:::          :::
Db      18857 TCAACAGAATATACATTCTGCTCAGCACCTCATTGCACTTATTCTAAAACTGATGACATA 18916

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
                  :::     :::   :::   ||||||
Db      18917 ATTGGAAGTAAAACACTCCTCAGCAAATGTAAA 18949

RESULT 73
US-09-925-065A-871294/c
; Sequence 871294, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 871294
;   LENGTH: 597
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-871294

Alignment Scores:
Pred. No.:              3.66          Length:          597
Score:                  93.00         Matches:         54
Percent Similarity:     44.1%         Conservative:    36
Best Local Similarity:  26.5%         Mismatches:      78
Query Match:            4.9%          Indels:          36
DB:                     6             Gaps:            8

US-10-757-745-2 (1-362) x US-09-925-065A-871294 (1-597)

Qy        110 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 129
              :::   |||      |||::::::|||   |||::::::||||||         |||
Db        595 TCAAATTCACACATATCAATACTAACCTTGAATGTAAATGGGCTA---------AATGCC 545

Qy        130 SerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePhe 149
              |||  :::          :::   |||::::::          |||   |||
Db        544 TCACTTAAAAGGCACAGAATGGCCAGTTGGATAAAAAAACAAGACCCAATGGTATGCTGT 485

Qy        150 LeuGlnGluVal-IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGl 169
              |||||||||    |||                 :::      :::      |||   ||
Db        484 CTTCAAGAGACCTATCTCACATGTAATGAAACTCATAGGGTCAAAATAAAGAAATGGMGA 425

Qy        169 uIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVa 189
              |      |||          |||                          ||||||:::   ::
Db        424 AAAATCTACCAAGCAAATGGAAAAC-----------------AGGAAAAAGGCAGGGAT 383

Qy        189 lLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg-------- 206
              :  :::      :::          |||      ||||||:::   :::
Db        382 TGCAATCCTAATTTCAGACAAAACAGATTTCAAACCAACAAAGATCAAAAAAGACAAGGA 323

Qy        207 ----AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSe 225
                  :::   :::   |||      ::::::                    |||:::
Db        322 AGAGCATTACATAATGGTTAAGGATTCAATTCAACAAGAAAACCGAACTCTCCTAAATAT 263
```

```
Qy        225 rHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLe 245
Db        262 ATATGCACCCAACACAAGA---------GCATCCAGATTC------ATAAAGCAAGTCCT 218

Qy        245 uLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn-L 265
Db        217 TAGAGCCCTACAAAGAGGCATAGACTCCCACACAATAATA---GCGGGAGACTTCAACAC 161

Qy        265 eu------------------------ArgAspArgGluValThr-ArgCysGlyGlyLeu 276
Db        160 TCCACTGACAGTATTAGACAGATCCTCGAAGCAGAAAATTAACAAAGATATTCAGGACTT 101

Qy        277 ----------------ProAsn-AsnIleValAspValTrpGluPheLeuGlyLysProLy 291
Db        100 AAACTCAGCATTGGACCAAATGGATCTGATAGACCTTTACAGAAATCTCCAACCTCCCCC 41

Qy        291 sHis 292
Db         40 GCAC 37
```

RESULT 74
US-11-121-086-70
; Sequence 70, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 70
; LENGTH: 180574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-70

Alignment Scores:

Pred. No.:	5.52e+03	Length:	180574
Score:	93.00	Matches:	51
Percent Similarity:	41.8%	Conservative:	43
Best Local Similarity:	22.7%	Mismatches:	97
Query Match:	4.9%	Indels:	35
DB:	12	Gaps:	8

US-10-757-745-2 (1-362) x US-11-121-086-70 (1-180574)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db     149607 GAAGAAACTGCATCAACTAACGAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 149665

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db     149666 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTCCA 149716

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db     149717 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 149776

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db     149777 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 149818

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db     149819 ATAAAAGGATGGAGGAAGGTCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 149878

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db     149879 TTCCTAGTCTCTGATAAAACAGACTTTAAACTAACAAAGATCAAAAGAGACAAAGAAGGC 149938

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
Db     149939 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATAC 149998

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db     149999 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 150043

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn------ 264
Db     150044 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACAACCCA 150100

Qy        265 ----------------------------------LeuArgAspArgGluValThrArg 272
Db     150101 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTTAACAAGGATACCCAGTAATTGAAC 150160

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db     150161 TCAGCTCTGCACCAAGGGGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 150217

Qy        293 CysGlnTyrThrTrp 297
Db     150218 ACAGAATATACATTT 150232
```

RESULT 75
US-11-112-908-19
; Sequence 19, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19
; LENGTH: 212805
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-19

Alignment Scores:

Pred. No.:	6.81e+03	Length:	212805
Score:	93.00	Matches:	64
Percent Similarity:	41.1%	Conservative:	47
Best Local Similarity:	23.7%	Mismatches:	101
Query Match:	4.9%	Indels:	61
DB:	12	Gaps:	11

US-10-757-745-2 (1-362) x US-11-112-908-19 (1-212805)

```
Qy         54 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAla--------- 70
Db      40611 CTGAAGGAAGCGCTAAACATGGAAAG-GAACAACCAGTACCAGCCGCTGCAAAATCATGC 40669

Qy         71 ---------------LeuGluArgArgProGluThrIleSerGluProLysThrTyrVal 85
Db      40670 CAAAATGTAAAGACCATCGAGACTAGGAA-GAAACTGCATCA----------------- 40710

Qy         86 AspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 105
```

```
Db      40711 ------ACTAACGAGCAAAATAACCAGCTAACATCATCA---------------------- 40743

Qy        106 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 125
Db      40744 ---TGACAGGATCA-AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA--- 40796

Qy        126 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
Db      40797 ------AATGCTCCAATTAAAAGACACAGACTGGCAAGTTGGATAAAGAGTCAAGACCCA 40850

Qy        146 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 165
Db      40851 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 40892

Qy        166 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 185
Db      40893 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 40952

Qy        186 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 205
Db      40953 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACGTTAAACCAACAAAGATCAAA 41012

Qy        206 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 221
Db      41013 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACT 41072

Qy        222 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 241
Db      41073 ATACTAAATATATATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 41117

Qy        242 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 261
Db      41118 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 41174

Qy        262 AspThrAsnLeu-------------------------------------------ArgAsp 267
Db      41175 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 41234

Qy        268 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
Db      41235 ACCCAGGAATTTAACTCATCTCCGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC 41294

Qy        288 GlyLysProLysHisCysGlnTyrThrTrp 297
Db      41295 ---CATCCCAAATCAACAGAATATACGTTT 41321
```

Search completed: March 11, 2006, 15:36:25
Job time : 7355.64 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 15:40:34 ; Search time 241.043 Seconds
(without alignments)
1746.920 Million cell updates/sec

Title: US-10-757-745-2_COPY_145_153
Perfect score: 45
Sequence: 1 PDVIFLQEV 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170329_4058/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_12397_@runat_09032006_170329_4058 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	45	100.0	149	3	BQ336306	BQ336306 PM1-MT014
c	2	45	100.0	156	3	BQ331200	BQ331200 MR4-ET014
	3	45	100.0	247	2	BF768814	BF768814 PM0-IT001
	4	45	100.0	252	1	AA325145	AA325145 EST28102
	5	45	100.0	296	2	BF992136	BF992136 QV3-GN020
	6	45	100.0	313	2	BF992133	BF992133 QV3-GN020
c	7	45	100.0	368	2	BE694706	BE694706 PM0-BT075
	8	45	100.0	388	1	AV659558	AV659558 AV659558
	9	45	100.0	394	1	AV661333	AV661333 AV661333
	10	45	100.0	413	3	BQ332322	BQ332322 MR4-ET014
	11	45	100.0	419	2	BF853399	BF853399 MR3-EN008
c	12	45	100.0	422	2	BG982044	BG982044 MR3-CN014
	13	45	100.0	429	5	BQ582059	BQ582059 il12c11.y
c	14	45	100.0	435	3	BQ332331	BQ332331 MR4-ET014
c	15	45	100.0	437	2	BF768820	BF768820 PM0-IT001
c	16	45	100.0	440	2	BG982029	BG982029 MR3-CN014
c	17	45	100.0	440	2	BG982036	BG982036 MR3-CN014
	18	45	100.0	441	2	BF892415	BF892415 PM1-MT014
	19	45	100.0	441	2	BF893925	BF893925 PM1-MT014
	20	45	100.0	444	2	BF893187	BF893187 PM1-MT014
	21	45	100.0	447	2	BF891509	BF891509 PM1-MT014
c	22	45	100.0	453	2	BF773259	BF773259 PM0-IT001
c	23	45	100.0	453	3	BQ315535	BQ315535 PM0-IT001
c	24	45	100.0	454	2	BG982031	BG982031 MR3-CN014
c	25	45	100.0	456	2	BF996980	BF996980 QV3-GN020
	26	45	100.0	456	3	BQ332321	BQ332321 MR4-ET014
c	27	45	100.0	458	3	BQ366782	BQ366782 QV3-GN020
c	28	45	100.0	459	2	BG982034	BG982034 MR3-CN014
	29	45	100.0	460	2	BF893150	BF893150 PM1-MT014
c	30	45	100.0	461	2	BG982012	BG982012 MR3-CN014
	31	45	100.0	463	2	BF852284	BF852284 MR3-EN008
c	32	45	100.0	463	2	BF852788	BF852788 MR3-EN008
	33	45	100.0	463	2	BF893919	BF893919 PM1-MT014
c	34	45	100.0	471	2	BF893924	BF893924 PM1-MT014
	35	45	100.0	471	7	CR543841	CR543841 DKFZp459O
	36	45	100.0	473	6	CB120234	CB120234 K-EST0167
	37	45	100.0	480	1	AI750554	AI750554 cn04a01.y
	38	45	100.0	485	1	AA486032	AA486032 ab40b10.r
	39	45	100.0	491	1	AJ681912	AJ681912 AJ681912
c	40	45	100.0	491	2	BI016235	BI016235 MR4-ET014
	41	45	100.0	530	3	BM846221	BM846221 K-EST0125
	42	45	100.0	534	7	CN298921	CN298921 170005315
	43	45	100.0	546	2	BG541031	BG541031 602570190
	44	45	100.0	549	1	AW669835	AW669835 113421 MA
	45	45	100.0	554	3	BP220985	BP220985 BP220985
	46	45	100.0	565	3	BP226620	BP226620 BP226620
	47	45	100.0	566	3	BP220508	BP220508 BP220508
	48	45	100.0	568	3	BP257211	BP257211 BP257211
	49	45	100.0	570	3	BM172060	BM172060 imageqc_3
	50	45	100.0	570	3	BP221518	BP221518 BP221518
	51	45	100.0	570	3	BP258309	BP258309 BP258309
	52	45	100.0	572	3	BP262485	BP262485 BP262485
	53	45	100.0	573	1	AU279894	AU279894 AU279894
	54	45	100.0	573	3	BP221260	BP221260 BP221260
	55	45	100.0	573	3	BP261500	BP261500 BP261500
	56	45	100.0	576	3	BP262642	BP262642 BP262642
	57	45	100.0	577	7	CN298919	CN298919 170006001
	58	45	100.0	580	3	BP261787	BP261787 BP261787
	59	45	100.0	580	3	BP274368	BP274368 BP274368
	60	45	100.0	581	1	AV717253	AV717253 AV717253
	61	45	100.0	581	3	BP207751	BP207751 BP207751
	62	45	100.0	581	3	BP225088	BP225088 BP225088
	63	45	100.0	581	3	BP262101	BP262101 BP262101
	64	45	100.0	581	3	BP262638	BP262638 BP262638
	65	45	100.0	581	3	BP270744	BP270744 BP270744
	66	45	100.0	581	3	BP349936	BP349936 BP349936
	67	45	100.0	581	3	BP379659	BP379659 BP379659
	68	45	100.0	581	6	CB286391	CB286391 CMD36_C09
	69	45	100.0	582	3	BM507091	BM507091 ih24h01.y
	70	45	100.0	582	3	BP219740	BP219740 BP219740
	71	45	100.0	582	3	BP233592	BP233592 BP233592
	72	45	100.0	582	3	BP234218	BP234218 BP234218
	73	45	100.0	582	3	BP261141	BP261141 BP261141
	74	45	100.0	582	3	BP261789	BP261789 BP261789
	75	45	100.0	582	3	BP263516	BP263516 BP263516

ALIGNMENTS

RESULT 1

```
BQ336306
LOCUS       BQ336306                149 bp    mRNA    linear   EST 20-MAY-2002
DEFINITION  PM1-MT0143-100701-007-e12 MT0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ336306
VERSION     BQ336306.1  GI:20992407
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 149)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-MT0143-
            100701-007-e12&t3=2001-07-10&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 22
            High quality sequence stop: 148.
FEATURES             Location/Qualifiers
     source          1. .149
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="MT0143"
                     /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                 5.86          Length:          149
Score:                     45.00         Matches:         9
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               100.0%        Indels:          0
DB:                        3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BQ336306 (1-149)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         91 CCAGATGTGATATTTCTACAGGAAGTT 117

RESULT 2
BQ331200/c
LOCUS       BQ331200                156 bp    mRNA    linear   EST 17-MAY-2002
DEFINITION  MR4-ET0140-220101-004-h09_1 ET0140 Homo sapiens cDNA, mRNA
            sequence.
ACCESSION   BQ331200
VERSION     BQ331200.1  GI:20972241
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 156)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-ET0140-
            220101-004-h09_1&t3=2001-01-22&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .156
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="ET0140"
                     /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                 6.17          Length:          156
Score:                     45.00         Matches:         9
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               100.0%        Indels:          0
DB:                        3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BQ331200 (1-156)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         76 CCAGATGTGATATTTCTACAGGAAGTT 50

RESULT 3
BF768814
LOCUS       BF768814                247 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  PM0-IT0018-091100-001-d12 IT0018 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF768814
VERSION     BF768814.1  GI:12116818
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 247)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM0&t2=PM0-IT0018-
            091100-001-d12&t3=2000-11-09&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 9
            High quality sequence stop: 246.
FEATURES             Location/Qualifiers
     source          1. .247
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="IT0018"
                     /note="Organ: epid_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                10.4          Length:          247
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       2             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BF768814 (1-247)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        110 CCAGATGTGATATTTCTACAGGAAGTT 136

RESULT 4
AA325145
LOCUS       AA325145                252 bp    mRNA    linear   EST 20-APR-1997
DEFINITION  EST28102 Cerebellum II Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION   AA325145
VERSION     AA325145.1  GI:1977400
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 252)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  PUBMED    7566098
COMMENT     Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .252
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="ATCC (inhost):125652"
                     /db_xref="taxon:9606"
                     /tissue_type="cerebellum"
                     /dev_stage="adult"
                     /clone_lib="Cerebellum II"
                     /note="Organ: brain; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
ORIGIN

Alignment Scores:
Pred. No.:                10.7          Length:          252
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       1             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AA325145 (1-252)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          2 CCAGATGTGATATTTCTACAGGAAGTT 28

RESULT 5
BF992136
LOCUS       BF992136                296 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  QV3-GN0201-301000-454-b08 GN0201 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF992136
VERSION     BF992136.1  GI:12398459
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 296)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
```

```
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-GN0201-
            301000-454-b08&t3=2000-10-30&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 296.
FEATURES             Location/Qualifiers
     source          1. .296
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0201"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                12.8          Length:         296
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BF992136 (1-296)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         99 CCAGATGTGATATTTCTACAGGAAGTT 125

RESULT 6
BF992133
LOCUS       BF992133                 313 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  QV3-GN0201-301000-454-g09 GN0201 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF992133
VERSION     BF992133.1  GI:12398456
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 313)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-GN0201-
            301000-454-g09&t3=2000-10-30&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 45
            High quality sequence stop: 311.
FEATURES             Location/Qualifiers
     source          1. .313
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0201"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                13.6          Length:         313
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BF992133 (1-313)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         99 CCAGATGTGATATTTCTACAGGAAGTT 125

RESULT 7
BE694706/c
LOCUS       BE694706                 368 bp    mRNA    linear   EST 11-SEP-2000
DEFINITION  PM0-BT0757-140800-006-d10 BT0757 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE694706
VERSION     BE694706.1  GI:10081866
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 368)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
```

```
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=PM0-BT0757-140
            800-006-d10&t3=2000-08-14&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 37
            High quality sequence stop: 368.
FEATURES              Location/Qualifiers
     source           1. .368
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="BT0757"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No.
                      196,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                16.4          Length:         368
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BE694706 (1-368)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        176 CCAGATGTGATATTTCTACAGGAAGTT 150

RESULT 8
AV659558
LOCUS       AV659558                388 bp    mRNA    linear   EST 16-JAN-2002
DEFINITION  AV659558 GLC Homo sapiens cDNA clone GLCFYD04 3', mRNA sequence.
ACCESSION   AV659558
VERSION     AV659558.1  GI:9880572
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 388)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
   PUBMED   11752456
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES              Location/Qualifiers
     source           1. .388
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="GLCFYD04"
                      /tissue_type="corresponding non cancerous liver tissue"
                      /dev_stage="Adult"
                      /lab_host="SOLR"
                      /clone_lib="GLC"
                      /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                      XhoI"
ORIGIN

Alignment Scores:
Pred. No.:                17.4          Length:         388
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       1             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AV659558 (1-388)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        315 CCAGATGTGATATTTCTACAGGAAGTT 341

RESULT 9
AV661333
LOCUS       AV661333                394 bp    mRNA    linear   EST 16-JAN-2002
DEFINITION  AV661333 GLC Homo sapiens cDNA clone GLCGRD01 3', mRNA sequence.
ACCESSION   AV661333
VERSION     AV661333.1  GI:9882347
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 394)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
   PUBMED   11752456
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES              Location/Qualifiers
     source           1. .394
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="GLCGRD01"
                      /tissue_type="corresponding non cancerous liver tissue"
                      /dev_stage="Adult"
                      /lab_host="SOLR"
                      /clone_lib="GLC"
                      /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                      XhoI"
ORIGIN

Alignment Scores:
Pred. No.:                17.7          Length:         394
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
```

```
Query Match:            100.0%      Indels:        0
DB:                     1           Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x AV661333 (1-394)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db        221 CCAGATGTGATATTTCTACAGGAAGTT 247

RESULT 10
BQ332322
LOCUS       BQ332322                  413 bp    mRNA    linear   EST 17-MAY-2002
DEFINITION  MR4-ET0140-240501-015-d11 ET0140 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ332322
VERSION     BQ332322.1  GI:20973745
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-ET0140-
            240501-015-d11&t3=2001-05-24&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 6
            High quality sequence stop: 413.
FEATURES             Location/Qualifiers
     source          1..413
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="ET0140"
                     /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:              18.7          Length:          413
Score:                  45.00         Matches:         9
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BQ332322 (1-413)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db         53 CCAGATGTGATATTTCTACAGGAAGTT 79

RESULT 11
BF853399
LOCUS       BF853399                  419 bp    mRNA    linear   EST 16-JAN-2001
DEFINITION  MR3-EN0087-191200-015-a07 EN0087 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF853399
VERSION     BF853399.1  GI:12241143
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 419)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-EN0087-
            191200-015-a07&t3=2000-12-19&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 419.
FEATURES             Location/Qualifiers
     source          1. .419
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="EN0087"
                     /note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:              19            Length:          419
Score:                  45.00         Matches:         9
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     2             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BF853399 (1-419)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db         59 CCAGATGTGATATTTCTACAGGAAGTT 85

RESULT 12
```

```
BG982044/c
LOCUS       BG982044                422 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION  MR3-CN0147-120201-011-h12 CN0147 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG982044
VERSION     BG982044.1  GI:14384779
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-CN0147-
            120201-011-h12&t3=2001-02-12&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 47
            High quality sequence stop: 422.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CN0147"
                     /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                19.2          Length:         422
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BG982044 (1-422)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        342 CCAGATGTGATATTTCTACAGGAAGTT 316

RESULT 13
BQ582059
LOCUS       BQ582059                429 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  il12c11.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6029589
            5' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION   BQ582059
VERSION     BQ582059.1  GI:21494955
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: il12c11.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6029589"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:                19.5          Length:         429
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       5             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BQ582059 (1-429)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        376 CCAGATGTGATATTTCTACAGGAAGTT 402

RESULT 14
BQ332331/c
LOCUS       BQ332331                435 bp    mRNA    linear   EST 17-MAY-2002
DEFINITION  MR4-ET0140-240501-015-g10 ET0140 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ332331
VERSION     BQ332331.1  GI:20973758
KEYWORDS    EST.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 435)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-ET0140-
            240501-015-g10&t3=2001-05-24&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 132
            High quality sequence stop: 435.
FEATURES             Location/Qualifiers
     source          1. .435
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="ET0140"
                     /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                19.8          Length:          435
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BQ332331 (1-435)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        355 CCAGATGTGATATTTCTACAGGAAGTT 329

RESULT 15
BF768820/c
LOCUS       BF768820                437 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  PM0-IT0018-091100-001-a01 IT0018 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF768820
VERSION     BF768820.1  GI:12116824
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 437)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM0&t2=PM0-IT0018-
            091100-001-a01&t3=2000-11-09&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 28
            High quality sequence stop: 437.
FEATURES             Location/Qualifiers
     source          1. .437
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="IT0018"
                     /note="Organ: epid_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                19.9          Length:          437
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       2             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BF768820 (1-437)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        357 CCAGATGTGATATTTCTACAGGAAGTT 331

RESULT 16
BG982029/c
LOCUS       BG982029                440 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION  MR3-CN0147-120201-011-b10 CN0147 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG982029
VERSION     BG982029.1  GI:14384764
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
```

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-CN0147-120201-011-b10&t3=2001-02-12&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 440.
FEATURES Location/Qualifiers
source 1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CN0147"
/note="Organ: colon_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
ORIGIN

Alignment Scores:

Pred. No.:	20.1	Length:	440
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BG982029 (1-440)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        360 CCAGATGTGATATTTCTACAGGAAGTT 334
```

RESULT 17
BG982036/c
LOCUS BG982036 440 bp mRNA linear EST 12-JUN-2001
DEFINITION MR3-CN0147-120201-011-d12 CN0147 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG982036
VERSION BG982036.1 GI:14384771
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 440)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-CN0147-120201-011-d12&t3=2001-02-12&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 440.
FEATURES Location/Qualifiers
source 1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CN0147"
/note="Organ: colon_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
ORIGIN

Alignment Scores:

Pred. No.:	20.1	Length:	440
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BG982036 (1-440)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        360 CCAGATGTGATATTTCTACAGGAAGTT 334
```

RESULT 18
BF892415
LOCUS BF892415 441 bp mRNA linear EST 18-JAN-2001
DEFINITION PM1-MT0143-041100-002-f12 MT0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF892415
VERSION BF892415.1 GI:12283874
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 441)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922

Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-MT0143-041100-002-f12&t3=2000-11-04&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 441.

FEATURES Location/Qualifiers
source 1. .441
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0143"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	20.1	Length:	441
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BF892415 (1-441)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         81 CCAGATGTGATATTTCTACAGGAAGTT 107
```

RESULT 19
BF893925
LOCUS BF893925 441 bp mRNA linear EST 18-JAN-2001
DEFINITION PM1-MT0143-131100-004-g08 MT0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF893925
VERSION BF893925.1 GI:12285384
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 441)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-MT0143-131100-004-g08&t3=2000-11-13&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 441.

FEATURES Location/Qualifiers
source 1. .441
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0143"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	20.1	Length:	441
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BF893925 (1-441)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         81 CCAGATGTGATATTTCTACAGGAAGTT 107
```

RESULT 20
BF893187
LOCUS BF893187 444 bp mRNA linear EST 18-JAN-2001
DEFINITION PM1-MT0143-101100-003-h07 MT0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF893187
VERSION BF893187.1 GI:12284646
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 444)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-MT0143-101100-003-h07&t3=2000-11-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 444.

FEATURES Location/Qualifiers
source 1. .444
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0143"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	20.3	Length:	444
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BF893187 (1-444)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         84 CCAGATGTGATATTTCTACAGGAAGTT 110
```

RESULT 21
BF891509
LOCUS BF891509 447 bp mRNA linear EST 18-JAN-2001
DEFINITION PM1-MT0143-281000-001-h02 MT0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF891509
VERSION BF891509.1 GI:12282981
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 447)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-MT0143-281000-001-h02&t3=2000-10-28&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 27
High quality sequence stop: 447.

FEATURES Location/Qualifiers
source 1. .447
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0143"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	20.5	Length:	447
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BF891509 (1-447)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         88 CCAGATGTGATATTTCTACAGGAAGTT 114
```

RESULT 22
BF773259/c
LOCUS BF773259 453 bp mRNA linear EST 12-JAN-2001
DEFINITION PM0-IT0018-151200-002-d10 IT0018 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF773259
VERSION BF773259.1 GI:12121159
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 453)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM0&t2=PM0-IT0018-151200-002-d10&t3=2000-12-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 28
High quality sequence stop: 133.

FEATURES Location/Qualifiers
source 1. .453
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="IT0018"
/note="Organ: epid_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	20.8	Length:	453
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BF773259 (1-453)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        373 CCAGATGTGATATTTCTACAGGAAGTT 347
```

RESULT 23
BQ315535/c
LOCUS BQ315535 453 bp mRNA linear EST 16-MAY-2002
DEFINITION PM0-IT0018-151200-002-d06 IT0018 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ315535
VERSION BQ315535.1 GI:20874720
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 453)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM0&t2=PM0-IT0018-151200-002-d06&t3=2000-12-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 20.

FEATURES Location/Qualifiers
source 1. .453
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="IT0018"
/note="Organ: epid_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	20.8	Length:	453
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BQ315535 (1-453)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        373 CCAGATGTGATATTTCTACAGGAAGTT 347
```

RESULT 24
BG982031/c
LOCUS BG982031 454 bp mRNA linear EST 12-JUN-2001
DEFINITION MR3-CN0147-120201-011-b12 CN0147 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG982031
VERSION BG982031.1 GI:14384766
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 454)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-CN0147-120201-011-b12&t3=2001-02-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 453.

FEATURES Location/Qualifiers
source 1. .454
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CN0147"
/note="Organ: colon_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	20.8	Length:	454
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

```
US-10-757-745-2_COPY_145_153 (1-9) x BG982031 (1-454)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db        373 CCAGATGTGATATTTCTACAGGAGGTT 347

RESULT 25
BF996980/c
LOCUS       BF996980                456 bp    mRNA    linear   EST 23-JAN-2001
DEFINITION  QV3-GN0201-081100-468-c10 GN0201 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF996980
VERSION     BF996980.1  GI:12403303
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-GN0201-
            081100-468-c10&t3=2000-11-08&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 456.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0201"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                20.9          Length:          456
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       2             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BF996980 (1-456)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||
Db        376 CCAGATGTGATATTTCTACAGGAAGTT 350

RESULT 26
BQ332321
LOCUS       BQ332321                456 bp    mRNA    linear   EST 17-MAY-2002
DEFINITION  MR4-ET0140-240501-015-d10 ET0140 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ332321
VERSION     BQ332321.1  GI:20973744
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-ET0140-
            240501-015-d10&t3=2001-05-24&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 456.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="ET0140"
                     /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                20.9          Length:          456
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BQ332321 (1-456)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||
Db         96 CCAGATGTGATATTTCTACAGGAAGTT 122

RESULT 27
BQ366782/c
LOCUS       BQ366782                458 bp    mRNA    linear   EST 21-MAY-2002
DEFINITION  QV3-GN0201-081100-437-c05 GN0201 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ366782
VERSION     BQ366782.1  GI:21042296
```

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 458)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV3&t2=QV3-GN0201-
            081100-437-c05&t3=2000-11-08&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 12.
FEATURES             Location/Qualifiers
     source          1. .458
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="GN0201"
                     /note="Organ: placenta_normal; Vector: puc18; Site_1:
                     SmaI; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                 21            Length:          458
Score:                     45.00         Matches:         9
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               100.0%        Indels:          0
DB:                        3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BQ366782 (1-458)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        378 CCAGATGTGATATTTCTACAGGAAGTT 352

RESULT 28
BG982034/c
LOCUS       BG982034                 459 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION  MR3-CN0147-120201-011-c09 CN0147 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG982034
VERSION     BG982034.1  GI:14384769
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 459)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-CN0147-
            120201-011-c09&t3=2001-02-12&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 19
            High quality sequence stop: 459.
FEATURES             Location/Qualifiers
     source          1. .459
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CN0147"
                     /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                 21.1          Length:          459
Score:                     45.00         Matches:         9
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               100.0%        Indels:          0
DB:                        2             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BG982034 (1-459)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        379 CCAGATGTGATATTTCTACAGGAAGTT 353

RESULT 29
BF893150
LOCUS       BF893150                 460 bp    mRNA    linear   EST 18-JAN-2001
DEFINITION  PM1-MT0143-101100-003-b07 MT0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF893150
VERSION     BF893150.1  GI:12284609
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
```

```
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-MT0143-
            101100-003-b07&t3=2000-11-10&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 7
            High quality sequence stop: 434.
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="MT0143"
                     /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                21.1          Length:         460
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BF893150 (1-460)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        100 CCAGATGTGATATTTCTACAGGAAGTT 126

RESULT 30
BG982012/c
LOCUS       BG982012                461 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION  MR3-CN0147-120201-011-b03 CN0147 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG982012
VERSION     BG982012.1  GI:14384747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-CN0147-
            120201-011-b03&t3=2001-02-12&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 191.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CN0147"
                     /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                21.2          Length:         461
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BG982012 (1-461)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        381 CCAGATGTGATATTTCTGCAGGAGGTT 355

RESULT 31
BF852284
LOCUS       BF852284                463 bp    mRNA    linear   EST 16-JAN-2001
DEFINITION  MR3-EN0087-121200-007-d01 EN0087 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF852284
VERSION     BF852284.1  GI:12239655
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 463)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
```

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-EN0087-121200-007-d01&t3=2000-12-12&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 463.

FEATURES Location/Qualifiers
source 1. .463
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0087"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	21.3	Length:	463
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BF852284 (1-463)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         95 CCAGATGTGATATTTCTACAGGAAGTT 121
```

RESULT 32
BF852788/c
LOCUS BF852788 463 bp mRNA linear EST 16-JAN-2001
DEFINITION MR3-EN0087-151200-012-a09 EN0087 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF852788
VERSION BF852788.1 GI:12240532
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 463)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-EN0087-151200-012-a09&t3=2000-12-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 31
High quality sequence stop: 463.

FEATURES Location/Qualifiers
source 1. .463
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0087"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	21.3	Length:	463
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BF852788 (1-463)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        383 CCAGATGTGATATTTCTACAGGAAGTT 357
```

RESULT 33
BF893919
LOCUS BF893919 463 bp mRNA linear EST 18-JAN-2001
DEFINITION PM1-MT0143-131100-004-e06 MT0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF893919
VERSION BF893919.1 GI:12285378
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 463)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-MT0143-131100-004-e06&t3=2000-11-13&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 463.

FEATURES Location/Qualifiers

```
     source          1. .463
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="MT0143"
                     /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                21.3          Length:         463
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BF893919 (1-463)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        103 CCAGATGTGATATTTCTACAGGAAGTT 129

RESULT 34
BF893924/c
LOCUS       BF893924                471 bp    mRNA    linear   EST 18-JAN-2001
DEFINITION  PM1-MT0143-131100-004-g07 MT0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF893924
VERSION     BF893924.1  GI:12285383
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 471)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-MT0143-
            131100-004-g07&t3=2000-11-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 33
            High quality sequence stop: 471.
FEATURES             Location/Qualifiers
     source          1. .471
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="MT0143"
                     /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                21.7          Length:         471
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BF893924 (1-471)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        391 CCAGATGTGATATTTCTACAGGAAGTT 365

RESULT 35
CR543841
LOCUS       CR543841                471 bp    mRNA    linear   EST 07-JUL-2004
DEFINITION  DKFZp459O0742_r1 459 (synonym: pcor1) Pongo pygmaeus cDNA clone
            DKFZp459O0742 5', mRNA sequence.
ACCESSION   CR543841
VERSION     CR543841.1  GI:49895253
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 471)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C.,
            Osanger,A., Fobo,G., Han,M. and Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Bloecker,H., Boecher,M., Brandt,P., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National
            Research Centre for Biotechnology Ltd., Braunschweig/Germany)
            within the cDNA sequencing consortium of the German Genome Project.
            This clone (DKFZp459O0742) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .471
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp459O0742"
                     /tissue_type="cortex"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="459 (synonym: pcor1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:                21.7          Length:         471
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
```

```
Query Match:            100.0%          Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CR543841 (1-471)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        434 CCAGATGTGATATTTCTACAGGAAGTT 460

RESULT 36
CB120234
LOCUS       CB120234                  473 bp    mRNA    linear   EST 28-JAN-2003
DEFINITION  K-EST0167337 L8SCK0 Homo sapiens cDNA clone L8SCK0-29-F06 5', mRNA
            sequence.
ACCESSION   CB120234
VERSION     CB120234.1  GI:27946036
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 473)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 29  row: F  column: 06
            High quality sequence stop: 473.
FEATURES             Location/Qualifiers
     source          1. .473
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L8SCK0-29-F06"
                     /sex="M"
                     /cell_line="SCK"
                     /lab_host="Top10F'"
                     /clone_lib="L8SCK0"
                     /note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI;
                     Site_2: NotI; The library was contributed by the Soares
                     laboratory and it was constructed as described by Bonaldo,
                     M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
                     6(9): 791-806. RNA was prepared from harvested cell
                     culture."
ORIGIN

Alignment Scores:
Pred. No.:              21.8            Length:         473
Score:                  45.00           Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CB120234 (1-473)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        382 CCAGATGTGATATTTCTACAGGAAGTT 408

RESULT 37
AI750554
LOCUS       AI750554                  480 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  cn04a01.y1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
            clone NHTBC_cn04a01 random, mRNA sequence.
ACCESSION   AI750554
VERSION     AI750554.1  GI:5128818
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
            Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
            Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
  TITLE     SGAP: The Skeletal Genome Anatomy Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Libin Jia
            Medical Genetics Branch
            National Human Genome Research Institute
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libin@helix.nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 04  row: a  column: 01
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NHTBC_cn04a01"
                     /sex="Female"
                     /tissue_type="Bone"
                     /cell_type="Trabecular Bone Cells"
                     /lab_host="SURE"
                     /clone_lib="Normal Human Trabecular Bone Cells"
                     /note="Organ: Hip; Vector: pBluescript; Site_1: EcoRI;
                     Library constructed by Dr. Marian Young and Dr. Pamela
                     Gehron Robey (NIDCR)"
ORIGIN

Alignment Scores:
Pred. No.:              22.2            Length:         480
Score:                  45.00           Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AI750554 (1-480)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        427 CCAGATGTGATATTTCTACAGGAAGTT 453

RESULT 38
AA486032
LOCUS       AA486032                  485 bp    mRNA    linear   EST 06-MAR-1998
DEFINITION  ab40b10.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone
            IMAGE:843259 5', mRNA sequence.
ACCESSION   AA486032
VERSION     AA486032.1  GI:2216248
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
```

```
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1913   Std Error: 0.00
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:843259"
                     /sex="female"
                     /dev_stage="HeLa S3 cell line"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene HeLa cell s3 937216"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. HeLa S3
                     epithelioid carcinoma cells grown to semi-confluency
                     without induction. Average insert size: 1.5 kb; Uni-ZAP XR
                     Vector. ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

Alignment Scores:
Pred. No.:                22.4          Length:          485
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       1             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AA486032 (1-485)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        423 CCAGATGTGATATTTCTACAGGAAGTT 449

RESULT 39
AJ681912
LOCUS       AJ681912                 491 bp    mRNA    linear   EST 29-JUN-2004
DEFINITION  AJ681912 CSEQRAN04 Sus scrofa cDNA clone C0001797_I20, mRNA
            sequence.
ACCESSION   AJ681912
VERSION     AJ681912.1  GI:49414502
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 491)
  AUTHORS   Anderson,S.I., Finlayson,H.A. and Archibald,A.L.
  TITLE     Development of cDNA and EST resources for studying reproduction and
            embryo development in pigs and cattle
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Anderson SI
            Genomics and Bioinformatics
            Roslin Institute
            Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
            Single pass sequencing. Bases called and trimmed with phred
            v0.020425.c. Vector identified by cross_match with the -minscore 20
            and -minmatch 12 options. Vector:pBlueScriptII(KS+) R. Site1: EcoRI
            R. Site2: NotI 5' Seq Primer M13F Normalised library constructed
            from pig uterus. Clones available from UK Centre for Functional
            Genomics in Farm Animals, Roslin Institite, Roslin, Midlothian, UK,
            EH25 9PS, www.arkgenomics.org.
FEATURES             Location/Qualifiers
     source          1. .491
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /clone="C0001797_I20"
                     /tissue_type="uterus"
                     /clone_lib="CSEQRAN04"
                     /note="Vector: pBlueScriptII(KS+); Site_1: EcoRI; Site_2:
                     NotI; Single pass sequencing. Normalised library
                     constructed from pig uterus."
ORIGIN

Alignment Scores:
Pred. No.:                22.8          Length:          491
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       1             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AJ681912 (1-491)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        127 CCAGATGTGATATTTCTACAGGAAGTT 153

RESULT 40
BI016235/c
LOCUS       BI016235                 491 bp    mRNA    linear   EST 13-JUN-2001
DEFINITION  MR4-ET0140-130301-009-a09 ET0140 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI016235
VERSION     BI016235.1  GI:14420306
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 491)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-ET0140-
            130301-009-a09&t3=2001-03-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 7
            High quality sequence stop: 232.
FEATURES             Location/Qualifiers
     source          1. .491
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
```

/clone_lib="ET0140"
/note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	22.8	Length:	491
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BI016235 (1-491)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        410 CCAGATGTGATATTTCTACAGGAAGTT 384
```

RESULT 41
BM846221
LOCUS BM846221 530 bp mRNA linear EST 06-MAR-2002
DEFINITION K-EST0125204 S13KMS5 Homo sapiens cDNA clone S13KMS5-59-C04 5', mRNA sequence.
ACCESSION BM846221
VERSION BM846221.1 GI:19202620
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 530)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 59 row: C column: 04
High quality sequence stop: 530.

FEATURES Location/Qualifiers
source 1. .530
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S13KMS5-59-C04"
/tissue_type="myeloma"
/cell_line="KMS-5"
/lab_host="Top10F'"
/clone_lib="S13KMS5"
/note="Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Alignment Scores:

Pred. No.:	24.8	Length:	530
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BM846221 (1-530)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        437 CCAGATGTGATATTTCTACAGGAAGTT 463
```

RESULT 42
CN298921
LOCUS CN298921 534 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000531590248 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN298921
VERSION CN298921.1 GI:47315335
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 534)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 534 Std Error: 0.00.

FEATURES Location/Qualifiers
source 1. .534
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="oligo dT primed, full-length enriched cDNA library from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."

ORIGIN

Alignment Scores:

Pred. No.:	25	Length:	534
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x CN298921 (1-534)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db         40 CCAGATGTGATATTTCTACAGGAAGTT 66

RESULT 43
BG541031
LOCUS       BG541031                546 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602570190F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4694374 5',
            mRNA sequence.
ACCESSION   BG541031
VERSION     BG541031.1  GI:13533264
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1518  row: d  column: 23
            High quality sequence stop: 545.
FEATURES             Location/Qualifiers
     source          1. .546
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4694374"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_77"
                     /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.9
                     kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                25.7          Length:         546
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       2             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BG541031 (1-546)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db         39 CCAGATGTGATATTTCTACAGGAAGTT 65

RESULT 44
AW669835
LOCUS       AW669835                549 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  113421 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   AW669835
VERSION     AW669835.1  GI:7526349
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 549)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
            Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
            Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,
            Quackenbush,J. and Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
   PUBMED   11282978
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 108  row: M  column: 18
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 1BOV"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from lymph node, ovary,
                     fat, hypothalamus, and pituitary."
ORIGIN

Alignment Scores:
Pred. No.:                25.9          Length:         549
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       1             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AW669835 (1-549)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db        498 CCTGATGTGATATTTCTACAGGAAGTC 524

RESULT 45
BP220985
LOCUS       BP220985                554 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP220985 Sugano cDNA library, colon Homo sapiens cDNA clone
            COL06450, mRNA sequence.
ACCESSION   BP220985
VERSION     BP220985.1  GI:52093890
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 554)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
```

```
          block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .554
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL06450"
                     /tissue_type="colon"
                     /clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:
Pred. No.:              26.1          Length:         554
Score:                  45.00         Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP220985 (1-554)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        342 CCAGATGTGATATTTCTACAGGAAGTT 368

RESULT 46
BP226620
LOCUS       BP226620                565 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP226620 Sugano cDNA library, dermoid cancer Homo sapiens cDNA
          clone DMC03190, mRNA sequence.
ACCESSION   BP226620
VERSION     BP226620.1  GI:52099525
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE   1  (bases 1 to 565)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
          Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
          block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .565
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DMC03190"
                     /clone_lib="Sugano cDNA library, dermoid cancer"
                     /note="dermoid cancer"
ORIGIN

Alignment Scores:
Pred. No.:              26.7          Length:         565
Score:                  45.00         Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP226620 (1-565)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        434 CCAGATGTGATATTTCTACAGGAAGTT 460

RESULT 47
BP220508
LOCUS       BP220508                566 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP220508 Sugano cDNA library, colon Homo sapiens cDNA clone
          COL03160, mRNA sequence.
ACCESSION   BP220508
VERSION     BP220508.1  GI:52093413
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
          Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
          block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .566
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL03160"
                     /tissue_type="colon"
                     /clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:
Pred. No.:              26.8          Length:         566
Score:                  45.00         Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP220508 (1-566)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        352 CCAGATGTGATATTTCTACAGGAAGTT 378

RESULT 48
BP257211
LOCUS       BP257211                568 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP257211 Sugano cDNA library, heart Homo sapiens cDNA clone
          HRT00430, mRNA sequence.
ACCESSION   BP257211
VERSION     BP257211.1  GI:52172441
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
          Mizushima-Sugano,J., Nakai,K. and Sugano,S.
```

```
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .568
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRT00430"
                     /tissue_type="heart"
                     /clone_lib="Sugano cDNA library, heart"
ORIGIN

Alignment Scores:
Pred. No.:                 26.9          Length:          568
Score:                     45.00         Matches:         9
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               100.0%        Indels:          0
DB:                        3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BP257211 (1-568)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        403 CCAGATGTGATATTTCTACAGGAAGTT 429

RESULT 49
BM172060
LOCUS       BM172060                 570 bp    mRNA    linear   EST 04-DEC-2001
DEFINITION  imageqc_3_2001/sml293bdff41.x1 NIH_MGC_76 Homo sapiens cDNA clone
            IMAGE:4700059 5', mRNA sequence.
ACCESSION   BM172060
VERSION     BM172060.1  GI:17311623
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Kale,P.I., Harsch,T.J., Folta,P.A., Nelson,D.O., Sanders,C.G. and
            Prange,C.K.
  TITLE     The I.M.A.G.E. Consortium quality control effort: clone
            resequencing for verification
  JOURNAL   Unpublished (2001)
COMMENT     Other_ESTs: BG533717
            Contact: Prange CK
            The I.M.A.G.E. Consortium
            Lawrence Livermore National Laboratory
            Livermore, CA, USA
            Email: help@image.llnl.gov
            This read has been verified (found to hit its original self in the
            correct orientation), as part of the I.M.A.G.E. Consortium quality
            control effort. High quality sequence is defined as having 100 or
            more base pairs with a phred quality value of 20 or greater, where
            a sliding window of 4 base pairs with a phred quality value of 15
            or greater marks the beginning and end of the sequence. For
            information on obtaining this clone, please contact
            info@image.llnl.gov.
            Plate: LLCM1533  row: a  column: 20
            Seq primer: -21m13
            High quality sequence stop: 570.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4700059"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_76"
                     /note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.85
                     kb (range 1.0-4.0 kb). 15/15 colonies contained  inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                 27            Length:          570
Score:                     45.00         Matches:         9
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               100.0%        Indels:          0
DB:                        3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BM172060 (1-570)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        285 CCAGATGTGATATTTCTACAGGAAGTT 311

RESULT 50
BP221518
LOCUS       BP221518                 570 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP221518 Sugano cDNA library, colon Homo sapiens cDNA clone
            COL09203, mRNA sequence.
ACCESSION   BP221518
VERSION     BP221518.1  GI:52094423
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL09203"
                     /tissue_type="colon"
                     /clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:
Pred. No.:                 27            Length:          570
Score:                     45.00         Matches:         9
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               100.0%        Indels:          0
DB:                        3             Gaps:            0
```

US-10-757-745-2_COPY_145_153 (1-9) x BP221518 (1-570)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db        341 CCAGATGTGATATTTCTACAGGAAGTT 367
```

RESULT 51
BP258309

```
LOCUS       BP258309                570 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP258309 Sugano cDNA library, heart Homo sapiens cDNA clone
            HRT03149, mRNA sequence.
ACCESSION   BP258309
VERSION     BP258309.1  GI:52173539
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRT03149"
                     /tissue_type="heart"
                     /clone_lib="Sugano cDNA library, heart"
ORIGIN
```

Alignment Scores:

Pred. No.:	27	Length:	570
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BP258309 (1-570)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db         78 CCAGATGTGATATTTCTACAGGAAGTT 104
```

RESULT 52
BP262485

```
LOCUS       BP262485                572 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262485 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI07111, mRNA sequence.
ACCESSION   BP262485
VERSION     BP262485.1  GI:52177716
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 572)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .572
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI07111"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN
```

Alignment Scores:

Pred. No.:	27.1	Length:	572
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BP262485 (1-572)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db        389 CCAGATGTGATATTTCTACAGGAAGTT 415
```

RESULT 53
AU279894

```
LOCUS       AU279894                573 bp    mRNA    linear   EST 31-JUL-2003
DEFINITION  AU279894 CHONS2 Homo sapiens cDNA clone CHONS2002038 5', mRNA
            sequence.
ACCESSION   AU279894
VERSION     AU279894.1  GI:28299121
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Imabayashi,H., Mori,T., Gojo,S., Kiyono,T., Sugiyama,T., Irie,R.,
            Isogai,T., Hata,J., Tomoya,Y. and Umezawa,A.
  TITLE     Redifferentiation of dedifferentiated chondrocytes and
            chondrogenesis of human bone marrow stromal cells via chondrosphere
            formation with expression profiling by large-scale cDNA analysis
  JOURNAL   Exp. Cell Res. 288 (1), 35-50 (2003)
   PUBMED   12878157
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA Project, Sugiyama,T.; Wakamatsu,A.; Irie,R.;
            Umezawa,A.; Fukuma,M.; Kusakari,S.; Hata,J.; Ishii,S.; Yamamoto,J.;
            Isono,Y.; Saito,K.; Nakamura,Y.; Masuho,Y.; Nagai,K.; Isogai,T.
             HRI human cDNA project; cDNA library construction & 5'-end one
            pass sequencing: Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CHONS2002038"
                     /cell_type="chondrocytes"
                     /clone_lib="CHONS2"
                     /note="Vector: pME18SFL3"
```

```
ORIGIN

Alignment Scores:
Pred. No.:                27.1          Length:         573
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       1             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AU279894 (1-573)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        470 CCAGATGTGATATTTCTACAGGAAGTT 496

RESULT 54
BP221260
LOCUS       BP221260                 573 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP221260 Sugano cDNA library, colon Homo sapiens cDNA clone
            COL08083, mRNA sequence.
ACCESSION   BP221260
VERSION     BP221260.1  GI:52094165
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL08083"
                     /tissue_type="colon"
                     /clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:
Pred. No.:                27.1          Length:         573
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP221260 (1-573)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        355 CCAGATGTGATATTTCTACAGGAAGTT 381

RESULT 55
BP261500
LOCUS       BP261500                 573 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP261500 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI03816, mRNA sequence.
ACCESSION   BP261500
VERSION     BP261500.1  GI:52176731
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI03816"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                27.1          Length:         573
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP261500 (1-573)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        339 CCAGATGTGATATTTCTACAGGAAGTT 365

RESULT 56
BP262642
LOCUS       BP262642                 576 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262642 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI07584, mRNA sequence.
ACCESSION   BP262642
VERSION     BP262642.1  GI:52177873
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 576)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .576
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI07584"
                     /tissue_type="small intestine"
```

/clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:

Pred. No.:	27.3	Length:	576
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BP262642 (1-576)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        357 CCAGATGTGATATTTCTACAGGAAGTT 383
```

RESULT 57
CN298919
LOCUS CN298919 577 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000600175228 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN298919
VERSION CN298919.1 GI:47315333
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 577)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 577 Std Error: 0.00.
FEATURES Location/Qualifiers
source 1. .577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, DMSO-treated H9 cell line"
/clone_lib="GRN_PREHEP"
/note="oligo dT primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"
ORIGIN

Alignment Scores:

Pred. No.:	27.4	Length:	577
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x CN298919 (1-577)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459
```

RESULT 58
BP261787
LOCUS BP261787 580 bp mRNA linear EST 16-SEP-2004
DEFINITION BP261787 Sugano cDNA library, small intestine Homo sapiens cDNA clone HSI05008, mRNA sequence.
ACCESSION BP261787
VERSION BP261787.1 GI:52177018
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HSI05008"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:

Pred. No.:	27.5	Length:	580
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BP261787 (1-580)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        368 CCAGATGTGATATTTCTACAGGAAGTT 394
```

RESULT 59
BP274368
LOCUS BP274368 580 bp mRNA linear EST 16-SEP-2004
DEFINITION BP274368 Sugano cDNA library, kidney Homo sapiens cDNA clone KDN00894, mRNA sequence.
ACCESSION BP274368
VERSION BP274368.1 GI:52188100
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.

```
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KDN00894"
                     /tissue_type="kidney"
                     /clone_lib="Sugano cDNA library, kidney"
ORIGIN

Alignment Scores:
Pred. No.:              27.5          Length:         580
Score:                  45.00         Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP274368 (1-580)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        453 CCAGATGTGATATTTCTACAGGAAGTT 479

RESULT 60
AV717253
LOCUS       AV717253                581 bp    mRNA    linear   EST 16-OCT-2000
DEFINITION  AV717253 DCB Homo sapiens cDNA clone DCBBMA04 5', mRNA sequence.
ACCESSION   AV717253
VERSION     AV717253.1  GI:10814405
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
  TITLE     Homo sapiens cDNA DCB clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBBMA04"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

Alignment Scores:
Pred. No.:              27.6          Length:         581
Score:                  45.00         Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     1             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AV717253 (1-581)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        276 CCAGATGTGATATTTCTACAGGAAGTT 302

RESULT 61
BP207751
LOCUS       BP207751                581 bp    mRNA    linear   EST 14-SEP-2004
DEFINITION  BP207751 Sugano cDNA library, coronary artery smooth muscle cell
            Homo sapiens cDNA clone CAS11735, mRNA sequence.
ACCESSION   BP207751
VERSION     BP207751.1  GI:52064157
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CAS11735"
                     /tissue_type="coronary artery"
                     /cell_type="smooth muscle cell"
                     /clone_lib="Sugano cDNA library, coronary artery smooth
                     muscle cell"
ORIGIN

Alignment Scores:
Pred. No.:              27.6          Length:         581
Score:                  45.00         Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP207751 (1-581)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        341 CCAGATGTGATATTTCTACAGGAAGTT 367

RESULT 62
BP225088
LOCUS       BP225088                581 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP225088 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA
            clone DAT05184, mRNA sequence.
ACCESSION   BP225088
VERSION     BP225088.1  GI:52097993
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
```

```
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DAT05184"
                     /cell_type="lymphocyte"
                     /cell_line="Daudi"
                     /clone_lib="Sugano cDNA library, lymphocyte Daudi"
                     /note="Burkitt's lymphoma"
ORIGIN

Alignment Scores:
Pred. No.:                 27.6          Length:          581
Score:                     45.00         Matches:         9
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               100.0%        Indels:          0
DB:                        3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BP225088 (1-581)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        454 CCAGATGTGATATTTCTACAGGAAGTT 480

RESULT 63
BP262101
LOCUS       BP262101                581 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262101 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI05893, mRNA sequence.
ACCESSION   BP262101
VERSION     BP262101.1  GI:52177332
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI05893"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                 27.6          Length:          581
Score:                     45.00         Matches:         9
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               100.0%        Indels:          0
DB:                        3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BP262101 (1-581)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         61 CCAGATGTGATATTTCTACAGGAAGTT 87

RESULT 64
BP262638
LOCUS       BP262638                581 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262638 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI07576, mRNA sequence.
ACCESSION   BP262638
VERSION     BP262638.1  GI:52177869
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI07576"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                 27.6          Length:          581
Score:                     45.00         Matches:         9
Percent Similarity:        100.0%        Conservative:    0
Best Local Similarity:     100.0%        Mismatches:      0
Query Match:               100.0%        Indels:          0
DB:                        3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BP262638 (1-581)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        350 CCAGATGTGATATTTCTACAGGAAGTT 376

RESULT 65
BP270744
LOCUS       BP270744                581 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP270744 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone KAR05146, mRNA sequence.
ACCESSION   BP270744
VERSION     BP270744.1  GI:52220093
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
```

```
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAR05146"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                27.6          Length:         581
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP270744 (1-581)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        341 CCAGATGTGATATTTCTACAGGAAGTT 367

RESULT 66
BP349936
LOCUS       BP349936                581 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP349936 Sugano cDNA library, brain Homo sapiens cDNA clone
            SZR06923, mRNA sequence.
ACCESSION   BP349936
VERSION     BP349936.1  GI:52279921
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SZR06923"
                     /tissue_type="brain"
                     /clone_lib="Sugano cDNA library, brain"
ORIGIN

Alignment Scores:
Pred. No.:                27.6          Length:         581
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP349936 (1-581)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        210 CCAGATGTGATATTTCTACAGGAAGTT 236

RESULT 67
BP379659
LOCUS       BP379659                581 bp    mRNA    linear   EST 21-SEP-2004
DEFINITION  BP379659 Sugano cDNA library, uterus Homo sapiens cDNA clone
            WMD06642, mRNA sequence.
ACCESSION   BP379659
VERSION     BP379659.1  GI:52412762
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="WMD06642"
                     /tissue_type="uterus"
                     /clone_lib="Sugano cDNA library, uterus"
ORIGIN

Alignment Scores:
Pred. No.:                27.6          Length:         581
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP379659 (1-581)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        342 CCAGATGTGATATTTCTACAGGAAGTT 368

RESULT 68
CB286391
LOCUS       CB286391                581 bp    mRNA    linear   EST 27-FEB-2003
DEFINITION  CMD36_C09_69 UMNMPM3 Sus scrofa cDNA clone PPSUBLIB_21C09 5', mRNA
            sequence.
ACCESSION   CB286391
VERSION     CB286391.1  GI:28577345
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
```

REFERENCE 1 (bases 1 to 581)
AUTHORS Dvorak,C.M.T., Hyland,K.A., Machado,J.G., Zhang,Y., Fahrenkrug,S.C. and Murtaugh,M.P.
TITLE Gene discovery and expression profiling in porcine Peyer's patch
JOURNAL Vet. Immunol. Immunopathol. 105 (3-4), 301-315 (2005)
PUBMED 15808308
COMMENT Contact: Murtaugh, M.P.
Department of Veterinary PathoBiology
University of Minnesota
1971 Commonwealth Ave., St. Paul, MN 55108, USA
Tel: 612-625-6735
Fax: 612-625-5203
Email: murta001@umn.edu
The PHRED quality scores of this sequence are supplied below: 35 40
45 42 44 42 51 46 42 37 37 37 35 40 35 35 22 35 35 35 29 35 35 35
30 36 28 33 33 45 45 45 45 51 51 46 46 36 42 47 37 37 46 46 46
46 51 51 46 43 42 42 33 46 26 33 33 45 45 45 45 40 37 37 40 40 40
37 50 56 56 50 50 42 37 35 45 45 45 37 50 44 50 50 56 51 45 45 37
37 37 40 34 34 28 28 28 28 42 56 50 56 56 46 46 40 38 38 38 56 56
56 56 50 50 44 42 33 46 46 51 43 51 51 51 51 51 51 39 35 32 26 26
29 35 33 33 27 42 50 51 56 56 56 56 56 56 56 43 40 38 35 35 35 41 41
45 46 51 51 56 56 56 45 41 41 41 45 46 56 56 56 56 50 50 50 50 36
36 36 36 36 36 40 40 35 50 35 42 40 41 41 41 41 41 43 46 51 56 56
56 56 56 56 51 46 46 42 42 42 40 40 37 35 35 35 40 40 40 43 43 42
42 40 40 35 40 40 40 42 37 35 40 42 51 51 51 56 56 56 56 56 56 56
46 42 43 43 46 46 51 37 37 32 35 35 39 50 37 35 32 35 35 35 35 40
35 35 40 45 45 51 51 45 45 45 45 45 51 43 46 42 42 42 44 48 56
40 40 40 40 40 40 56 56 56 46 42 42 42 37 40 40 40 40 39 40 40 42
42 44 42 42 42 48 44 56 56 56 56 56 56 46 46 42 34 33 40 40 40 35
35 35 35 35 35 51 51 35 35 35 35 35 35 39 37 37 35 35 35 31 34 39
40 40 40 40 40 37 34 22 25 32 32 32 39 39 29 29 29 29 29 35 29 35
35 35 40 40 42 37 42 46 46 51 45 51 51 46 51 42 42 42 51 56 43 43
42 51 51 51 51 51 51 51 51 51 51 35 35 35 35 33 35 40 40 35 35 35
34 34 29 48 36 40 31 48 27 26 19 25 27 27 32 34 29 29 31 34 46 46
56 56 51 40 40 40 40 40 45 45 45 35 35 35 46 46 51 51 56 56 56 37
32 22 22 22 32 46 56 48 48 48 48 48 32 27 23 18 15 15 29 48 40 40
40 44 44 40 40 26 29 24 18 15 24 29 48 48 48 40 29 25 19 20 24 23
23 21 18 25 48 48 48 48 40 40 32 29 22 25 15 16 11 10 10 20 24 26
25 29 29 25 27 25 22 22 32 32 27 29 29 34 34 40 40 29 20 15 11 11
19 24 33 40 40 46 40 37 37 22 22 21 27 23 22 22 25 29 29 29 34 40
40 32 32 26 32 26 24 Sequences were generated on ABI 3100 and 3700 capillary sequencers. Raw sequence data was batch processed analyzed, groomed and submitted by the Animal Biotechnology Center (URL: http://primer.ansci.umn.edu/software1.html),University of Minnesota. Trace files have been submitted to the NCBI trace archive(http://www.ncbi.nlm.nih.gov/Traces/trace.cgi?).
Chromatograms were analyzed with Phred Version: 0.000925.c . Vector (pCMV_SPORT6, Invitrogen), Sus scrofa mitochondrion (gi|5835862|ref|NC_000845.1),Escherichia coli K12(gi|5835862|ref|NC_000913.1|),Porcine reproductive and respiratory syndrome virus(gi|11878202|gb|AF303357.1|AF303357),Alteromonas sp.16S rRNA gene(gi|4218471|emb|Y15322.1|PLY15322),Sus scrofa 28S ribosomal RNA(gi|3388170|gb|AF080393.1|AF080393),and Sus scrofa 18S ribosomal RNA(gi|3885982|gb|AF102857.1|AF102857) Sequences were identified and masked by Crossmatch -minscore=20, -minmatch=12, -penalty=-8. SWAT alignment scores were set to match score=1, gap initial penalty=-3 and gap extension penalty=-2. A continuous stretch of at least 50 bases of a PHRED quality 20 were required for submission.5' and 3' TERMINI were defined by a continuous stretch of 10 nucleotides with a PHRED quality not less than 20. A terminal stretch of at most 40 un-masked bases were automatically trimmed when flanking masked sequence.Sequences with an INTERNAL continuous stretch of at most 20 bases with PHRED quality less than 20 were automatically prepared for submission. HIQUAL_START and HIQUAL_STOP refer to the coordinates on CLIP_QUALITY and CLIP_VECTOR at the NCBI trace archive.
Plate: 21 row: C column: 09
Seq primer: SP6 primer
High quality sequence start: 67
High quality sequence stop: 302.

FEATURES Location/Qualifiers
source 1. .581
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="PPSUBLIB_21C09"
/tissue_type="Peyer's patch"
/clone_lib="UMNMPM3"
/note="Organ: small intestine, jejunum; Vector: pCMVSPORT6; Site_1: NotI; Site_2: SalI; Jejunal Peyer's patches were isolated from healthy, 4-6 week old cross-bred pigs. RNA was extracted either immediately or after 3 hours stimulation in an Ussing chamber with one of the following treatments: 1) no treatment, 2) Salmonella choleraesuis vaccine strain SC-54, 3) phorbol myristate acetate, concanavalin A, and 8-bromo-cyclic AMP, and 4) lipopolysaccharide and cholera toxin. Each treatment was performed in the presence and absence of cycloheximide. Purified poly A + RNA from each of the treated tissues (2-4) was combined together, reverse transcribed, and cloned in to pCMVSPORT6 to make a library of approximately 530,000 recombinant clones with an average insert size of 1.0 kbp. Poly A+ RNA from freshly isolated and non-treated tissue from an Ussing chamber (treatment 1) was cloned in the same manner to produce an unstimulated cDNA library of approximately 900,000 clones with an average insert size of 1.5 kbp. Equal portions of the two libraries were pooled and then subtracted with porcine ST fibroblast RNA to create a subtracted porcine Peyer's patch library of approximately 6,000 clones with an average insert size of 1.0 kbp."

ORIGIN

Alignment Scores:

Pred. No.:	27.6	Length:	581
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x CB286391 (1-581)

```
QY          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        265 CCAGATGTGATATTTCTACAGGAAGTT 291
```

RESULT 69
BM507091
LOCUS BM507091 582 bp mRNA linear EST 15-FEB-2002
DEFINITION ih24h01.y1 Human insulinoma Homo sapiens cDNA 5' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION BM507091
VERSION BM507091.1 GI:18678234
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other_ESTs: ih24h01.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812

Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. J. Ferrer In vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 269.

FEATURES Location/Qualifiers
source 1. .582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="insulinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Human insulinoma"
/note="Organ: pancreas; Vector: pBluescript SK-; Site_1: XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library. "

ORIGIN

Alignment Scores:

Pred. No.:	27.6	Length:	582
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BM507091 (1-582)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        235 CCAGATGTGATATTTCTACAGGAAGTT 261
```

RESULT 70
BP219740
LOCUS BP219740 582 bp mRNA linear EST 15-SEP-2004
DEFINITION BP219740 Sugano cDNA library, caudate nucleus Homo sapiens cDNA clone CNR08985, mRNA sequence.
ACCESSION BP219740
VERSION BP219740.1 GI:52092643
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES Location/Qualifiers
source 1. .582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CNR08985"
/tissue_type="caudate nucleus"
/clone_lib="Sugano cDNA library, caudate nucleus"

ORIGIN

Alignment Scores:

Pred. No.:	27.6	Length:	582
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BP219740 (1-582)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        448 CCAGATGTGATATTTCTACAGGAAGTT 474
```

RESULT 71
BP233592
LOCUS BP233592 582 bp mRNA linear EST 15-SEP-2004
DEFINITION BP233592 Sugano cDNA library, coronary artery endothelial cell Homo sapiens cDNA clone HCR00774, mRNA sequence.
ACCESSION BP233592
VERSION BP233592.1 GI:52106502
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES Location/Qualifiers
source 1. .582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HCR00774"
/tissue_type="coronary artery"
/cell_type="endothelial cell"
/clone_lib="Sugano cDNA library, coronary artery endothelial cell"

ORIGIN

Alignment Scores:

Pred. No.:	27.6	Length:	582
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BP233592 (1-582)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        448 CCAGATGTGATATTTCTACAGGAAGTT 474
```

RESULT 72
BP234218
LOCUS BP234218 582 bp mRNA linear EST 15-SEP-2004

```
DEFINITION  BP234218 Sugano cDNA library, coronary artery endothelial cell Homo
            sapiens cDNA clone HCR02378, mRNA sequence.
ACCESSION   BP234218
VERSION     BP234218.1  GI:52107128
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
  PUBMED    15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HCR02378"
                     /tissue_type="coronary artery"
                     /cell_type="endothelial cell"
                     /clone_lib="Sugano cDNA library, coronary artery
                     endothelial cell"
ORIGIN

Alignment Scores:
Pred. No.:                27.6          Length:          582
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BP234218 (1-582)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        460 CCAGATGTGATATTTCTACAGGAAGTT 486

RESULT 73
BP261141
LOCUS       BP261141                582 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP261141 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI02606, mRNA sequence.
ACCESSION   BP261141
VERSION     BP261141.1  GI:52176372
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
  PUBMED    15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI02606"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                27.6          Length:          582
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BP261141 (1-582)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        339 CCAGATGTGATATTTCTACAGGAAGTT 365

RESULT 74
BP261789
LOCUS       BP261789                582 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP261789 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI05013, mRNA sequence.
ACCESSION   BP261789
VERSION     BP261789.1  GI:52177020
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
  PUBMED    15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI05013"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                27.6          Length:          582
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BP261789 (1-582)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         61 CCAGATGTGATATTTCTACAGGAAGTT 87
```

```
RESULT 75
BP263516
LOCUS       BP263516                582 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP263516 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI11929, mRNA sequence.
ACCESSION   BP263516
VERSION     BP263516.1  GI:52178747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI11929"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:              27.6          Length:         582
Score:                  45.00         Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BP263516 (1-582)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        399 CCAGATGTGATATTTCTACAGGAAGTT 425


Search completed: March 11, 2006, 03:39:39
Job time : 245.043 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 15:40:34 ; Search time 2330.09 Seconds
(without alignments)
1746.920 Million cell updates/sec

Title: US-10-757-745-2_COPY_54_140
Perfect score: 448
Sequence: 1 MERALNSYFEPPVEESALER.....IDGLDLNNLSERARGVCSYL 87

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170329_4058/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_12397_@runat_09032006_170329_4058 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	448	100.0	429	5	BQ582059	BQ582059 il12c11.y
2	448	100.0	471	7	CR543841	CR543841 DKFZp459O
3	448	100.0	473	6	CB120234	CB120234 K-EST0167
4	448	100.0	480	1	AI750554	AI750554 cn04a01.y
5	448	100.0	485	1	AA486032	AA486032 ab40b10.r
6	448	100.0	530	3	BM846221	BM846221 K-EST0125
7	448	100.0	554	3	BP220985	BP220985 BP220985
8	448	100.0	565	3	BP226620	BP226620 BP226620
9	448	100.0	566	3	BP220508	BP220508 BP220508
10	448	100.0	568	3	BP257211	BP257211 BP257211
11	448	100.0	570	3	BP221518	BP221518 BP221518
12	448	100.0	573	1	AU279894	AU279894 AU279894
13	448	100.0	573	3	BP221260	BP221260 BP221260
14	448	100.0	573	3	BP261500	BP261500 BP261500
15	448	100.0	576	3	BP262642	BP262642 BP262642
16	448	100.0	577	7	CN298919	CN298919 170006001
17	448	100.0	580	3	BP274368	BP274368 BP274368
18	448	100.0	581	3	BP207751	BP207751 BP207751
19	448	100.0	581	3	BP225088	BP225088 BP225088
20	448	100.0	581	3	BP262638	BP262638 BP262638
21	448	100.0	581	3	BP270744	BP270744 BP270744
22	448	100.0	581	3	BP379659	BP379659 BP379659
23	448	100.0	582	3	BP219740	BP219740 BP219740
24	448	100.0	582	3	BP261141	BP261141 BP261141
25	448	100.0	582	3	BP270415	BP270415 BP270415
26	448	100.0	582	3	BP275810	BP275810 BP275810
27	448	100.0	583	3	BP195721	BP195721 BP195721
28	448	100.0	583	3	BP262103	BP262103 BP262103
29	448	100.0	583	3	BP262741	BP262741 BP262741
30	448	100.0	583	3	BP319548	BP319548 BP319548
31	448	100.0	583	5	BU783229	BU783229 in01a08.y
32	448	100.0	584	3	BP263059	BP263059 BP263059
33	448	100.0	584	3	BP348623	BP348623 BP348623
34	448	100.0	588	3	BP236074	BP236074 BP236074
35	448	100.0	602	1	AL703449	AL703449 DKFZp686G
36	448	100.0	621	6	CB069952	CB069952 is31e10.y
37	448	100.0	634	6	CB157906	CB157906 K-EST0217
38	448	100.0	634	7	CV023369	CV023369 288 Full
39	448	100.0	644	3	BI333830	BI333830 602999339
40	448	100.0	652	7	CR753214	CR753214 DKFZp469I
41	448	100.0	657	7	CN298920	CN298920 170006001
42	448	100.0	678	2	BG719113	BG719113 602699045
43	448	100.0	701	7	CR767236	CR767236 DKFZp469E
44	448	100.0	710	5	BU661472	BU661472 cl72d11.z
45	448	100.0	711	7	CN298923	CN298923 170006000
46	448	100.0	729	7	CN298922	CN298922 170006000
47	448	100.0	729	7	CN298924	CN298924 170004244
48	448	100.0	731	8	CX760857	CX760857 AGENCOURT
49	448	100.0	767	2	BI258848	BI258848 602969663
50	448	100.0	774	3	BI754101	BI754101 603027659
51	448	100.0	786	2	BG719977	BG719977 602691335
52	448	100.0	799	8	CX756424	CX756424 AGENCOURT
53	448	100.0	810	8	CX166335	CX166335 HESC2_39_
54	448	100.0	838	3	BI908925	BI908925 603067028
55	448	100.0	846	7	CR765451	CR765451 DKFZp469E
56	448	100.0	853	2	BE784416	BE784416 601473891
57	448	100.0	858	3	BI553412	BI553412 603193358
58	448	100.0	870	2	BG740396	BG740396 602634171
59	448	100.0	870	5	BU159911	BU159911 AGENCOURT
60	448	100.0	883	5	BU169945	BU169945 AGENCOURT
61	448	100.0	883	5	BU179107	BU179107 AGENCOURT
62	448	100.0	906	6	CD251503	CD251503 AGENCOURT
63	448	100.0	937	2	BG541819	BG541819 602569712
64	448	100.0	948	1	AL555333	AL555333 AL555333
65	448	100.0	959	2	BI161201	BI161201 602865659
66	448	100.0	972	3	BM468826	BM468826 AGENCOURT
67	448	100.0	981	5	BX338160	BX338160 BX338160
68	448	100.0	995	2	BE892886	BE892886 601435730
69	448	100.0	1018	5	BX337141	BX337141 BX337141
70	448	100.0	1035	3	BM554324	BM554324 AGENCOURT
71	448	100.0	1067	5	BX337905	BX337905 BX337905
72	448	100.0	1081	5	BX444691	BX444691 BX444691
73	448	100.0	1089	11	DQ049205	DQ049205 Homo sapi
74	448	100.0	1103	3	BM555041	BM555041 AGENCOURT
75	448	100.0	1168	4	CR601303	CR601303 full-leng

ALIGNMENTS

RESULT 1

```
BQ582059
LOCUS       BQ582059                 429 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  il12c11.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6029589
            5' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION   BQ582059
VERSION     BQ582059.1  GI:21494955
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: il12c11.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6029589"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:                5.92e-47        Length:          429
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BQ582059 (1-429)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 162

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        163 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 222

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        223 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 282

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        283 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 342

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        343 CGAGGGGTGTGTTCCTACTTA 363

RESULT 2
CR543841
LOCUS       CR543841                 471 bp    mRNA    linear   EST 07-JUL-2004
DEFINITION  DKFZp459O0742_r1 459 (synonym: pcor1) Pongo pygmaeus cDNA clone
            DKFZp459O0742 5', mRNA sequence.
ACCESSION   CR543841
VERSION     CR543841.1  GI:49895253
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 471)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C.,
            Osanger,A., Fobo,G., Han,M. and Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Bloecker,H., Boecher,M., Brandt,P., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National
            Research Centre for Biotechnology Ltd., Braunschweig/Germany)
            within the cDNA sequencing consortium of the German Genome Project.
            This clone (DKFZp459O0742) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .471
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp459O0742"
                     /tissue_type="cortex"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="459 (synonym: pcor1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:                6.69e-47        Length:          471
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x CR543841 (1-471)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 220

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 CGCCCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 280

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 340

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 400

Qy         81 ArgGlyValCysSerTyrLeu 87
                 |||||||||||||||||||
Db        401 CGAGGGGTGTGTTCCTACTTA 421

RESULT 3
CB120234
LOCUS       CB120234                473 bp    mRNA    linear   EST 28-JAN-2003
DEFINITION  K-EST0167337 L8SCK0 Homo sapiens cDNA clone L8SCK0-29-F06 5', mRNA
            sequence.
ACCESSION   CB120234
VERSION     CB120234.1  GI:27946036
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 473)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 29  row: F  column: 06
            High quality sequence stop: 473.
FEATURES             Location/Qualifiers
     source          1. .473
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L8SCK0-29-F06"
                     /sex="M"
                     /cell_line="SCK"
                     /lab_host="Top10F'"
                     /clone_lib="L8SCK0"
                     /note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI;
                     Site_2: NotI; The library was contributed by the Soares
                     laboratory and it was constructed as described by Bonaldo,
                     M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
                     6(9): 791-806. RNA was prepared from harvested cell
                     culture."
ORIGIN

Alignment Scores:
Pred. No.:              6.73e-47        Length:         473
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x CB120234 (1-473)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 168

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 228

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 288

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 348

Qy         81 ArgGlyValCysSerTyrLeu 87
                |||||||||||||||||||
Db        349 CGAGGGGTGTGTTCCTACTTA 369

RESULT 4
AI750554
LOCUS       AI750554                480 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  cn04a01.y1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
            clone NHTBC_cn04a01 random, mRNA sequence.
ACCESSION   AI750554
VERSION     AI750554.1  GI:5128818
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
            Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
            Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
  TITLE     SGAP: The Skeletal Genome Anatomy Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Libin Jia
            Medical Genetics Branch
            National Human Genome Research Institute
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libin@helix.nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 04  row: a  column: 01
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NHTBC_cn04a01"
                     /sex="Female"
                     /tissue_type="Bone"
                     /cell_type="Trabecular Bone Cells"
                     /lab_host="SURE"
                     /clone_lib="Normal Human Trabecular Bone Cells"
                     /note="Organ: Hip; Vector: pBluescript; Site_1: EcoRI;
                     Library constructed by Dr. Marian Young and Dr. Pamela
                     Gehron Robey (NIDCR)"
ORIGIN

Alignment Scores:
Pred. No.:              6.86e-47        Length:         480
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AI750554 (1-480)
```

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 213

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 273

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 333

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 393

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        394 CGAGGGGTGTGTTCCTACTTA 414
```

RESULT 5
AA486032
LOCUS AA486032 485 bp mRNA linear EST 06-MAR-1998
DEFINITION ab40b10.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone IMAGE:843259 5', mRNA sequence.
ACCESSION AA486032
VERSION AA486032.1 GI:2216248
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 485)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1913 Std Error: 0.00
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 444.

FEATURES Location/Qualifiers
source 1. .485
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:843259"
/sex="female"
/dev_stage="HeLa S3 cell line"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene HeLa cell s3 937216"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. HeLa S3 epithelioid carcinoma cells grown to semi-confluency without induction. Average insert size: 1.5 kb; Uni-ZAP XR Vector. ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"

ORIGIN

Alignment Scores:
Pred. No.: 6.95e-47 Length: 485
Score: 448.00 Matches: 87

Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x AA486032 (1-485)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 209

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 269

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 329

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 389

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        390 CGAGGGGTGTGTTCCTACTTA 410
```

RESULT 6
BM846221
LOCUS BM846221 530 bp mRNA linear EST 06-MAR-2002
DEFINITION K-EST0125204 S13KMS5 Homo sapiens cDNA clone S13KMS5-59-C04 5', mRNA sequence.
ACCESSION BM846221
VERSION BM846221.1 GI:19202620
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 530)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 59 row: C column: 04
High quality sequence stop: 530.

FEATURES Location/Qualifiers
source 1. .530
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S13KMS5-59-C04"
/tissue_type="myeloma"
/cell_line="KMS-5"
/lab_host="Top10F'"
/clone_lib="S13KMS5"
/note="Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to

have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Alignment Scores:

Pred. No.:	7.81e-47	Length:	530
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x BM846221 (1-530)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 223

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 283

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 343

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 403

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        404 CGAGGGGTGTGTTCCTACTTA 424
```

RESULT 7
BP220985
LOCUS BP220985 554 bp mRNA linear EST 15-SEP-2004
DEFINITION BP220985 Sugano cDNA library, colon Homo sapiens cDNA clone COL06450, mRNA sequence.
ACCESSION BP220985
VERSION BP220985.1 GI:52093890
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 554)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .554
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL06450"
/tissue_type="colon"
/clone_lib="Sugano cDNA library, colon"

ORIGIN

Alignment Scores:

Pred. No.:	8.28e-47	Length:	554
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x BP220985 (1-554)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 128

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 188

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 248

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 308

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        309 CGAGGGGTGTGTTCCTACTTA 329
```

RESULT 8
BP226620
LOCUS BP226620 565 bp mRNA linear EST 15-SEP-2004
DEFINITION BP226620 Sugano cDNA library, dermoid cancer Homo sapiens cDNA clone DMC03190, mRNA sequence.
ACCESSION BP226620
VERSION BP226620.1 GI:52099525
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 565)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .565
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DMC03190"
/clone_lib="Sugano cDNA library, dermoid cancer"
/note="dermoid cancer"

ORIGIN

Alignment Scores:

Pred. No.:	8.49e-47	Length:	565
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0

```
DB:                    3              Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x BP226620 (1-565)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 220

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 280

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 340

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 400

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        401 CGAGGGGTGTGTTCCTACTTA 421
```

```
RESULT 9
BP220508
LOCUS       BP220508                 566 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP220508 Sugano cDNA library, colon Homo sapiens cDNA clone
            COL03160, mRNA sequence.
ACCESSION   BP220508
VERSION     BP220508.1  GI:52093413
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .566
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL03160"
                     /tissue_type="colon"
                     /clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:
Pred. No.:                 8.51e-47        Length:                 566
Score:                     448.00          Matches:                87
Percent Similarity:        100.0%          Conservative:           0
Best Local Similarity:     100.0%          Mismatches:             0
Query Match:               100.0%          Indels:                 0
DB:                        3               Gaps:                   0

US-10-757-745-2_COPY_54_140 (1-87) x BP220508 (1-566)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 138

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 198

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 258

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 318

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        319 CGAGGGGTGTGTTCCTACTTA 339
```

```
RESULT 10
BP257211
LOCUS       BP257211                 568 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP257211 Sugano cDNA library, heart Homo sapiens cDNA clone
            HRT00430, mRNA sequence.
ACCESSION   BP257211
VERSION     BP257211.1  GI:52172441
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .568
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRT00430"
                     /tissue_type="heart"
                     /clone_lib="Sugano cDNA library, heart"
ORIGIN

Alignment Scores:
Pred. No.:                 8.55e-47        Length:                 568
Score:                     448.00          Matches:                87
Percent Similarity:        100.0%          Conservative:           0
Best Local Similarity:     100.0%          Mismatches:             0
Query Match:               100.0%          Indels:                 0
DB:                        3               Gaps:                   0

US-10-757-745-2_COPY_54_140 (1-87) x BP257211 (1-568)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 189

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 249

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 309
```

```
Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 369

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        370 CGAGGGGTGTGTTCCTACTTA 390

RESULT 11
BP221518
LOCUS       BP221518                 570 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP221518 Sugano cDNA library, colon Homo sapiens cDNA clone
            COL09203, mRNA sequence.
ACCESSION   BP221518
VERSION     BP221518.1  GI:52094423
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL09203"
                     /tissue_type="colon"
                     /clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:
Pred. No.:              8.59e-47        Length:         570
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP221518 (1-570)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 127

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 187

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 247

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 307

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        308 CGAGGGGTGTGTTCCTACTTA 328
```

```
RESULT 12
AU279894
LOCUS       AU279894                 573 bp    mRNA    linear   EST 31-JUL-2003
DEFINITION  AU279894 CHONS2 Homo sapiens cDNA clone CHONS2002038 5', mRNA
            sequence.
ACCESSION   AU279894
VERSION     AU279894.1  GI:28299121
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Imabayashi,H., Mori,T., Gojo,S., Kiyono,T., Sugiyama,T., Irie,R.,
            Isogai,T., Hata,J., Tomoya,Y. and Umezawa,A.
  TITLE     Redifferentiation of dedifferentiated chondrocytes and
            chondrogenesis of human bone marrow stromal cells via chondrosphere
            formation with expression profiling by large-scale cDNA analysis
  JOURNAL   Exp. Cell Res. 288 (1), 35-50 (2003)
   PUBMED   12878157
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA Project, Sugiyama,T.; Wakamatsu,A.; Irie,R.;
            Umezawa,A.; Fukuma,M.; Kusakari,S.; Hata,J.; Ishii,S.; Yamamoto,J.;
            Isono,Y.; Saito,K.; Nakamura,Y.; Masuho,Y.; Nagai,K.; Isogai,T.
            HRI human cDNA project; cDNA library construction & 5'-end one
            pass sequencing: Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CHONS2002038"
                     /cell_type="chondrocytes"
                     /clone_lib="CHONS2"
                     /note="Vector: pME18SFL3"
ORIGIN

Alignment Scores:
Pred. No.:              8.65e-47        Length:         573
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AU279894 (1-573)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 256

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 316

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 376

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 436

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        437 CGAGGGGTGTGTTCCTACTTA 457
```

```
RESULT 13
BP221260
LOCUS       BP221260                 573 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP221260 Sugano cDNA library, colon Homo sapiens cDNA clone
            COL08083, mRNA sequence.
ACCESSION   BP221260
VERSION     BP221260.1  GI:52094165
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL08083"
                     /tissue_type="colon"
                     /clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:
Pred. No.:                8.65e-47        Length:         573
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP221260 (1-573)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 141

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 201

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 261

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 321

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        322 CGAGGGGTGTGTTCCTACTTA 342

RESULT 14
BP261500
LOCUS       BP261500                 573 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP261500 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI03816, mRNA sequence.
ACCESSION   BP261500
VERSION     BP261500.1  GI:52176731
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI03816"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                8.65e-47        Length:         573
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP261500 (1-573)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 125

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 185

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 245

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 305

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        306 CGAGGGGTGTGTTCCTACTTA 326

RESULT 15
BP262642
LOCUS       BP262642                 576 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262642 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI07584, mRNA sequence.
ACCESSION   BP262642
VERSION     BP262642.1  GI:52177873
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 576)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
```

```
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .576
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI07584"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                 8.71e-47        Length:          576
Score:                     448.00          Matches:         87
Percent Similarity:        100.0%          Conservative:    0
Best Local Similarity:     100.0%          Mismatches:      0
Query Match:               100.0%          Indels:          0
DB:                        3               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BP262642 (1-576)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 143

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 203

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 263

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 323

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        324 CGAGGGGTGTGTTCCTACTTA 344

RESULT 16
CN298919
LOCUS       CN298919                 577 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600175228 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN298919
VERSION     CN298919.1  GI:47315333
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 577   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

Alignment Scores:
Pred. No.:                 8.73e-47        Length:          577
Score:                     448.00          Matches:         87
Percent Similarity:        100.0%          Conservative:    0
Best Local Similarity:     100.0%          Mismatches:      0
Query Match:               100.0%          Indels:          0
DB:                        7               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x CN298919 (1-577)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGTTCCTACTTA 420

RESULT 17
BP274368
LOCUS       BP274368                 580 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP274368 Sugano cDNA library, kidney Homo sapiens cDNA clone
            KDN00894, mRNA sequence.
ACCESSION   BP274368
VERSION     BP274368.1  GI:52188100
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
```

```
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KDN00894"
                     /tissue_type="kidney"
                     /clone_lib="Sugano cDNA library, kidney"
ORIGIN

Alignment Scores:
Pred. No.:                8.79e-47       Length:          580
Score:                    448.00         Matches:         87
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       3              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BP274368 (1-580)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 239

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 299

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 359

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 419

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        420 CGAGGGGTGTGTTCCTACTTA 440

RESULT 18
BP207751
LOCUS       BP207751                  581 bp    mRNA    linear   EST 14-SEP-2004
DEFINITION  BP207751 Sugano cDNA library, coronary artery smooth muscle cell
            Homo sapiens cDNA clone CAS11735, mRNA sequence.
ACCESSION   BP207751
VERSION     BP207751.1  GI:52064157
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CAS11735"
                     /tissue_type="coronary artery"
                     /cell_type="smooth muscle cell"
                     /clone_lib="Sugano cDNA library, coronary artery smooth
                     muscle cell"
ORIGIN

Alignment Scores:
Pred. No.:                8.81e-47       Length:          581
Score:                    448.00         Matches:         87
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       3              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BP207751 (1-581)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 127

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 187

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 247

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 307

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        308 CGAGGGGTGTGTTCCTACTTA 328

RESULT 19
BP225088
LOCUS       BP225088                  581 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP225088 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA
            clone DAT05184, mRNA sequence.
ACCESSION   BP225088
VERSION     BP225088.1  GI:52097993
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DAT05184"
                     /cell_type="lymphocyte"
                     /cell_line="Daudi"
                     /clone_lib="Sugano cDNA library, lymphocyte Daudi"
                     /note="Burkitt's lymphoma"
ORIGIN

Alignment Scores:
Pred. No.:                8.81e-47       Length:          581
```

```
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP225088 (1-581)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 240

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 300

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 360

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 420

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        421 CGAGGGGTGTGTTCCTACTTA 441

RESULT 20
BP262638
LOCUS       BP262638                 581 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262638 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI07576, mRNA sequence.
ACCESSION   BP262638
VERSION     BP262638.1  GI:52177869
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI07576"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:              8.81e-47        Length:         581
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP262638 (1-581)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 136

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 196

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 256

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 316

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        317 CGAGGGGTGTGTTCCTACTTA 337

RESULT 21
BP270744
LOCUS       BP270744                 581 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP270744 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone KAR05146, mRNA sequence.
ACCESSION   BP270744
VERSION     BP270744.1  GI:52220093
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAR05146"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:              8.81e-47        Length:         581
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP270744 (1-581)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 127

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 187
```

```
Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 247

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 307

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        308 CGAGGGGTGTGTTCCTACTTA 328

RESULT 22
BP379659
LOCUS       BP379659                581 bp    mRNA    linear   EST 21-SEP-2004
DEFINITION  BP379659 Sugano cDNA library, uterus Homo sapiens cDNA clone
            WMD06642, mRNA sequence.
ACCESSION   BP379659
VERSION     BP379659.1  GI:52412762
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
  PUBMED    15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="WMD06642"
                     /tissue_type="uterus"
                     /clone_lib="Sugano cDNA library, uterus"
ORIGIN

Alignment Scores:
Pred. No.:                8.81e-47        Length:          581
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BP379659 (1-581)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 128

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 188

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 248

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 308

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        309 CGAGGGGTGTGTTCCTACTTA 329

RESULT 23
BP219740
LOCUS       BP219740                582 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP219740 Sugano cDNA library, caudate nucleus Homo sapiens cDNA
            clone CNR08985, mRNA sequence.
ACCESSION   BP219740
VERSION     BP219740.1  GI:52092643
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
  PUBMED    15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CNR08985"
                     /tissue_type="caudate nucleus"
                     /clone_lib="Sugano cDNA library, caudate nucleus"
ORIGIN

Alignment Scores:
Pred. No.:                8.83e-47        Length:          582
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BP219740 (1-582)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 234

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 294

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 354

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 414

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        415 CGAGGGGTGTGTTCCTACTTA 435

RESULT 24
BP261141
LOCUS       BP261141                582 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP261141 Sugano cDNA library, small intestine Homo sapiens cDNA
```

```
            clone HSI02606, mRNA sequence.
ACCESSION   BP261141
VERSION     BP261141.1  GI:52176372
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI02606"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:               8.83e-47       Length:          582
Score:                   448.00         Matches:         87
Percent Similarity:      100.0%         Conservative:    0
Best Local Similarity:   100.0%         Mismatches:      0
Query Match:             100.0%         Indels:          0
DB:                      3              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BP261141 (1-582)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 125

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 185

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 245

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 305

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        306 CGAGGGGTGTGTTCCTACTTA 326

RESULT 25
BP270415
LOCUS       BP270415                582 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP270415 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone KAR02166, mRNA sequence.
ACCESSION   BP270415
VERSION     BP270415.1  GI:52219764
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAR02166"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:               8.83e-47       Length:          582
Score:                   448.00         Matches:         87
Percent Similarity:      100.0%         Conservative:    0
Best Local Similarity:   100.0%         Mismatches:      0
Query Match:             100.0%         Indels:          0
DB:                      3              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BP270415 (1-582)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 217

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 277

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 337

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 397

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        398 CGAGGGGTGTGTTCCTACTTA 418

RESULT 26
BP275810
LOCUS       BP275810                582 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP275810 Sugano cDNA library, kidney Homo sapiens cDNA clone
            KDN04583, mRNA sequence.
ACCESSION   BP275810
VERSION     BP275810.1  GI:52189542
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
```

```
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KDN04583"
                     /tissue_type="kidney"
                     /clone_lib="Sugano cDNA library, kidney"
ORIGIN
```

Alignment Scores:

Pred. No.:	8.83e-47	Length:	582
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x BP275810 (1-582)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 208

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 268

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 328

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 388

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        389 CGAGGGGTGTGTTCCTACTTA 409
```

```
RESULT 27
BP195721
LOCUS       BP195721                583 bp    mRNA    linear   EST 14-SEP-2004
DEFINITION  BP195721 Sugano cDNA library, brain Homo sapiens cDNA clone
            ADB08961, mRNA sequence.
ACCESSION   BP195721
VERSION     BP195721.1  GI:52039998
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="ADB08961"
                     /tissue_type="brain"
                     /clone_lib="Sugano cDNA library, brain"
ORIGIN
```

Alignment Scores:

Pred. No.:	8.85e-47	Length:	583
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x BP195721 (1-583)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 276

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 336

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 396

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 456

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        457 CGAGGGGTGTGTTCCTACTTA 477
```

```
RESULT 28
BP262103
LOCUS       BP262103                583 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262103 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI05902, mRNA sequence.
ACCESSION   BP262103
VERSION     BP262103.1  GI:52177334
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI05902"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN
```

Alignment Scores:

```
Pred. No.:              8.85e-47        Length:         583
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP262103 (1-583)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 153

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 213

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 273

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 333

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        334 CGAGGGGTGTGTTCCTACTTA 354

RESULT 29
BP262741
LOCUS       BP262741                583 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262741 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI07919, mRNA sequence.
ACCESSION   BP262741
VERSION     BP262741.1  GI:52177972
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI07919"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:              8.85e-47        Length:         583
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP262741 (1-583)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 89

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 149

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 209

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 269

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        270 CGAGGGGTGTGTTCCTACTTA 290

RESULT 30
BP319548
LOCUS       BP319548                583 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP319548 Sugano cDNA library, pericardium Homo sapiens cDNA clone
            PCD11128, mRNA sequence.
ACCESSION   BP319548
VERSION     BP319548.1  GI:52248523
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PCD11128"
                     /tissue_type="pericardium"
                     /clone_lib="Sugano cDNA library, pericardium"
ORIGIN

Alignment Scores:
Pred. No.:              8.85e-47        Length:         583
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP319548 (1-583)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 268

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 328
```

```
Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         329 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 388

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         389 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 448

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         449 CGAGGGGTGTGTTCCTACTTA 469

RESULT 31
BU783229
LOCUS       BU783229                583 bp    mRNA    linear   EST 11-OCT-2002
DEFINITION  in01a08.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123158
            5' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION   BU783229
VERSION     BU783229.1  GI:23827207
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: in01a08.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 430.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6123158"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:

Pred. No.:                8.85e-47        Length:          583
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BU783229 (1-583)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         158 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 217

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         218 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 277

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         278 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 337

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         338 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 397

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         398 CGAGGGGTGTGTTCCTACTTA 418

RESULT 32
BP263059
LOCUS       BP263059                584 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP263059 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI10591, mRNA sequence.
ACCESSION   BP263059
VERSION     BP263059.1  GI:52178290
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
  PUBMED    15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .584
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI10591"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                8.87e-47        Length:          584
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BP263059 (1-584)
```

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 142

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 202

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 262

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 322

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        323 CGAGGGGTGTGTTCCTACTTA 343

RESULT 33
BP348623
LOCUS       BP348623                584 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP348623 Sugano cDNA library, brain Homo sapiens cDNA clone
            SZR01143, mRNA sequence.
ACCESSION   BP348623
VERSION     BP348623.1  GI:52278608
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .584
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SZR01143"
                     /tissue_type="brain"
                     /clone_lib="Sugano cDNA library, brain"
ORIGIN

Alignment Scores:
Pred. No.:              8.87e-47        Length:         584
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP348623 (1-584)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 249

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 309

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 369

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 429

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        430 CGAGGGGTGTGTTCCTACTTA 450

RESULT 34
BP236074
LOCUS       BP236074                588 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP236074 Sugano cDNA library, coronary artery endothelial cell Homo
            sapiens cDNA clone HCR06784, mRNA sequence.
ACCESSION   BP236074
VERSION     BP236074.1  GI:52108984
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 588)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .588
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HCR06784"
                     /tissue_type="coronary artery"
                     /cell_type="endothelial cell"
                     /clone_lib="Sugano cDNA library, coronary artery
                     endothelial cell"
ORIGIN

Alignment Scores:
Pred. No.:              8.95e-47        Length:         588
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BP236074 (1-588)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 255

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 315

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 375

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 435

Qy         81 ArgGlyValCysSerTyrLeu 87
                ||||||||||||||||||||
Db        436 CGAGGGGTGTGTTCCTACTTA 456

RESULT 35
AL703449
LOCUS       AL703449                602 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686G0621_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686G0621 5', mRNA sequence.
ACCESSION   AL703449
VERSION     AL703449.1  GI:19686804
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   Ansorge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Ansorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686G0621) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .602
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686G0621"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

Alignment Scores:
Pred. No.:              9.23e-47        Length:         602
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AL703449 (1-602)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 220

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 280

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 340

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 400

Qy         81 ArgGlyValCysSerTyrLeu 87
                ||||||||||||||||||||
Db        401 CGAGGGGTGTGTTCCTACTTA 421

RESULT 36
CB069952
LOCUS       CB069952                621 bp    mRNA    linear   EST 21-JAN-2003
DEFINITION  is31e10.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6554035 5'
            similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION   CB069952
VERSION     CB069952.1  GI:27814472
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: is31e10.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40UP from Gibco
            High quality sequence stop: 476.
FEATURES             Location/Qualifiers
     source          1. .621
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6554035"
                     /tissue_type="Purified pancreatic islet"
                     /lab_host="DH10B"
                     /clone_lib="HR85 islet"
                     /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                     NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                     Size-selected on agarose gel. Average insert size ~1kb. 5'
                     XhoI site was destroyed after directional cloning.
                     Amplified once. Contact information: Hiroshi Inoue, MD,
                     Metabolism Div. (Alan Permutt Lab), Washington University
                     School of Medicine, Box 8127, 660 South Euclid Ave., St.
                     Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                     314-362-1916, Fax: 314-747-2692."
ORIGIN

Alignment Scores:
Pred. No.:              9.61e-47        Length:         621
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x CB069952 (1-621)
```

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 225

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 285

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 345

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 405

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        406 CGAGGGGTGTGTTCCTACTTA 426
```

RESULT 37
CB157906

```
LOCUS       CB157906                634 bp    mRNA    linear   EST 29-JAN-2003
DEFINITION  K-EST0217150 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-5-F06
            5', mRNA sequence.
ACCESSION   CB157906
VERSION     CB157906.1  GI:28143040
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 5  row: F  column: 06
            High quality sequence stop: 634.
FEATURES             Location/Qualifiers
     source          1. .634
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L18POOL1n1-5-F06"
                     /cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
                     /lab_host="Top10F'"
                     /clone_lib="L18POOL1n1"
                     /note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI;
                     Site_2: NotI; The library was contributed by the Soares
                     laboratory and it was constructed as described by Bonaldo,
                     M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
                     6(9): 791-806. RNA was prepared from harvested cell
                     culture."
ORIGIN
```

Alignment Scores:

Pred. No.:	9.88e-47	Length:	634
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x CB157906 (1-634)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 171

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 231

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 291

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 351

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        352 CGAGGGGTGTGTTCCTACTTA 372
```

RESULT 38
CV023369

```
LOCUS       CV023369                634 bp    mRNA    linear   EST 20-AUG-2004
DEFINITION  288 Full Length cDNA from the Mammalian Gene Collection Homo
            sapiens cDNA 5' similar to BC017553, mRNA sequence.
ACCESSION   CV023369
VERSION     CV023369.1  GI:51481130
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
            Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
            Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
            Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
            Vandenhaute,J., Cusick,M.E., Albala,J.S., Hill,D.E. and Vidal,M.
  TITLE     Human ORFeome Version 1.1: a Platform for Reverse Proteomics
  JOURNAL   Genome Res. (2004) In press
COMMENT     Contact: Vidal M
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 5739
            Email: Marc_Vidal@dfci.harvard.edu
            ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
            results from a PCR reaction using an MGC full-length cDNA as
            template DNA and ORF specific primers
            PCR PRimers
            FORWARD: ATGGAGTTGGGGAGTTGCCTGGA
            BACKWARD: TACAATATTATATCTAAGTTGCACAGAAG
            Insert Length: 634    Std Error: 32.00
            Plate: 11008  row: 05  column: F
            Seq primer: ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
            High quality sequence start: 97
            High quality sequence stop: 633
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .634
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="mixed"
                     /clone_lib="Full Length cDNA from the Mammalian Gene
                     Collection"
                     /note="Vector: mixed; The ORFs were PCR amplified from the
                     MGC (Mammalian Gene Collection) as of April 2004 and
```

cloned by recombinational Gateway cloning into pDONR223 Donor vector. Reference : MGC (Mammalian Gene Collection) Program Team, Generation and Initial Analysis of more than 15,000 Full-Length Human and Mouse cDNA Sequences. PNAS, 2002, 99(26), 16899-16903"

ORIGIN

Alignment Scores:

Pred. No.:	9.88e-47	Length:	634
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x CV023369 (1-634)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGTTCCTACTTA 420
```

RESULT 39
BI333830

LOCUS	BI333830	644 bp mRNA linear	EST 30-JUL-2001
DEFINITION	602999339F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141577 5', mRNA sequence.		
ACCESSION	BI333830		
VERSION	BI333830.1 GI:15018487		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 644)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Incyte Genomics, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11348 row: n column: 10 High quality sequence stop: 531.		

FEATURES	Location/Qualifiers
source	1. .644 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5141577" /tissue_type="cervical carcinoma cell line" /lab_host="DH10B" /clone_lib="NIH_MGC_12" /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

ORIGIN

Alignment Scores:

Pred. No.:	1.01e-46	Length:	644
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x BI333830 (1-644)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 206

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 266

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 326

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 386

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        387 CGAGGGGTGTGTTCCTACTTA 407
```

RESULT 40
CR753214

LOCUS	CR753214	652 bp mRNA linear	EST 01-SEP-2004
DEFINITION	DKFZp469I195_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone DKFZp469I195 5', mRNA sequence.		
ACCESSION	CR753214		
VERSION	CR753214.1 GI:51845629		
KEYWORDS	EST.		
SOURCE	Pongo pygmaeus (orangutan)		
ORGANISM	Pongo pygmaeus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo.		
REFERENCE	1 (bases 1 to 652)		
AUTHORS	Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.		
TITLE	Pongo pygmaeus mRNA (Ansorge,W., Krieger,S., Regiert,T., et al.)		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: MIPS MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany This is the 5' sequence of the clone insert. Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; lin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469I195 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.		

FEATURES	Location/Qualifiers
source	1. .652 /organism="Pongo pygmaeus" /mol_type="mRNA" /db_xref="taxon:9600" /clone="DKFZp469I195" /tissue_type="kidney"

```
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="469 (synonym: pkid1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:                1.02e-46       Length:          652
Score:                    448.00         Matches:         87
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       7              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x CR753214 (1-652)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 236

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 CGCCCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 296

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 356

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 416

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        417 CGAGGGGTGTGTTCCTACTTA 437

RESULT 41
CN298920
LOCUS       CN298920                657 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600170525 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN298920
VERSION     CN298920.1  GI:47315334
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 657   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cell, retinoic acid and
                     mitogen-treated hES cell line H7"
                     /clone_lib="GRN_PRENEU"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from hES cell line H7 (p29) maintained in feeder-free
                     conditions. Embryoid bodies were generated in the presence
                     of all-trans retinoic and mitogens."
ORIGIN

Alignment Scores:
Pred. No.:                1.03e-46       Length:          657
Score:                    448.00         Matches:         87
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       7              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x CN298920 (1-657)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 175

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 235

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 295

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 355

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        356 CGAGGGGTGTGTTCCTACTTA 376

RESULT 42
BG719113
LOCUS       BG719113                678 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602699045F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4831186 5',
            mRNA sequence.
ACCESSION   BG719113
VERSION     BG719113.1  GI:13998300
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10753  row: i  column: 11
            High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4831186"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
```

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	1.08e-46	Length:	678
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x BG719113 (1-678)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 241

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 301

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 361

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 421

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        422 CGAGGGGTGTGTTCCTACTTA 442
```

RESULT 43
CR767236

LOCUS	CR767236 701 bp mRNA linear EST 23-SEP-2004
DEFINITION	DKFZp469E1236_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone DKFZp469E1236 5', mRNA sequence.
ACCESSION	CR767236
VERSION	CR767236.1 GI:52609173
KEYWORDS	EST.
SOURCE	Pongo pygmaeus (orangutan)
ORGANISM	Pongo pygmaeus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE	1 (bases 1 to 701)
AUTHORS	Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
TITLE	Pongo pygmaeus mRNA (Ansorge,W., Krieger,S., Regiert,T., et al.)
JOURNAL	Unpublished (2004)
COMMENT	Contact: MIPS MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany This is the 5' sequence of the clone insert. Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; rlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469E1236 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES	Location/Qualifiers
source	1. .701 /organism="Pongo pygmaeus" /mol_type="mRNA" /db_xref="taxon:9600" /clone="DKFZp469E1236" /tissue_type="kidney" /dev_stage="adult" /lab_host="DH10B" /clone_lib="469 (synonym: pkid1)" /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"

ORIGIN

Alignment Scores:

Pred. No.:	1.13e-46	Length:	701
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x CR767236 (1-701)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 147

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 CGCCCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 207

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 267

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 327

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        328 CGAGGGGTGTGTTCCTACTTA 348
```

RESULT 44
BU661472

LOCUS	BU661472 710 bp mRNA linear EST 30-SEP-2002
DEFINITION	cl72d11.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo sapiens cDNA clone cl72d11 5', mRNA sequence.
ACCESSION	BU661472
VERSION	BU661472.1 GI:23373654
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 710)
AUTHORS	Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
TITLE	Gene Expression in Human Erythroid Precursor Cells
JOURNAL	Unpublished (2002)
COMMENT	Contact: Jeffery L. Miller Laboratory of Chemical Biology National Institute of Diabetes and Digestive and Kidney Diseases Building 10, Room 9B17, National Institutes of Health, Bethesda, MD 20892, USA Tel: 301 402 2373 Fax: 301 435 5148 Email: jm7f@nih.gov The 'cl' library was constructed by Alexander Gubin, Ph.D. in the Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or analyses by National Institutes of Health Intramural Sequencing Center (NISC). More information available at: http://hembase.niddk.nih.gov Plate: 72 row: d column: 11 Seq primer: 5' lambda-TriplEx2 Sequencing Primer.

FEATURES	Location/Qualifiers

```
     source          1. .710
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl72d11"
                     /sex="unknown"
                     /tissue_type="blood"
                     /cell_type="Erythroid Precursor Cells"
                     /cell_line="Primary Culture of Peripheral Blood
                     Mononuclear Cells"
                     /dev_stage="Precursor erythroblasts; GPA++"
                     /lab_host="DH5alpha"
                     /clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl
                     library)"
                     /note="Organ: blood; Vector: pTriplEx2; Site_1: SfiI;
                     Site_2: SfiI; A complementary DNA (cDNA) library from
                     human erythroid precursor cells was constructed using
                     SMART PCR (polymerase chain reaction) cDNA Library
                     Construction Kit (Clontech, Palo Alto, CA) according to
                     the manufacturer's directions, but with slight
                     modifications. Briefly, reverse transcription was
                     performed in the presence of 1 umol/L peptide nucleic acid
                     (PNA) oligos
                     (N-terminal)-biotin-GTC-CAC-CCG-AAG-CTT-G-(C-terminal) and
                     (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)
                     . Synthesized cDNA was digested with SfiI and
                     size-selected on a 1% agarose gel (>800bp). Large-scale
                     sequencing of the library was performed by the NIH
                     Intramural Sequencing Center (NISC;
                     Http://www.nisc.nih.gov/)."
ORIGIN

Alignment Scores:
Pred. No.:                1.15e-46       Length:          710
Score:                    448.00         Matches:         87
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       5              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BU661472 (1-710)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 135

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 195

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 255

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 315

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        316 CGAGGGGTGTGTTCCTACTTA 336

RESULT 45
CN298923
LOCUS       CN298923                711 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600002122 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN298923
VERSION     CN298923.1  GI:47315337
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 711   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

Alignment Scores:
Pred. No.:                1.15e-46       Length:          711
Score:                    448.00         Matches:         87
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       7              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x CN298923 (1-711)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 184

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 244

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 304

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 364

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        365 CGAGGGGTGTGTTCCTACTTA 385

RESULT 46
CN298922
LOCUS       CN298922                729 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600077717 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN298922
VERSION     CN298922.1  GI:47315336
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 729)
```

```
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 729   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

Alignment Scores:
Pred. No.:                1.19e-46        Length:         729
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x CN298922 (1-729)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 211

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 271

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 331

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 391

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        392 CGAGGGGTGTGTTCCTACTTA 412

RESULT 47
CN298924
LOCUS       CN298924                729 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000424496911 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN298924
VERSION     CN298924.1  GI:47315338
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 729   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, cell lines H1, H7, and
                     H9"
                     /clone_lib="GRN_ES"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from undifferentiated hES cell lines H1 (p32), H7 (p29),
                     and H9 (p26) maintained in feeder-free conditions"
ORIGIN

Alignment Scores:
Pred. No.:                1.19e-46        Length:         729
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x CN298924 (1-729)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 153

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 213

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 273

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 333

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        334 CGAGGGGTGTGTTCCTACTTA 354

RESULT 48
CX760857
LOCUS       CX760857                731 bp    mRNA    linear   EST 24-JAN-2005
DEFINITION  AGENCOURT_40963383 NIH_MGC_281 Homo sapiens cDNA clone
            IMAGE:7783600 3', mRNA sequence.
ACCESSION   CX760857
VERSION     CX760857.1  GI:58057513
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 731)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Meri Firpo
             cDNA Library Preparation: Express Genomics
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15944  row: b  column: 14
            High quality sequence stop: 574.
FEATURES             Location/Qualifiers
     source          1. .731
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7783600"
                     /tissue_type="pluripotent cell line derived from
                     blastocyst inner cell mass"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_281"
                     /note="Organ: Blastocyst; Vector: pExpress-1; Site_1:
                     EcoRV; Site_2: NotI; RNA obtained from pluripotent cell
                     line derived from blastocyst inner cell mass (cell line
                     HSF-6, NIH Registry designation UC06. Positive for OCT4
                     expression by rtPCR, positive for SSEA-3, SSEA-4,
                     Tra-1-81, Tra-1-60 by immunofluorescence. Negative for
                     SSEA-1 by immunofluorescence Passage 62. cDNA was primed
                     using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. Size-selection >1.25
                     kb resulted in an average insert size of 2.0 kb. This
                     primary library is normalized (non-normalized primary
                     library is NIH_MGC_280) and was constructed by Express
                     Genomics (Frederick, MD). Note: this is a Mammalian Gene
                     Collection library."
ORIGIN

Alignment Scores:
Pred. No.:              1.19e-46      Length:         731
Score:                  448.00        Matches:        87
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     8             Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x CX760857 (1-731)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 223

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 283

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 343

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 403

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        404 CGAGGGGTGTGTTCCTACTTA 424

RESULT 49
BI258848

LOCUS       BI258848                767 bp    mRNA    linear   EST 17-JUL-2001
DEFINITION  602969663F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5109303 5',
            mRNA sequence.
ACCESSION   BI258848
VERSION     BI258848.1  GI:14815606
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 767)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11264  row: m  column: 16
            High quality sequence stop: 758.
FEATURES             Location/Qualifiers
     source          1. .767
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5109303"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_12"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
ORIGIN

Alignment Scores:
Pred. No.:              1.27e-46      Length:         767
Score:                  448.00        Matches:        87
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     2             Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BI258848 (1-767)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 156

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 216

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 276

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 336

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        337 CGAGGGGTGTGTTCCTACTTA 357

RESULT 50
BI754101
```

```
LOCUS       BI754101                774 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603027659F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5197838 5',
            mRNA sequence.
ACCESSION   BI754101
VERSION     BI754101.1  GI:15745679
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 774)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11495  row: f  column: 15
            High quality sequence stop: 756.
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5197838"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_114"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.28e-46        Length:         774
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BI754101 (1-774)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 278

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 338

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 398

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 458

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        459 CGAGGGGTGTGTTCCTACTTA 479

RESULT 51
BG719977
LOCUS       BG719977                786 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  602691335F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4823432 5',
            mRNA sequence.
ACCESSION   BG719977
VERSION     BG719977.1  GI:13999164
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 786)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10733  row: f  column: 09
            High quality sequence stop: 784.
FEATURES             Location/Qualifiers
     source          1. .786
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4823432"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.31e-46        Length:         786
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     2               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BG719977 (1-786)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 364

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 424

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 484
```

```
Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 TTCTCTCTCATTACCTGGAACATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 544

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        545 CGAGGGGTGTGTTCCTACTTA 565

RESULT 52
CX756424
LOCUS       CX756424                799 bp    mRNA    linear   EST 24-JAN-2005
DEFINITION  AGENCOURT_41337854 NIH_MGC_281 Homo sapiens cDNA clone
            IMAGE:7779791 3', mRNA sequence.
ACCESSION   CX756424
VERSION     CX756424.1  GI:58053080
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 799)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Meri Firpo
             cDNA Library Preparation: Express Genomics
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15934  row: c  column: 21
            High quality sequence stop: 579.
FEATURES             Location/Qualifiers
     source          1. .799
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7779791"
                     /tissue_type="pluripotent cell line derived from
                     blastocyst inner cell mass"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_281"
                     /note="Organ: Blastocyst; Vector: pExpress-1; Site_1:
                     EcoRV; Site_2: NotI; RNA obtained from pluripotent cell
                     line derived from blastocyst inner cell mass (cell line
                     HSF-6, NIH Registry designation UC06. Positive for OCT4
                     expression by rtPCR, positive for SSEA-3, SSEA-4,
                     Tra-1-81, Tra-1-60 by immunofluorescence. Negative for
                     SSEA-1 by immunofluorescence Passage 62. cDNA was primed
                     using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. Size-selection >1.25
                     kb resulted in an average insert size of 2.0 kb. This
                     primary library is normalized (non-normalized primary
                     library is NIH_MGC_280) and was constructed by Express
                     Genomics (Frederick, MD). Note: this is a Mammalian Gene
                     Collection library."
ORIGIN

Alignment Scores:
Pred. No.:              1.34e-46        Length:         799
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x CX756424 (1-799)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 236

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 296

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 356

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 416

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        417 CGAGGGGTGTGTTCCTACTTA 437

RESULT 53
CX166335
LOCUS       CX166335                810 bp    mRNA    linear   EST 23-DEC-2004
DEFINITION  HESC2_39_B09.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone
            IMAGE:7469923 5', mRNA sequence.
ACCESSION   CX166335
VERSION     CX166335.1  GI:56796415
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 810)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15771  row: d  column: 17
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 810.
FEATURES             Location/Qualifiers
     source          1. .810
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7469923"
                     /sex="male"
                     /tissue_type="human embryonic stem cells differentiated to
                     an early endodermal cell type"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_258"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early endodermal cell type. Cell line
```

id and NIH Registry designation is BGO1. Positive for GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP expression. Passage number 40. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Note: this is a Mammalian Gene Collection library."

ORIGIN

Alignment Scores:

Pred. No.:	1.36e-46	Length:	810
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x CX166335 (1-810)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 176

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 236

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 296

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 356

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        357 CGAGGGGTGTGTTCCTACTTA 377
```

RESULT 54
BI908925

```
LOCUS       BI908925                838 bp    mRNA    linear   EST 16-OCT-2001
DEFINITION  603067028F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5215926 5',
            mRNA sequence.
ACCESSION   BI908925
VERSION     BI908925.1  GI:16172029
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11542  row: h  column: 07
            High quality sequence stop: 773.
FEATURES             Location/Qualifiers
     source          1. .838
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5215926"
                     /tissue_type="leukocyte"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_118"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                     (destroyed); RNA source leukocytes from anonymous pool of
                     non-activated  adult donors.  Library is oligo-dT primed
                     and directionally cloned  (EcoRV site is destroyed upon
                     cloning).  Average insert size 1.7 kb, insert size range
                     1.2-3.3 kb.  Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking code 027. Note:
                     this is a NIH_MGC Library."
ORIGIN
```

Alignment Scores:

Pred. No.:	1.42e-46	Length:	838
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x BI908925 (1-838)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 220

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 280

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 340

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 400

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        401 CGAGGGGTGTGTTCCTACTTA 421
```

RESULT 55
CR765451

```
LOCUS       CR765451                846 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp469E2434_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469E2434 5', mRNA sequence.
ACCESSION   CR765451
VERSION     CR765451.1  GI:52605526
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,
            Deutschenbaur,S., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
```

Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix
(Martinsried/Germany) within the cDNA sequencing consortium of the
German Genome Project. This clone (DKFZp469E2434) is available at
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469E2434
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES Location/Qualifiers
source 1. .846
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469E2434"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469 (synonym: pkid1)"
/note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"

ORIGIN

Alignment Scores:

Pred. No.:	1.44e-46	Length:	846
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x CR765451 (1-846)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 122

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CGCCCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 182

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 242

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 302

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        303 CGAGGGGTGTGTTCCTACTTA 323
```

RESULT 56
BE784416
LOCUS BE784416 853 bp mRNA linear EST 20-OCT-2000
DEFINITION 601473891F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3876862 5',
mRNA sequence.
ACCESSION BE784416
VERSION BE784416.1 GI:10205614
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 853)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9638 row: e column: 23
High quality sequence stop: 660.

FEATURES Location/Qualifiers
source 1. .853
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3876862"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_68"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies. "

ORIGIN

Alignment Scores:

Pred. No.:	1.46e-46	Length:	853
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x BE784416 (1-853)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 223

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 283

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 343

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 403

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        404 CGAGGGGTGTGTTCCTACTTA 424
```

RESULT 57
BI553412
LOCUS BI553412 858 bp mRNA linear EST 05-SEP-2001
DEFINITION 603193358F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5264582 5',
mRNA sequence.
ACCESSION BI553412
VERSION BI553412.1 GI:15440724
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 858)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11667 row: c column: 15
High quality sequence stop: 829.

FEATURES Location/Qualifiers
source 1. .858
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5264582"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_95"
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.5 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	1.47e-46	Length:	858
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x BI553412 (1-858)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 239

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 299

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 359

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 419

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        420 CGAGGGGTGTGTTCCTACTTA 440
```

RESULT 58
BG740396
LOCUS BG740396 870 bp mRNA linear EST 15-MAY-2001
DEFINITION 602634171F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4779449 5', mRNA sequence.
ACCESSION BG740396
VERSION BG740396.1 GI:14051049
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 870)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10636 row: m column: 18
High quality sequence stop: 826.

FEATURES Location/Qualifiers
source 1. .870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4779449"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn3"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.:	1.49e-46	Length:	870
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x BG740396 (1-870)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 130

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 190

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 250

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 310

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        311 CGAGGGGTGTGTTCCTACTTA 331
```

RESULT 59
BU159911
LOCUS BU159911 870 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7933863 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6144208 5', mRNA sequence.
ACCESSION BU159911
VERSION BU159911.1 GI:22673821
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 870)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

```
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13468  row: n  column: 17
            High quality sequence stop: 677.
FEATURES             Location/Qualifiers
     source          1. .870
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6144208"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
ORIGIN

Alignment Scores:
Pred. No.:                1.49e-46       Length:          870
Score:                    448.00         Matches:         87
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       5              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BU159911 (1-870)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 154

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 214

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 274

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 334

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        335 CGAGGGGTGTGTTCCTACTTA 355

RESULT 60
BU169945
LOCUS       BU169945                 883 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7913097 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6024760
            5', mRNA sequence.
ACCESSION   BU169945
VERSION     BU169945.1  GI:22683929
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
```

```
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13235  row: m  column: 17
            High quality sequence stop: 413.
FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6024760"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

Alignment Scores:
Pred. No.:                1.52e-46       Length:          883
Score:                    448.00         Matches:         87
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       5              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BU169945 (1-883)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         40 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 99

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 159

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 219

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 279

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        280 CGAGGGGTGTGTTCCTACTTA 300

RESULT 61
BU179107
LOCUS       BU179107                 883 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7984768 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6174958
            5', mRNA sequence.
ACCESSION   BU179107
VERSION     BU179107.1  GI:22693091
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
```

```
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13548  row: o  column: 23
            High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6174958"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

Alignment Scores:
Pred. No.:                1.52e-46        Length:          883
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BU179107 (1-883)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 77

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 137

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 197

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 257

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        258 CGAGGGGTGTGTTCCTACTTA 278

RESULT 62
CD251503
LOCUS       CD251503                906 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  AGENCOURT_14212105 NIH_MGC_179 Homo sapiens cDNA clone
            IMAGE:30385385 5', mRNA sequence.
ACCESSION   CD251503
VERSION     CD251503.1  GI:31011969
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM452  row: f  column: 18
            High quality sequence stop: 588.
FEATURES             Location/Qualifiers
     source          1. .906
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30385385"
                     /tissue_type="Pituitary"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_179"
                     /note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV
                     (destroyed); Site_2: NotI; Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning).  Average insert size 1.1 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN

Alignment Scores:
Pred. No.:                1.58e-46        Length:          906
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x CD251503 (1-906)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 260

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 320

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 380

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 440

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        441 CGAGGGGTGTGTTCCTACTTA 461

RESULT 63
BG541819
LOCUS       BG541819                937 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602569712F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4694066 5',
            mRNA sequence.
ACCESSION   BG541819
VERSION     BG541819.1  GI:13534052
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 937)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1517  row: h  column: 03
            High quality sequence stop: 714.
FEATURES             Location/Qualifiers
     source          1. .937
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4694066"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_77"
                     /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.9
                     kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                1.65e-46        Length:          937
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       2               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BG541819 (1-937)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 235

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 295

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 355

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 415

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        416 CGAGGGGTGTGTTCCTACTTA 436

RESULT 64
AL555333
LOCUS       AL555333                948 bp    mRNA    linear   EST 30-MAR-2004
DEFINITION  AL555333 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK007YK05 5-PRIME, mRNA sequence.
ACCESSION   AL555333
VERSION     AL555333.3  GI:45860070
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:31277141.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DK007AF03QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK007YK05"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:                1.67e-46        Length:          948
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       1               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x AL555333 (1-948)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 238

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 298

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 358

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 418

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        419 CGAGGGGTGTGTTCCTACTTA 439

RESULT 65
BI161201
LOCUS       BI161201                959 bp    mRNA    linear   EST 05-JUL-2001
DEFINITION  602865659F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5019639 5',
```

```
            mRNA sequence.
ACCESSION   BI161201
VERSION     BI161201.1  GI:14621202
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 959)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1835  row: e  column: 16
            High quality sequence stop: 551.
FEATURES             Location/Qualifiers
     source          1. .959
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5019639"
                     /tissue_type="epithelioid carcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_42"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
ORIGIN

Alignment Scores:
Pred. No.:                1.7e-46        Length:          959
Score:                    448.00         Matches:         87
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       2              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BI161201 (1-959)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 237

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 297

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 357

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 417

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        418 CGAGGGGTGTGTTCCTACTTA 438
```

```
RESULT 66
BM468826
LOCUS       BM468826                 972 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6445782 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5587154
            5', mRNA sequence.
ACCESSION   BM468826
VERSION     BM468826.1  GI:18517868
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 972)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12356  row: d  column: 03
            High quality sequence stop: 707.
FEATURES             Location/Qualifiers
     source          1. .972
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5587154"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_92"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                1.73e-46       Length:          972
Score:                    448.00         Matches:         87
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       3              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BM468826 (1-972)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 114

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 174

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 234

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 294

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
```

```
Db        295 CGAGGGGTGTGTTCCTACTTA 315

RESULT 67
BX338160
LOCUS       BX338160                981 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX338160 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI056YC23 5-PRIME, mRNA sequence.
ACCESSION   BX338160
VERSION     BX338160.2  GI:46283046
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 981)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30345671.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI056AB12QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .981
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI056YC23"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:                1.75e-46      Length:          981
Score:                    448.00        Matches:         87
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BX338160 (1-981)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 232

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 292

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 352

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 412

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        413 CGAGGGGTGTGTTCCTACTTA 433

RESULT 68
BE892886
LOCUS       BE892886                995 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601435730F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920790 5',
            mRNA sequence.
ACCESSION   BE892886
VERSION     BE892886.1  GI:10353510
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 995)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9752  row: 1  column: 07
            High quality sequence stop: 573.
FEATURES             Location/Qualifiers
     source          1. .995
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3920790"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

Alignment Scores:
Pred. No.:                1.78e-46      Length:          995
Score:                    448.00        Matches:         87
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       2             Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BE892886 (1-995)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 248

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 308

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 368

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 428

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
```

```
                    ||||||||||||||||||||
Db        429 CGAGGGGTGTGTTCCTACTTA 449

RESULT 69
BX337141
LOCUS       BX337141                1018 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX337141 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI036YJ02 5-PRIME, mRNA sequence.
ACCESSION   BX337141
VERSION     BX337141.2  GI:46271144
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1018)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30337585.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI036DE01QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .1018
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI036YJ02"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:                1.84e-46        Length:          1018
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BX337141 (1-1018)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 209

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 269

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 329

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 389

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        390 CGAGGGGTGTGTTCCTACTTA 410

RESULT 70
BM554324
LOCUS       BM554324                1035 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6546794 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742466
            5', mRNA sequence.
ACCESSION   BM554324
VERSION     BM554324.1  GI:18793827
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1035)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12760  row: k  column: 11
            High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1. .1035
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5742466"
                     /tissue_type="medulla"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_119"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source normal medulla from
                     anonymous male age 27.  Library  is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed  upon
                     cloning).  Average insert size 1.3 kb, insert size range
                     0.9-3 kb.  Library is normalized and enriched for
                     full-length clones and  was constructed by C. Gruber
                     (Invitrogen).  Research Genetics  tracking code 013. Note:
                     this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                1.88e-46        Length:          1035
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x BM554324 (1-1035)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 391

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 451

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 511
```

```
Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 571

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        572 CGAGGGGTGTGTTCCTACTTA 592

RESULT 71
BX337905
LOCUS       BX337905               1067 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX337905 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI052YN13 5-PRIME, mRNA sequence.
ACCESSION   BX337905
VERSION     BX337905.2  GI:46273926
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1067)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30339657.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI052CG07QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .1067
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI052YN13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:              1.95e-46        Length:         1067
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BX337905 (1-1067)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 243

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 303

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 363

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 423

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        424 CGAGGGGTGTGTTCCTACTTA 444

RESULT 72
BX444691
LOCUS       BX444691               1081 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX444691 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN005YN02 5-PRIME, mRNA sequence.
ACCESSION   BX444691
VERSION     BX444691.2  GI:47009162
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1081)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30780264.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DN005DG01QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .1081
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

Alignment Scores:
Pred. No.:              1.99e-46        Length:         1081
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BX444691 (1-1081)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 245

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 305
```

```
Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         306 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 365

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         366 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 425

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         426 CGAGGGGTGTGTTCCTACTTA 446

RESULT 73
DQ049205
LOCUS       DQ049205               1089 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Homo sapiens TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ049205
VERSION     DQ049205.1  GI:66902404
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1089)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
    PUBMED  15869325
REFERENCE   2  (bases 1 to 1089)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.  Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1. .1089
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
     gene            <1. .>1089
                     /gene="TTRAP"
                     /locus_tag="HC17203"
ORIGIN

Alignment Scores:
Pred. No.:              2.01e-46        Length:         1089
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x DQ049205 (1-1089)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         400 CGAGGGGTGTGTTCCTACTTA 420

RESULT 74
BM555041
LOCUS       BM555041               1103 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6545705 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5737150
            5', mRNA sequence.
ACCESSION   BM555041
VERSION     BM555041.1  GI:18795166
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1103)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12746  row: m  column: 23
            High quality sequence stop: 678.
FEATURES             Location/Qualifiers
     source          1. .1103
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5737150"
                     /tissue_type="duodenal adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_88"
                     /note="Organ: small intestine; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                     oligo-dT primed. Average insert size 1.767 kb. Library
                     enriched for full-length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              2.04e-46        Length:         1103
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x BM555041 (1-1103)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         189 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 248

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         249 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 308
```

```
Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 368

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 428

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        429 CGAGGGGTGTGTTCCTACTTA 449

RESULT 75
CR601303
LOCUS       CR601303               1168 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DN005YN02 of Adult brain of Homo sapiens
            (human).
ACCESSION   CR601303
VERSION     CR601303.1  GI:50482110
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1168)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1168)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1168
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="Adult brain"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Alignment Scores:
Pred. No.:              2.2e-46        Length:         1168
Score:                  448.00         Matches:        87
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     4              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x CR601303 (1-1168)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

Search completed: March 11, 2006, 03:40:21
Job time : 2340.09 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:38:30 ; Search time 27.087 Seconds
(without alignments)
1722.336 Million cell updates/sec

Title: US-10-757-745-2_COPY_115_121
Perfect score: 38
Sequence: 1 SLITWNI 7

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170325_3993/app_query.fasta_1
-DB=N_Geneseq -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1589_@runat_09032006_170325_3993 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	38	100.0	391	11	ADT95307	Adt95307 Colon can
c	2	38	100.0	391	11	ADX41789	Adx41789 Human cDN
	3	38	100.0	483	2	AAX40590	Aax40590 Human sec
	4	38	100.0	752	4	AAH08073	Aah08073 Human cDN
	5	38	100.0	1296	3	AAC98160	Aac98160 Human col
	6	38	100.0	1898	4	AAH15146	Aah15146 Human cDN
	7	38	100.0	1920	3	AAZ47118	Aaz47118 Human CD4
	8	38	100.0	1936	10	ADD19013	Add19013 Human dis
	9	38	100.0	1936	13	ADP25361	Adp25361 PRO polyp
	10	38	100.0	1940	14	ADX06356	Adx06356 Cyclin-de
	11	38	100.0	1948	4	AAI58997	Aai58997 Human pol
	12	38	100.0	1948	5	ADQ99219	Adq99219 DNA encod
	13	38	100.0	1948	9	ADB48979	Adb48979 Novel hum
	14	38	100.0	1958	4	AAI60783	Aai60783 Human pol
	15	38	100.0	2499	2	AAX28153	Aax28153 Topoisome
	16	38	100.0	3152	9	ADA10970	Ada10970 Human cDN
	17	38	100.0	166910	12	ADN01278	Adn01278 Human end
c	18	37	97.4	1168	4	AAI99213	Aai99213 Human exc
c	19	37	97.4	1168	4	AAI99212	Aai99212 Human exc
c	20	37	97.4	1168	5	AAI63563	Aai63563 Human kid
c	21	37	97.4	1168	5	AAI63562	Aai63562 Human kid
	22	37	97.4	2000	11	ACL36488	Acl36488 Rice stre
	23	37	97.4	2000	12	ADJ41437	Adj41437 Plant cDN
	24	36	94.7	47	3	AAZ69238	Aaz69238 Human map
	25	36	94.7	151	4	AAI25284	Aai25284 Probe #15
	26	36	94.7	151	4	ABA71190	Aba71190 Human foe
	27	36	94.7	151	4	AAI51424	Aai51424 Probe #20
	28	36	94.7	151	4	ABA37514	Aba37514 Probe #15
	29	36	94.7	151	4	AAK45479	Aak45479 Human bon
	30	36	94.7	151	4	AAK19488	Aak19488 Human bra
	31	36	94.7	151	4	ABS45165	Abs45165 Human liv
	32	36	94.7	151	6	ABS19747	Abs19747 Human gen
c	33	36	94.7	326	3	AAC31176	Aac31176 Human sec
c	34	36	94.7	404	5	AAF66437	Aaf66437 Novel hum
c	35	36	94.7	582	6	ABN64179	Abn64179 Human can
c	36	36	94.7	582	14	ACL57959	Acl57959 Human col
c	37	36	94.7	672	5	AAS34165	Aas34165 Human cDN
	38	36	94.7	947	14	ADZ63029	Adz63029 Murine Il
c	39	36	94.7	972	6	ABQ61082	Abq61082 Mouse put
c	40	36	94.7	1030	14	ADY19001	Ady19001 DNA encod
c	41	36	94.7	1123	8	ABZ36177	Abz36177 Human sec
c	42	36	94.7	1183	8	ABX13151	Abx13151 Human cDN
	43	36	94.7	1271	12	ADH70063	Adh70063 Human Vbe
	44	36	94.7	1500	3	AAC51584	Aac51584 Arabidops
	45	36	94.7	1503	3	AAC38030	Aac38030 Arabidops
	46	36	94.7	1630	13	ADT17920	Adt17920 Plant cDN
c	47	36	94.7	1661	11	ADM77255	Adm77255 Human mac
c	48	36	94.7	2087	10	ADB61930	Adb61930 Human cDN
	49	36	94.7	2100	14	AEB53889	Aeb53889 DNA encod
	50	36	94.7	2303	10	ADB63770	Adb63770 Human cDN
c	51	36	94.7	2397	8	AAD54616	Aad54616 Human dia
c	52	36	94.7	2560	6	ABQ61204	Abq61204 Protein e
	53	36	94.7	2727	14	AEB53893	Aeb53893 DNA encod
	54	36	94.7	2820	14	AEB53895	Aeb53895 DNA encod
	55	36	94.7	3063	14	AEB53891	Aeb53891 DNA encod
	56	36	94.7	3185	5	AAS91627	Aas91627 DNA encod
c	57	36	94.7	3185	10	ADA53068	Ada53068 Human cod
c	58	36	94.7	3466	12	ADP72747	Adp72747 Renal tox
c	59	36	94.7	3466	12	ADP72511	Adp72511 Renal tox
c	60	36	94.7	3492	8	AAD54615	Aad54615 Human dia
c	61	36	94.7	3556	5	AAF75055	Aaf75055 TR16-long
c	62	36	94.7	3556	10	AAD62893	Aad62893 Human TR1
c	63	36	94.7	3556	13	ADT89650	Adt89650 Human TR1
c	64	36	94.7	3758	6	ABL59504	Abl59504 Human dia
c	65	36	94.7	3758	8	AAD54613	Aad54613 Human dia
c	66	36	94.7	3758	8	AAD54614	Aad54614 Human dia
c	67	36	94.7	3758	10	ADE31700	Ade31700 Human 348
c	68	36	94.7	3759	10	ABX74401	Abx74401 Human cDN
	69	36	94.7	3900	5	AAS64300	Aas64300 DNA encod
	70	36	94.7	3900	13	ADS98356	Ads98356 Protein f
	71	36	94.7	3962	6	ABQ93362	Abq93362 Human cDN
	72	36	94.7	3962	13	ADS98424	Ads98424 Protein f
c	73	36	94.7	4438	2	AAT93393	Aat93393 Bloom's s
c	74	36	94.7	6499	4	AAK91262	Aak91262 Human dig
c	75	36	94.7	11529	14	AEA15412	Aea15412 Human pol

ALIGNMENTS

RESULT 1
ADT95307/c
ID ADT95307 standard; cDNA; 391 BP.
XX
AC ADT95307;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #826.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 826; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 391 BP; 83 A; 100 C; 103 G; 104 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	341	Length:	391
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADT95307 (1-391)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||||
Db        103 TCTCTCATTACCTGGAATATT 83
```

RESULT 2
ADX41789/c
ID ADX41789 standard; cDNA; 391 BP.
XX
AC ADX41789;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 826.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 826; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 391 BP; 83 A; 100 C; 103 G; 104 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	341	Length:	391
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADX41789 (1-391)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        103 TCTCTCATTACCTGGAATATT 83
```

RESULT 3
AAX40590
ID AAX40590 standard; cDNA; 483 BP.
XX
AC AAX40590;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human secreted protein 5' EST SEQ ID No: 190.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS Homo sapiens.
XX
PN WO9906550-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB001232.
XX
PR 01-AUG-1997; 97US-00905144.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
DR WPI; 1999-153780/13.
DR P-PSDB; AAY11868.
XX
PT New isolated prostate-derived nucleic acids - used to develop products
PT which may have cytokine, immune regulatory, haematopoiesis regulating,
PT anti-inflammatory or tumour inhibition activity.
XX
PS Claim 1; Page 298; 675pp; English.
XX
CC AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins expressed in prostate, and encode the proteins
CC given in AAY11716 to AAY11993 respectively. The proteins given represent
CC the signal peptide and an N-terminal fragment of a secreted protein. The
CC nucleic acid sequences can be used for producing secreted human gene
CC products. They can also be used to develop products for diagnosis and
CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation and differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptides can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 483 BP; 123 A; 111 C; 139 G; 110 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	432	Length:	483
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAX40590 (1-483)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363
```

RESULT 4
AAH08073
ID AAH08073 standard; cDNA; 752 BP.
XX
AC AAH08073;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:4908.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 1; SEQ ID NO 4908; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 752 BP; 225 A; 148 C; 200 G; 176 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	710	Length:	752

```
Score:                  38.00          Matches:       7
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            100.0%         Indels:        0
DB:                     4              Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x AAH08073 (1-752)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363
```

RESULT 5
AAC98160
ID AAC98160 standard; cDNA; 1296 BP.
XX
AC AAC98160;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human colon cancer antigen nucleotide sequence SEQ ID NO:170.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200055351-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005883.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587534/55.
DR P-PSDB; AAB53403.
XX
PT Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer.
XX
PS Claim 1; Page 597; 2104pp; English.
XX
CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins may
CC also be used to prevent diseases such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, wounds, renal disorders, infectious diseases, and
CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 1296 BP; 376 A; 250 C; 326 G; 333 T; 0 U; 11 Other;

```
Alignment Scores:
Pred. No.:              1.31e+03       Length:        1296

Score:                  38.00          Matches:       7
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            100.0%         Indels:        0
DB:                     3              Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x AAC98160 (1-1296)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        366 TCTCTCATTACCTGGAATATT 386
```

RESULT 6
AAH15146
ID AAH15146 standard; cDNA; 1898 BP.
XX
AC AAH15146;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:13209.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 13209; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the

CC present invention
XX
SQ Sequence 1898 BP; 590 A; 327 C; 426 G; 555 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.01e+03	Length:	1898
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAH15146 (1-1898)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363
```

RESULT 7
AAZ47118
ID AAZ47118 standard; cDNA; 1920 BP.
XX
AC AAZ47118;
XX
DT 15-MAR-2000 (first entry)
XX
DE Human CD40 receptor associated protein gene.
XX
KW Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW human; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW graft versus host disease; autoimmune disease; ds.
XX
OS Homo sapiens.
XX
PN WO9955859-A2.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1999; 99WO-EP003025.
XX
PR 29-APR-1998; 98EP-00201392.
XX
PA (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Pype SMC, Remacle JEFJG, Huylebroeck DFE;
XX
DR WPI; 2000-062029/05.
DR P-PSDB; AAY56019.
XX
PT Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT diseases and for improvement of anti-tumor treatments.
XX
PS Claim 9; Page 37-39; 48pp; English.
XX
CC This sequence represents the gene encoding human CD40 receptor associated
CC protein (CRAP). CRAP is a functional protein capable of interacting with
CC the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC I, where the protein has no homology to TNF receptor associated factor
CC (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC diseases, and for the improvement of anti-tumour diseases. Diseases which
CC may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC systemic lupus erythematosus, graft rejection, graft versus host disease,
CC allergy, and autoimmune disease. The proteins can be used to sensitize
CC tumour cells to anti-tumour treatments and to screen for compounds which
CC interfere with the interaction of the proteins with other protein
CC components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ Sequence 1920 BP; 599 A; 327 C; 435 G; 557 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	2.03e+03	Length:	1920
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAZ47118 (1-1920)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        362 TCTCTCATTACCTGGAATATT 382
```

RESULT 8
ADD19013
ID ADD19013 standard; DNA; 1936 BP.
XX
AC ADD19013;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human disease related protein DNA sequence SeqID502.
XX
KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW antiarteriosclerotic; vulnerary; gene therapy;
KW hypoxia-regulated condition; tumourigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003018621-A2.
XX
PD 06-MAR-2003.
XX
PF 23-AUG-2002; 2002WO-GB003892.
XX
PR 23-AUG-2001; 2001GB-00020558.
PR 05-OCT-2001; 2001GB-00024037.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
DR WPI; 2003-290046/28.
DR P-PSDB; ADD19012.
XX
PT New substantially purified polypeptide, useful for diagnosing or treating
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT wound healing.
XX
PS Claim 27; SEQ ID NO 502; 424pp; English.
XX
CC This invention relates to novel human genes and gene product which are
CC implicated in certain disease states. Compounds which modulate the
CC proteins of the invention may have cytostatic, antiinflammatory,
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The
CC sequences of the invention may be useful for gene therapy. The invention
CC may be useful for diagnosing or treating a hypoxia-regulated condition,
CC such as tumourigenesis, angiogenesis, apoptosis, inflammation,
CC erythropoiesis, or the biological response to hypoxia conditions
CC including processes such as glycolysis, gluconeogenesis, glucose
CC transportation, catecholamine synthesis, iron transport or nitric oxide
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion

CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC inflammatory conditions or wound healing. The present sequence is that of
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.05e+03	Length:	1936
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADD19013 (1-1936)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        359 TCTCTCATTACCTGGAATATT 379
```

RESULT 9
ADP25361
ID ADP25361 standard; cDNA; 1936 BP.
XX
AC ADP25361;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide encoding cDNA SEQ ID NO:475.
XX
KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-419628/39.
DR P-PSDB; ADP25362.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 1; SEQ ID NO 475; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence encodes a PRO protein of
CC the invention.
XX
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.05e+03	Length:	1936
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	13	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADP25361 (1-1936)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        359 TCTCTCATTACCTGGAATATT 379
```

RESULT 10
ADX06356
ID ADX06356 standard; DNA; 1940 BP.
XX
AC ADX06356;
XX
DT 21-APR-2005 (first entry)
XX
DE Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 921.
XX
KW cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2005012875-A2.
XX
PD 10-FEB-2005.
XX
PF 29-JUL-2004; 2004WO-US024424.
XX
PR 29-JUL-2003; 2003US-0490890P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
XX
DR WPI; 2005-163068/17.
DR P-PSDB; ADX06357.
XX
PT Biomarkers useful for predicting or determining the response of a mammal
PT to a cancer treatment comprising administration of a modulator of cyclin-
PT dependent kinase activity.
XX
PS Claim 5; SEQ ID NO 921; 141pp; English.
XX
CC This invention describes a novel method of predicting or determining
CC whether a mammal will respond or is responding to an anti-cancer agent
CC that modulates cyclin-dependent kinase (cdk) activity. The method
CC comprises measuring the level of one or more biomarkers selected from
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC invention is utilized in a kit for determining or predicting whether

```
CC   patient would be susceptible or resistant to treatment by an agent
CC   modulating cdk activity. The invention also describes a method for
CC   utilizing individualized genetic profiles for treating diseases and
CC   disorders based on patient's response and molecular level, specialized
CC   microarrays comprising the biomarkers described, antibodies directed
CC   against the biomarkers and a cell culture model to identify biomarkers.
CC   The cdk modulator is preferably N-5-[[5-(1,1-Dimethylethyl)-2-
CC   oxazolyl]methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC   tartaric acid salt. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC   sequence encodes a biomarker used in the method of the invention.
XX
SQ   Sequence 1940 BP; 613 A; 330 C; 439 G; 558 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.06e+03        Length:          1940
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       14              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x ADX06356 (1-1940)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        368 TCTCTCATTACCTGGAATATT 388

RESULT 11
AAI58997
ID   AAI58997 standard; cDNA; 1948 BP.
XX
AC   AAI58997;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 1200.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J,  Zhao QA;
PI   Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39841.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders such
PT   as central nervous system injuries.
XX
PS   Claim 1; SEQ ID NO 1200; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC   encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders. Note: The sequence data for this patent did not form
CC   part of the printed specification
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.07e+03        Length:          1948
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       4               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AAI58997 (1-1948)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        381 TCTCTCATTACCTGGAATATT 401

RESULT 12
ADQ99219
ID   ADQ99219 standard; cDNA; 1948 BP.
XX
AC   ADQ99219;
XX
DT   23-SEP-2004  (first entry)
XX
DE   DNA encoding human GPCR-like protein seqid 889.
XX
KW   ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
KW   antidiabetic; GPCR-like protein; ophthalmic disorder;
KW   neurological disorder; immunological disorder; nephritic disorder;
KW   hormonal dysfunction; cancer; atherosclerosis; diabetes;
KW   molecular weight marker; food supplement; human; ss.
XX
OS   Homo sapiens.
XX
PN   US6569662-B1.
XX
PD   27-MAY-2003.
XX
PF   19-JUL-2000; 2000US-00620312.
XX
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Zhou P,  Drmanac RT;
XX
DR   WPI; 2001-442255/47.
XX
PT   New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT   useful for treating diseases of ophthalmic, neurological, immunological
PT   and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
```

PT and diabetes.
XX
PS Example 2; SEQ ID NO 889; 92pp; English.
XX
CC The invention describes an isolated polynucleotide (I) comprising a fully
CC defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC given in the specification, its translated or protein coding portion, its
CC extracellular portion or its active domain. The GPCR-like polypeptides
CC and polynucleotides are useful for the treatment of diseases of
CC ophthalmic, neurological, immunological and nephritic systems. They may
CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC diabetes. The antibodies are useful for detecting or quantitating the
CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement. This sequence represents a human
CC polynucleotide of the invention.
XX
SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.07e+03	Length:	1948
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADQ99219 (1-1948)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        381 TCTCTCATTACCTGGAATATT 401
```

RESULT 13
ADB48979
ID ADB48979 standard; cDNA; 1948 BP.
XX
AC ADB48979;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human cDNA SEQ ID NO 889.
XX
KW ss; cancer; neurodegenerative disease; human.
XX
OS Homo sapiens.
XX
PN US2003104529-A1.
XX
PD 05-JUN-2003.
XX
PF 04-JAN-2002; 2002US-00037270.
XX
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 19-JUL-2000; 2000US-00620312.
XX
PA (ZHOU/) ZHOU P.
PA (TANG/) TANG Y T.
PA (LIUC/) LIU C.
PA (ASUN/) ASUNDI V.
PA (DRMA/) DRMANAC R T.
XX
PI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
XX
DR WPI; 2003-678194/64.
XX
PT New polynucleotide, useful for treating diseases e.g., cancer or
PT neurodegenerative diseases.
XX
PS Claim 1; SEQ ID NO 889; 99pp; English.
XX
CC The invention relates to a polynucleotide comprising a sequence given in
CC the specification, or its mature protein-coding portion, or its
CC complement. The polynucleotide is useful for treating diseases e.g.,
CC cancer or neurodegenerative diseases and many others listed in the
CC specification. The present sequence represents a novel human cDNA. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=20030104529.
XX
SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.07e+03	Length:	1948
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADB48979 (1-1948)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        381 TCTCTCATTACCTGGAATATT 401
```

RESULT 14
AAI60783
ID AAI60783 standard; cDNA; 1958 BP.
XX
AC AAI60783;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4772.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41627.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 4772; 10078pp; English.
XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1958 BP; 613 A; 339 C; 442 G; 564 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.08e+03	Length:	1958
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAI60783 (1-1958)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        393 TCTCTCATTACCTGGAATATT 413
```

RESULT 15
AAX28153
ID AAX28153 standard; DNA; 2499 BP.
XX
AC AAX28153;
XX
DT 16-JUN-1999 (first entry)
XX
DE Topoisomerase II binding protein 1 coding sequence.
XX
KW Topoisomerase II binding protein; TopBP; anticancer agent; ds.
XX
OS Homo sapiens.
XX
PN JP11075856-A.
XX
PD 23-MAR-1999.
XX
PF 17-SEP-1997; 97JP-00251544.
XX
PR 17-SEP-1997; 97JP-00251544.
XX
PA (TSUR/) TSURUO T.
PA (CHUS) CHUGAI PHARM CO LTD.
XX
DR WPI; 1999-257704/22.
DR P-PSDB; AAY03182.
XX
PT New Topoisomerase II- binding protein - useful as an anticancer agent.
XX
PS Disclosure; Page 18-19; 28pp; Japanese.
XX
CC This sequence encodes the topoisomerase II binding protein (TopBP) of the
CC invention. The TopBP protein is useful as an anticancer agent. TopBP can
CC be used as the target molecule for anticancer agent
XX
SQ Sequence 2499 BP; 720 A; 505 C; 587 G; 687 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.73e+03	Length:	2499
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAX28153 (1-2499)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        824 TCTCTCATTACCTGGAATATT 844
```

RESULT 16
ADA10970
ID ADA10970 standard; cDNA; 3152 BP.
XX
AC ADA10970;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human cDNA differentially expressed in colon cancer #60.
XX
KW ss; differential expression; colon cancer; cancer; human.
XX
OS Homo sapiens.
XX
PN US2002160382-A1.
XX
PD 31-OCT-2002.
XX
PF 11-OCT-2001; 2001US-00981353.
XX
PR 11-OCT-2000; 2000US-0239841P.
XX
PA (LASE/) LASEK A W.
PA (JONE/) JONES D A.
XX
PI Lasek AW, Jones DA;
XX
DR WPI; 2003-265756/26.
XX
PT New combination comprising cDNAs that are differentially expressed in
PT colon disorder, useful for diagnosing, treating, staging or monitoring
PT treatment for colon cancers.
XX
PS Claim 1; SEQ ID NO 88; 231pp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in colon disorder. The methods and compositions
CC of the present invention are useful for diagnosing, treating, staging or
CC monitoring treatment for colon cancer. They are also useful in high
CC throughput methods for using cDNAs to detect differential expression of
CC nucleic acids in a sample, screening molecules or compounds to identify a
CC ligand which specifically binds a cDNA and using a protein to screen
CC molecules or compounds to identify at least one ligand which specifically
CC binds the protein. The present sequence represents a human cDNA
CC differentially expressed in colon cancer.
XX
SQ Sequence 3152 BP; 875 A; 643 C; 707 G; 926 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	3.54e+03	Length:	3152
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADA10970 (1-3152)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db       1594 TCTCTCATTACCTGGAATATT 1614
```

RESULT 17
ADN01278
ID ADN01278 standard; DNA; 166910 BP.
XX
AC ADN01278;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human endothelial differentiation gene 2 related DNA #1.
XX
KW cytostatic; gene therapy; endothelial differentiation gene 2;
KW antisense technology; disease identification;
KW hyperproliferative disorder; human; ds.
XX
OS Homo sapiens.
XX
PN US2004092462-A1.
XX
PD 13-MAY-2004.
XX
PF 13-NOV-2002; 2002US-00292337.
XX
PR 13-NOV-2002; 2002US-00292337.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Bennett CF, Dean NM, Dobie KW;
XX
DR WPI; 2004-374980/35.
XX
PT New compound that modulates endothelial differentiation gene 2
PT expression, useful in treating an animal having a disease or condition,
PT i.e. hyperproliferative disorder.
XX
PS Example 15; SEQ ID NO 11; 121pp; English.
XX
CC The invention describes a compound targeted to a nucleic acid molecule
CC encoding endothelial differentiation gene 2. The compound, 8-80
CC nucleobases in length, is targeted to a nucleic acid molecule encoding
CC endothelial differentiation gene 2, where the compound specifically
CC hybridises with the nucleic acid molecule encoding the endothelial
CC differentiation gene 2 comprising a sequence of 1576 bp (SEQ ID NO: 4)
CC and inhibits the expression of endothelial differentiation gene 2. Also
CC described are: a method of inhibiting the expression of the endothelial
CC differentiation gene 2 in cells or tissues; a method of screening for a
CC modulator of endothelial differentiation gene 2; a diagnostic method for
CC identifying a disease state; a kit or assay device comprising the
CC compound; and treating an animal having a disease or condition associated
CC with endothelial differentiation gene 2 compound so that expression of
CC endothelial differentiation gene 2 is inhibited. The compound and the
CC methods are useful in treating an animal having a disease or condition,
CC i.e. hyperproliferative disorder. This sequence represents a human
CC endothelial differentiation gene 2 related polynucleotide.
XX
SQ Sequence 166910 BP; 46470 A; 30922 C; 33180 G; 56338 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3e+05	Length:	166910
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADN01278 (1-166910)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      64005 TCTTTGATCACATGGAATATT 64025
```

RESULT 18
AAI99213/c
ID AAI99213 standard; DNA; 1168 BP.
XX
AC AAI99213;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human excretory related polynucleotide SEQ ID NO 977.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW excretory system; ds.
XX
OS Homo sapiens.
XX
PN WO200155313-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001323.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465569/50.
XX
PT Isolated nucleic acid molecule encoding excretory system antigen is used
PT in preventing, treating or ameliorating a medical condition.
XX
PS Example 2; SEQ ID NO 977; 574pp + Sequence Listing; English.
XX
CC The invention relates to novel excretory system related human
CC polynucleotides (AAI98567-AAI99503) and the encoded proteins (AAM99594-
CC AAM99913) useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy, espeically disorders related
CC to the excretory system. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment and
CC prevention of: (a) cancer, e.g. breast and ovarian cancer and other
CC cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal
CC tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's
CC disease, allergies, autoimmune haemolytic anaemia, autoimmune
CC thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,
CC rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders
CC such as myocardial ischaemias; (d) wound healing; (e) neurological
CC diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases
CC such as viral, bacterial, fungal and parasitic infections. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1168 BP; 420 A; 196 C; 282 G; 270 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.81e+03	Length:	1168
Score:	37.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAI99213 (1-1168)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||:::|||||||||||||
Db       1014 TCTCTTGTTACCTGGAATATT 994
```

RESULT 19
AAI99212/c
ID AAI99212 standard; DNA; 1168 BP.
XX
AC AAI99212;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human excretory related polynucleotide SEQ ID NO 976.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;

KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW excretory system; ds.
XX
OS Homo sapiens.
XX
PN WO200155313-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001323.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465569/50.
XX
PT Isolated nucleic acid molecule encoding excretory system antigen is used
PT in preventing, treating or ameliorating a medical condition.
XX
PS Example 2; SEQ ID NO 976; 574pp + Sequence Listing; English.
XX
CC The invention relates to novel excretory system related human
CC polynucleotides (AAI98567-AAI99503) and the encoded proteins (AAM99594-
CC AAM99913) useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy, espeically disorders related
CC to the excretory system. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment and
CC prevention of: (a) cancer, e.g. breast and ovarian cancer and other
CC cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal
CC tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's
CC disease, allergies, autoimmune haemolytic anaemia, autoimmune
CC thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,
CC rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders
CC such as myocardial ischaemias; (d) wound healing; (e) neurological
CC diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases
CC such as viral, bacterial, fungal and parasitic infections. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1168 BP; 420 A; 196 C; 281 G; 271 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.81e+03	Length:	1168
Score:	37.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAI99212 (1-1168)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db       1014 TCTCTTGTTACCTGGAATATT 994
```

RESULT 20
AAI63563/c
ID AAI63563 standard; DNA; 1168 BP.
XX
AC AAI63563;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human kidney related polynucleotide SEQ ID NO 878.
XX
KW Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW nootropic; neuroprotective; antibacterial; virucide; fungicide;
KW opthalmological; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antiulcer; vulnerary; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; ds.
XX
OS Homo sapiens.

XX
PN WO200155323-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001343.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488784/53.
XX
PT New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating and/or preventing human diseases and disorders.
XX
PS Disclosure; SEQ ID NO 878; 564pp + Sequence Listing; English.
XX
CC The invention relates to novel kidney related polynucleotides (AAI62971-
CC AAI63793) and the encoded polypeptides (AAM42417-AAM42691) collectively
CC known as kidney antigens and the use of such kidney antigens for
CC detecting disorders of the kidney, especially kidney cancer and kidney
CC cancer metastases. The polynucleotides and proteins are also useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. The genes are isolated from a range of human tissues
CC disclosed in the specification. The nucleic acids, proteins, antibodies
CC and (ant)agonists are useful in the diagnosis, treatment and prevention
CC of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1168 BP; 420 A; 196 C; 282 G; 270 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.81e+03	Length:	1168
Score:	37.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAI63563 (1-1168)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db       1014 TCTCTTGTTACCTGGAATATT 994
```

RESULT 21
AAI63562/c
ID AAI63562 standard; DNA; 1168 BP.
XX
AC AAI63562;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human kidney related polynucleotide SEQ ID NO 877.
XX
KW Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW nootropic; neuroprotective; antibacterial; virucide; fungicide;
KW opthalmological; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antiulcer; vulnerary; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; ds.
XX
OS Homo sapiens.
XX
PN WO200155323-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001343.
XX

PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488784/53.
XX
PT New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating and/or preventing human diseases and disorders.
XX

```
PS   Disclosure; SEQ ID NO 877; 564pp + Sequence Listing; English.
XX
CC   The invention relates to novel kidney related polynucleotides (AAI62971-
CC   AAI63793) and the encoded polypeptides (AAM42417-AAM42691) collectively
CC   known as kidney antigens and the use of such kidney antigens for
CC   detecting disorders of the kidney, especially kidney cancer and kidney
CC   cancer metastases. The polynucleotides and proteins are also useful for
CC   preventing, treating or ameliorating medical conditions e.g. by protein
CC   or gene therapy. The genes are isolated from a range of human tissues
CC   disclosed in the specification. The nucleic acids, proteins, antibodies
CC   and (ant)agonists are useful in the diagnosis, treatment and prevention
CC   of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
CC   adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC   lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC   allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC   diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC   arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC   myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC   cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC   bacterial, fungal and parasitic infections. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1168 BP; 420 A; 196 C; 281 G; 271 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.81e+03        Length:          1168
Score:                    37.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              97.4%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AAI63562 (1-1168)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       1014 TCTCTTGTTACCTGGAATATT 994

RESULT 22
ACL36488
ID   ACL36488 standard; cDNA; 2000 BP.
XX
AC   ACL36488;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Rice stress-regulated promoter SEQ ID NO:15051.
XX
KW   ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW   agriculture.
XX
OS   Oryza sativa.
XX
PN   WO2003008540-A2.
XX
PD   30-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019668.
XX
PR   22-JUN-2001; 2001US-0300112P.
PR   24-AUG-2001; 2001US-0314662P.
PR   26-SEP-2001; 2001US-0325277P.
PR   21-NOV-2001; 2001US-0332132P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Kreps J,  Briggs SP,  Cooper B,  Glazebrook J,  Goff SA,  Katagiri F;
PI   Moughamer T,  Provart N,  Ricke D,  Zhu T;
XX
DR   WPI; 2003-248011/24.
XX
PT   New stress-responsive nucleic acid, useful for altering the
PT   responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT   stress, salt stress or osmotic stress.
XX
PS   Claim 48; SEQ ID NO 15051; 89pp; English.
XX
CC   The invention relates to novel abiotic stress responsive polynucleotides
CC   and polypeptides. Also disclosed are vectors, expression cassettes, host
CC   cells, and plants containing such polynucleotides. Also disclosed are
CC   methods for using the polynucleotides and polypeptides to alter the
CC   responsiveness of a plant to abiotic stress. The invention is useful in
CC   agriculture. The nucleic acid is useful for determining whether a test
CC   plant has been exposed to an abiotic stress condition. It is also useful
CC   for selecting an agent that alters abiotic stress regulated
CC   polynucleotide expression in a plant cell, and to identify a homolog or
CC   ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC   molecule and the polypeptide encoded by it are useful in altering the
CC   responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC   stress, osmotic stress or any of their combinations. The present sequence
CC   is used in the exemplification of the invention
XX
SQ   Sequence 2000 BP; 471 A; 534 C; 514 G; 476 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.:                3.3e+03         Length:          2000
Score:                    37.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              97.4%           Indels:          0
DB:                       11              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x ACL36488 (1-2000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       1442 TCGTTAGTTACCTGGAACATT 1462

RESULT 23
ADJ41437
ID   ADJ41437 standard; cDNA; 2000 BP.
XX
AC   ADJ41437;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Plant cDNA #2437.
XX
KW   Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW   soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW   maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW   stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW   plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW   antifungal.
XX
OS   Eukaryota.
XX
PN   US2004016025-A1.
XX
PD   22-JAN-2004.
XX
PF   26-SEP-2002; 2002US-00260238.
XX
PR   26-SEP-2001; 2001US-0325277P.
PR   26-SEP-2001; 2001US-0325448P.
PR   04-APR-2002; 2002US-0370620P.
XX
PA   (BUDW/) BUDWORTH P.
PA   (MOUG/) MOUGHAMER T.
PA   (BRIG/) BRIGGS S P.
PA   (COOP/) COOPER B.
PA   (GLAZ/) GLAZEBROOK J.
PA   (GOFF/) GOFF S A.
PA   (KATA/) KATAGIRI F.
```

PA (KREP/) KREPS J.
PA (PROV/) PROVART N.
PA (RICK/) RICKE D.
PA (ZHUT/) ZHU T.
XX
PI Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
DR WPI; 2004-190374/18.
XX
PT New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT or high nutritional value.
XX
PS Claim 1; SEQ ID NO 2437; 230pp; English.
XX
CC The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2000 BP; 471 A; 534 C; 514 G; 476 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	3.3e+03	Length:	2000
Score:	37.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADJ41437 (1-2000)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db       1442 TCGTTAGTTACCTGGAACATT 1462
```

RESULT 24
AAZ69238
ID AAZ69238 standard; DNA; 47 BP.
XX
AC AAZ69238;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human map-related biallelic marker SEQ ID NO:3594.
XX
KW Human genome; biallelic marker; high density disequilibrium map;
KW genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW haplotyping; hybridisation; identification; characterisation; diagnosis;
KW single nucleotide polymorphism; SNP; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(24,C)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO9954500-A2.
XX
PD 28-OCT-1999.
XX
PF 21-APR-1999; 99WO-IB000822.
XX
PR 21-APR-1998; 98US-0082614P.
PR 23-NOV-1998; 98US-0109732P.
XX
PA (GEST) GENSET.
XX
PI Cohen D, Blumenfeld M, Chumakov I;
XX
DR WPI; 2000-013267/01.
XX
PT Novel biallelic markers used to construct a high density disequilibrium
PT map of the human genome.
XX
PS Claim 3; Page 997; 2745pp; English.
XX
CC AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC invention, which contain a polymorphic base at position 24 of their
CC nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC primers for the biallelic markers. The biallelic markers of the invention
CC have a variety of uses: they can be used for high density mapping of the
CC human genome, and in complex association studies and haplotyping studies
CC which are useful in determining the genetic basis for disease states.
CC Compositions and methods of the invention can also be useful for the
CC identification of the targets for the development of pharmaceutical
CC agents and diagnostic methods, as well as the characterisation of the
CC differential efficacious responses to and side effects from
CC pharmaceutical agents acting on a disease as well as other treatment.
CC N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
CC 3367, are not actually given a sequence in the Sequence Listing from the
CC present invention
XX
SQ Sequence 47 BP; 15 A; 4 C; 11 G; 17 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	76.2	Length:	47
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAZ69238 (1-47)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db         20 TCTTTAGTGACATGGAATGTA 40
```

RESULT 25
AAI25284
ID AAI25284 standard; DNA; 151 BP.
XX
AC AAI25284;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #15217 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.

```
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 15217; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 151 BP; 34 A; 32 C; 25 G; 60 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                283             Length:          151
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       4               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AAI25284 (1-151)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        120 TCATTACTAACTTGGAATATT 140

RESULT 26
ABA71190
ID   ABA71190 standard; DNA; 151 BP.
XX
AC   ABA71190;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #19495.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 19495; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 151 BP; 34 A; 32 C; 25 G; 60 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                283             Length:          151
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       4               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x ABA71190 (1-151)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        120 TCATTACTAACTTGGAATATT 140

RESULT 27
AAI51424
ID   AAI51424 standard; DNA; 151 BP.
XX
AC   AAI51424;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #20110 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
```

PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 20110; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 151 BP; 34 A; 32 C; 25 G; 60 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	283	Length:	151
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAI51424 (1-151)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        120 TCATTACTAACTTGGAATATT 140
```

RESULT 28
ABA37514
ID ABA37514 standard; DNA; 151 BP.
XX
AC ABA37514;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #15980 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 4; SEQ ID NO 15980; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 151 BP; 34 A; 32 C; 25 G; 60 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	283	Length:	151
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ABA37514 (1-151)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        120 TCATTACTAACTTGGAATATT 140
```

RESULT 29
AAK45479
ID AAK45479 standard; DNA; 151 BP.
XX
AC AAK45479;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 20036.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 20036; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 151 BP; 34 A; 32 C; 25 G; 60 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	283	Length:	151
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAK45479 (1-151)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        120 TCATTACTAACTTGGAATATT 140
```

RESULT 30
AAK19488
ID AAK19488 standard; DNA; 151 BP.
XX
AC AAK19488;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 19479.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 19479; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
SQ Sequence 151 BP; 34 A; 32 C; 25 G; 60 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	283	Length:	151
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAK19488 (1-151)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        120 TCATTACTAACTTGGAATATT 140
```

RESULT 31
ABS45165
ID ABS45165 standard; DNA; 151 BP.
XX
AC ABS45165;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 20155.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 4; SEQ ID NO 20155; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 151 BP; 34 A; 32 C; 25 G; 60 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	283	Length:	151
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ABS45165 (1-151)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        120 TCATTACTAACTTGGAATATT 140

RESULT 32
ABS19747
ID   ABS19747 standard; DNA; 151 BP.
XX
AC   ABS19747;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 19738.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 4; SEQ ID NO 19738; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe open reading frame of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 151 BP; 34 A; 32 C; 25 G; 60 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              283          Length:          151
Score:                  36.00        Matches:         6
Percent Similarity:     100.0%       Conservative:    1
Best Local Similarity:  85.7%        Mismatches:      0
Query Match:            94.7%        Indels:          0
DB:                     6            Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x ABS19747 (1-151)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        120 TCATTACTAACTTGGAATATT 140

RESULT 33
AAC31176/c
ID   AAC31176 standard; cDNA; 326 BP.
XX
AC   AAC31176;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 35251.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 35251; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
```

CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 326 BP; 83 A; 83 C; 82 G; 76 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	670	Length:	326
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAC31176 (1-326)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        140 TCTCTCATCACATGGAATCTG 120
```

RESULT 34
AAF66437/c
ID AAF66437 standard; cDNA; 404 BP.
XX
AC AAF66437;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 2193.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US018374.
XX
PR 02-JUL-1999; 99US-0142310P.
PR 02-JUL-1999; 99US-0142311P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J;
PI Kassam A, Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G;
PI Drmanac R, Crkenjakov R, Drmanac S, Dickson M, Labat I;
PI Leshkowitiz D, Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences.
XX
PS Claim 9; Page 864; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia
XX
SQ Sequence 404 BP; 134 A; 68 C; 98 G; 103 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	853	Length:	404
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAF66437 (1-404)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        230 TCATTACTAACTTGGAATATT 210
```

RESULT 35
ABN64179/c
ID ABN64179 standard; cDNA; 582 BP.
XX
AC ABN64179;
XX
DT 28-JUN-2002 (first entry)
XX
DE Human cancer related polynucleotide SEQ ID NO 4146.
XX
KW Human; cytostatic; gene expression; gene mapping; tissue profiling;
KW gene therapy; cancer; tumour; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200214500-A2.
XX
PD 21-FEB-2002.
XX
PF 16-AUG-2001; 2001WO-US025840.
XX
PR 16-AUG-2000; 2000US-0226326P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
XX
DR WPI; 2002-241905/29.
XX
PT New nucleic acid for producing a polypeptide, detecting differentially
PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth.
XX
PS Claim 1; SEQ ID NO 4146; 883pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC with cytostatic activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumour growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The
CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

```
SQ   Sequence 582 BP; 239 A; 97 C; 105 G; 141 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.28e+03        Length:         582
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ABN64179 (1-582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        436 TCTTTACTTACCTGGAATATT 416

RESULT 36
ACL57959/c
ID   ACL57959 standard; cDNA; 582 BP.
XX
AC   ACL57959;
XX
DT   24-MAR-2005  (first entry)
XX
DE   Human colon cancer differentially expressed polynucleotide, SEQ ID:4094.
XX
KW   Differential expression; diagnosis; therapy; drug screening; cancer;
KW   neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO2005000087-A2.
XX
PD   06-JAN-2005.
XX
PF   13-MAY-2004; 2004WO-US015421.
XX
PR   03-JUN-2003; 2003US-0475872P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Randazzo F,  Moler E,  Escobedo J,  Garcia PD;
XX
DR   WPI; 2005-075421/08.
XX
PT   New isolated polynucleotides, which are differentially expressed in colon
PT   cancer cell, useful for treating cancer, e.g. colon cancer, breast
PT   cancer, or pancreatic cancer.
XX
PS   Claim 1; SEQ ID NO 4094; 97pp; English.
XX
CC   The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
CC   are differentially expressed in colon cancer cells. The invention also
CC   relates to vectors and host cells comprising a differentially expressed
CC   polynucleotide of the invention; a method for detecting a cancerous cell
CC   by detection of a gene product of the polynucleotides; a method for
CC   inhibiting a cancerous phenotype of a cell by inhibiting a gene product
CC   of the polynucleotides; a method of treating an individual with cancer by
CC   administration of a modulator of a gene product of the polynucleotides;
CC   and an isolated antibody that specifically binds to a polypeptide encoded
CC   by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
CC   antibodies, and methods are useful for the detection of cancerous cells;
CC   for the diagnosis, prognosis and management of cancer; for the
CC   identification of agents that modulate the phenotype of cancerous cells;
CC   for the identification of therapeutic targets for cancer chemotherapy;
CC   and for the treatment of cancer, especially colon cancer and metastasized
CC   colon cancer, but also breast or pancreatic cancer. The polynucleotides
CC   are also useful as a source of probes or primers for use in diagnostic
CC   methods. The differentially expressed polynucleotides or their encoded
CC   proteins can additionally be used as vaccines to modulate primary immune
CC   responses for the prevention or treatment of cancer. The present sequence
CC   represents a specifically claimed polynucleotide which is differentially
CC   expressed in colon cancer. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 582 BP; 239 A; 97 C; 105 G; 141 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.28e+03        Length:         582
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ACL57959 (1-582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        436 TCTTTACTTACCTGGAATATT 416

RESULT 37
AAS34165/c
ID   AAS34165 standard; cDNA; 672 BP.
XX
AC   AAS34165;
XX
DT   17-DEC-2001  (first entry)
XX
DE   Human cDNA encoding a novel foetal antigen, SEQ ID No 689.
XX
KW   Human; foetal tissue antigen; ss; antiinflammatory; neuroprotective;
KW   immunomodulator; cardiovascular; cytostatic; nephrothropic;
KW   cardiovascular; autoimmune disease; rheumatoid arthritis;
KW   hyperproliferative disorder; breast neoplasm; cancer;
KW   cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW   cerebral ischaemia; angiogenesis; nervous system disorder;
KW   Alzheimer's disease; infection; ocular disorder; corneal infection;
KW   wound healing; epithelial cell proliferation; food additive.
XX
OS   Homo sapiens.
XX
PN   WO200155312-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001321.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
```

PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488782/53.
DR P-PSDB; AAU21345.
XX
PT New polynucleotides and polypeptides for diagnosing, treating, preventing
PT or prognosing e.g. diseases or disorders of the nervous, musculoskeletal,
PT excretory, gastrointestinal, reproductive, and respiratory systems.
XX
PS Claim 1; SEQ ID NO 689; 642pp; English.
XX
CC The invention relates to novel nucleic acids encoding novel human foetal
CC antigens. The nucleic acids and proteins are used to prevent, treat (e.g.
CC by gene therapy) or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. The antibodies to the antigens can also be used
CC in alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities, fat content, lipid, protein, carbohydrate,
CC vitamins, minerals, cofactors and other nutritional components. Numerous

CC examples of diseases and disorders treated by the nucleic acids and
CC proteins are given in the specification. The present sequence encodes a
CC foetal antigen of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in

Alignment Scores:

Pred. No.:	1.51e+03	Length:	672
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAS34165 (1-672)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db         24 TCATTACTAACTTGGAATATT 4
```

RESULT 38
ADZ63029
ID ADZ63029 standard; DNA; 947 BP.
XX
AC ADZ63029;
XX
DT 30-JUN-2005 (first entry)
XX
DE Murine Il1rap gene, SEQ ID 2753.
XX
KW Drug screening; Peroxisome Proliferator-Activated Receptor alpha;
KW PPAR alpha; gene; ds.
XX
OS Mus musculus.
XX
PN US2005084872-A1.
XX
PD 21-APR-2005.
XX
PF 23-JAN-2004; 2004US-00764420.
XX
PR 24-JAN-2003; 2003US-0442797P.
PR 30-MAY-2003; 2003US-0474413P.
XX
PA (LUMP/) LUM P Y.
PA (TANY/) TAN Y.
PA (DAIH/) DAI H.
PA (MUIS/) MUISE E S.
PA (BERG/) BERGER J P.
PA (THOM/) THOMPSON J R.
XX
PI Lum PY, Tan Y, Dai H, Muise ES, Berger JP, Thompson JR;
XX
DR WPI; 2005-313963/32.
DR GENBANK; AI255955.
XX
PT Determining whether agent has defined biological activity, by making
PT comparison e.g. comparing efficacy value, toxicity value and classifier
PT value of agent to reference values, and using results to determine
PT whether agent has activity.
XX
PS Example 4; SEQ ID NO 2753; 51pp; English.
XX
CC The present invention relates to a methods for screening biologically
CC active agents, such as candidate drugs, to determine whether the agent
CC possesses a defined biological activity. The methods involve making one
CC or more comparisons chosen from comparing efficacy value of agent to
CC reference efficacy value, comparing a toxicity value of the agent to
CC reference toxicity value, comparing a classifier value of the agent to
CC reference classifier value, and using the comparison result(s) to
CC determine whether the agent possesses the defined biological activity.
CC Also claimed is a population of oligonucleotide probes (I) specific for
CC measuring the expression levels of members of a classifier population of
CC genes or a toxicity-related population of genes. (I) are useful for
CC measuring the expression levels of genes that are useful for identifying
CC agonists or partial agonists of Peroxisome Proliferator-Activated
CC Receptor (PPAR) gamma. ADZ63072-ADZ63959 are oligonucleotide probes which
CC are useful in the method of the invention to measure the expression
CC pattern of mouse genes ADZ62273-ADZ63071 that yield an expression pattern
CC that correlates with the stimulation of PPARalpha receptors by an agent.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=20050084872.
XX
SQ Sequence 947 BP; 262 A; 185 C; 146 G; 350 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.:	2.22e+03	Length:	947
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	14	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADZ63029 (1-947)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        429 TCTTTACTCACTTGGAATATT 449
```

RESULT 39
ABQ61082/c
ID ABQ61082 standard; cDNA; 972 BP.
XX
AC ABQ61082;
XX
DT 26-FEB-2003 (first entry)
XX
DE Mouse putative protein #14 encoding sequence.
XX
KW Neuroprotective; immunomodulator; cancer; cytostatic; anti-inflammatory;
KW gene therapy; nutritional supplement; wound; burn; ulcer;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW autoimmune disorder; inflammation; vulnerary; gene; ss.
XX
OS Mus musculus.
XX
PN WO200231111-A2.
XX
PD 18-APR-2002.
XX
PF 11-OCT-2001; 2001WO-US027760.
XX
PR 12-OCT-2000; 2000US-00687527.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-426278/45.
DR N-PSDB; ABP43838.
XX
PT New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation.
XX
PS Claim 1; SEQ ID # 295; 357pp + Sequence Listing; English.
XX
CC The invention relates to 446 newly isolated polynucleotide sequences. The
CC activity of polynucleotides of the invention may be described as,
CC vulnerary, neuroprotective, immunomodulator, cytostatic and anti-
CC inflammatory. Compositions comprising nucleic acids of the invention are
CC useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,

```
CC   amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC   inflammation. The nucleic acids and polypeptides are also useful in
CC   diagnostic and research methods. The sequences given in records ABQ60788-
CC   ABQ61233 represent polynucleotides of the invention. NOTE: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 972 BP; 300 A; 159 C; 225 G; 288 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.28e+03        Length:         972
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ABQ61082 (1-972)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        199 AGCATTATTACGTGGAACATC 179

RESULT 40
ADY19001/c
ID   ADY19001 standard; DNA; 1030 BP.
XX
AC   ADY19001;
XX
DT   05-MAY-2005  (first entry)
XX
DE   DNA encoding a PRO polypeptide, SEQ ID NO 4807.
XX
KW   Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW   Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW   Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
KW   Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW   Antiallergic; ds; gene; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2005016962-A2.
XX
PD   24-FEB-2005.
XX
PF   11-AUG-2004; 2004WO-US026249.
XX
PR   11-AUG-2003; 2003US-0493546P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Abbas A,  Clark H,  Ouyang W,  Williams MP,  Wood WI,  Wu TD;
XX
DR   WPI; 2005-182330/19.
XX
PT   New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT   treating an immune related disorder, e.g. systemic lupus erythematosus,
PT   rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
PS   Claim 1; SEQ ID NO 4807; 158pp; English.
XX
CC   The invention relates to an isolated nucleic acid encoding a PRO
CC   polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC   composition, and method are useful for diagnosing and treating an immune
CC   related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC   arthritis. The present sequence represents a PRO polypeptide.
XX
SQ   Sequence 1030 BP; 346 A; 154 C; 205 G; 325 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.44e+03        Length:         1030
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ADY19001 (1-1030)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db         69 AGCATTATTACGTGGAACATC 49

RESULT 41
ABZ36177/c
ID   ABZ36177 standard; cDNA; 1123 BP.
XX
AC   ABZ36177;
XX
DT   10-FEB-2003  (first entry)
XX
DE   Human secretory polynucleotide SPTM SEQ ID NO 341.
XX
KW   Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
KW   asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
KW   Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
KW   multiple sclerosis; Parkinson's disease; cell proliferative disorder;
KW   anti-inflammatory; immunosuppressive; neuroprotective; nootropic;
KW   neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
KW   antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
KW   secretory polynucleotide; secretory protein; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200283876-A2.
XX
PD   24-OCT-2002.
XX
PF   27-MAR-2002; 2002WO-US009921.
XX
PR   29-MAR-2001; 2001US-0280067P.
PR   29-MAR-2001; 2001US-0280068P.
PR   16-MAY-2001; 2001US-0291280P.
PR   17-MAY-2001; 2001US-0291829P.
PR   17-MAY-2001; 2001US-0291849P.
PR   19-JUN-2001; 2001US-0299428P.
PR   20-JUN-2001; 2001US-0299776P.
PR   20-JUN-2001; 2001US-0300001P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Daffo A,  Jones AL,  Tran AB,  Dahl CR,  Gietzen D,  Chinn J;
PI   Dufour GE,  Hillman JL,  Yu JY,  Tuason O,  Yap PE,  Amshey SR;
PI   Daughtery SC,  Dam TC,  Liu TF,  Nguyen DA,  Kleefeld Y,  Gerstin EH;
PI   Peralta CH,  David MH,  Lewis SA,  Chen AJ,  Panzer SR,  Harris B;
PI   Flores V,  Marwaha R,  Lo A,  Lan RY,  Urashka ME;
XX
DR   WPI; 2003-075543/07.
DR   P-PSDB; ABP75734.
XX
PT   New human secretory proteins and polynucleotides, useful for diagnosing,
PT   treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
PT   neurological disorders (e.g. Alzheimer's), or cell proliferations or
PT   cancers.
XX
PS   Claim 1; SEQ ID NO 341; 458pp + Sequence Listing; English.
XX
CC   The invention relates to a secretory polynucleotide (designated sptm)
CC   comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
CC   naturally occurring polynucleotide sequence at least 90 % identical to
CC   the polynucleotide sequence, a polynucleotide complementary to them or an
CC   RNA equivalent of them. The polypeptide or polynucleotide are useful for
CC   treating, preventing or diagnosing a disease or condition associated with
CC   the expression of functional SPTM. These are particularly useful for
CC   diagnosing, treating or preventing autoimmune/inflammatory disorders
```

```
CC   (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
CC   disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC   dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
CC   multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
CC   schizophrenia or amnesia), or cell proliferative disorders (e.g.
CC   psoriasis, polycythemia vera, or cancers including adenocarcinoma,
CC   leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
CC   breast, cervix or prostate). Note: The sequence data for this patent did
CC   not form part of the printed specfication, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1123 BP; 358 A; 179 C; 247 G; 339 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.69e+03        Length:         1123
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ABZ36177 (1-1123)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        119 AGCATTATTACCTGGAACATC 99

RESULT 42
ABX13151/c
ID   ABX13151 standard; cDNA; 1183 BP.
XX
AC   ABX13151;
XX
DT   10-MAY-2003  (first entry)
XX
DE   Human cDNA for kinase and phosphatase, KPP-5, INCYTE No. 603340CB1.
XX
KW   Human; ss; gene; kinase and phosphatase; KPP; cancer; cirrhosis;
KW   cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW   hepatitis; paroxysmal nocturnal haemoglobinuria; polycythaemia vera;
KW   psoriasis; primary thromobocytopaenia; developmental disorder;
KW   renal tubular acidosis; anaemia; mental retardation; AIDS; epilepsy;
KW   neurological disorder; Alzheimer's disease; Parkinson's disease;
KW   autoimmune disorder; inflammatory disorder; allergy; asthma;
KW   acquired immunodeficiency syndrome; autoimmune thyroiditis;
KW   contact dermatitis; Crohn's disease; diabetes mellitus;
KW   glomerulonephritis; Goodpasture's syndrome; gout; Graves' disease;
KW   Hashimoto's thyroiditis; irritable bowel syndrome; multiple sclerosis;
KW   osteoarthritis; osteoporosis; pancreatitis; Reiter's syndrome;
KW   rheumatoid arthritis; Sjogren's syndrome; uveitis; infection.
XX
OS   Homo sapiens.
XX
PN   WO2003012065-A2.
XX
PD   13-FEB-2003.
XX
PF   01-AUG-2002; 2002WO-US024521.
XX
PR   02-AUG-2001; 2001US-0309627P.
PR   07-AUG-2001; 2001US-0310933P.
PR   09-AUG-2001; 2001US-0311323P.
PR   07-SEP-2001; 2001US-0317820P.
PR   14-SEP-2001; 2001US-0322264P.
PR   28-SEP-2001; 2001US-0326098P.
PR   19-DEC-2001; 2001US-0343007P.
PR   15-MAR-2002; 2002US-0364494P.
PR   24-APR-2002; 2002US-0375539P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Baughn MR,  Yue H,  Walia NK,  He A,  Au-Young JK,  Lee SY;
PI   Gietzen KJ,  Lal PG,  Elliott VS,  Ison CH,  Yang J,  Lee EA,  Li JX;
PI   Emerling BM,  Richardson TW,  Warren BA,  Hafalia AJA,  Marquis JP;
XX
DR   WPI; 2003-239519/23.
DR   P-PSDB; ABU08104.
XX
PT   New human kinases and phosphatases and polynucleotides, useful for
PT   diagnosing, treating or preventing autoimmune or inflammatory disorders
PT   (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
PT   cancer or hepatitis.
XX
PS   Claim 5; Page 186-187; 199pp; English.
XX
CC   The invention relates to an isolated polypeptide, which is a human kinase
CC   and phosphatase, KPP (KPP-1 to KPP-18). Also included are the encoding
CC   polynucleotides KPP NA, recombinant polynucleotide comprising a promoter
CC   sequence operably linked to KPP NA, a cell transformed with the
CC   recombinant polynucleotide, a transgenic organism comprising the
CC   recombinant polynucleotide;, an anti-KPP antibody, screening for
CC   ant/agonists of KPP, screening for compounds which bind to or alter the
CC   activity or expression of KPP, microarray where at least one element is
CC   KPP NA, generating an expression profile of a sample containing
CC   polynucleotides and an array comprising different nucleotide molecules
CC   affixed in distinct physical locations on a solid substrate (where at
CC   least one of the nucleotide molecules comprises a first oligonucleotide
CC   or polynucleotide sequence specifically hybridisable with at least 30
CC   contiguous nucleotides of a target KPP NA). The kinases and phosphatases
CC   (KPP) polypeptides, polynucleotides, agonists and antagonists are useful
CC   for diagnosing, treating or preventing disorders associated with aberrant
CC   expression of KPP, particularly cell proliferative disorders (e.g.
CC   arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal
CC   nocturnal haemoglobinuria, polycythaemia vera, psoriasis, primary
CC   thromobocytopaenia or cancer), developmental disorders (renal tubular
CC   acidosis, anaemia or mental retardation), neurological disorders (e.g.
CC   Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/
CC   inflammatory disorders (e.g. AIDS, acquired immunodeficiency syndrome,
CC   allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's
CC   disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,
CC   gout, Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
CC   multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
CC   syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
CC   bacterial, fungal, parasitic, protozoan or helminthic infections. They
CC   are also useful in the assessment of the effects of exogenous compounds
CC   on the expression of nucleic acid and aa sequences of proteins associated
CC   with KPP. The polynucleotides encoding KPP are useful for creating
CC   transgenic animals to model human disease. The present sequence encodes a
CC   KPP protein of the invention
XX
SQ   Sequence 1183 BP; 366 A; 195 C; 269 G; 353 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.85e+03        Length:         1183
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ABX13151 (1-1183)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        181 AGCATTATTACGTGGAACATC 161

RESULT 43
ADH70063
ID   ADH70063 standard; DNA; 1271 BP.
XX
AC   ADH70063;
XX
DT   25-MAR-2004  (first entry)
XX
DE   Human Vbeta gene family genomic sequence #25.
XX
```

KW human; T-cell associated disease; Vbeta; autoimmune disease;
KW degenerative nervous system disease; graft versus host disease;
KW hypersensitivity disease; infectious disease; neoplastic disease;
KW Addison's disease; atrophic gastritis;
KW degenerative nervous system disease; multiple sclerosis;
KW Alzheimer's disease; hypersensitivity disease; type I hypersensitivity;
KW allergy; type II hypersensitivity; Goodpasture's syndrome;
KW type IV hypersensitivity; leprosy; infectious disease; viral infection;
KW HIV; fungal infection; Candida; parasitic infection; schistosome;
KW filaria; bacterial infection; Mycobacterium; neoplastic disease;
KW lymphoproliferative disease; leukaemia; lymphoma; cancer; brain cancer;
KW breast cancer; ds.
XX
OS Homo sapiens.
XX
PN US2002150891-A1.
XX
PD 17-OCT-2002.
XX
PF 05-MAR-1999; 99US-00263959.
XX
PR 19-SEP-1994; 94US-00309335.
PR 19-SEP-1995; 95US-00531241.
XX
PA (HOOD/) HOOD L E.
PA (ROWE/) ROWEN L.
XX
PI Hood LE, Rowen L;
XX
DR WPI; 2004-059052/06.
XX
PT Kit for diagnozing and treating T-cell associated diseases e.g.
PT autoimmune, degenerative nervous system and infectious disease, comprises
PT nucleic acid primers specifically priming and allowing amplification of a
PT Vbeta gene.
XX
PS Disclosure; SEQ ID NO 257; 164pp; English.
XX
CC The invention relates to a kit for diagnosing and treating T-cell
CC associated diseases which comprises a panel of nucleic acid primers
CC specifically priming and allowing amplification of each Vbeta gene,
CC VbetaRNA or cDNA. The kit is useful for diagnosing organ transplant
CC rejection and diagnosing and treating T-cell associated diseases
CC including autoimmune diseases, degenerative nervous system diseases,
CC graft versus host disease, hypersensitivity diseases, infectious diseases
CC and neoplastic diseases. Autoimmune diseases include Addison's disease,
CC atrophic gastritis. Degenerative nervous system diseases include multiple
CC sclerosis and Alzheimer's disease. Hypersensitivity diseases include Type
CC I hypersensitivities such as contact with allergens that lead to
CC allergies, Type II hypersensitivities such as those present in
CC Goodpasture's syndrome and Type IV hypersensitivities such as those
CC manifested in leprosy. Infectious diseases include viral infections
CC caused by viruses such as HIV, fungal infections such as those caused by
CC the yeast genus Candida, parasitic infections such as those caused by
CC schistosomes, filaria and bacterial infections such as those caused by
CC Mycobacterium. Neoplastic diseases include lymphoproliferative diseases
CC such as leukaemias, lymphomas and cancers such as cancer of the brain,
CC breast. The present sequence represents a Vbeta gene family genomic
CC sequence.
XX
SQ Sequence 1271 BP; 339 A; 290 C; 279 G; 363 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.09e+03	Length:	1271
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADH70063 (1-1271)

Qy 1 SerLeuIleThrTrpAsnIle 7

||||||||||||||||||:::
Db 266 TCTTTAATTACCTGGAATTTG 286

RESULT 44
AAC51584
ID AAC51584 standard; DNA; 1500 BP.
XX
AC AAC51584;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69049.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	3.72e+03	Length:	1500
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAC51584 (1-1500)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        705 TCTTTGCTTACATGGAACATT 725
```

```
RESULT 45
AAC38030
ID   AAC38030 standard; DNA; 1503 BP.
XX
AC   AAC38030;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 19536.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
PR   27-JUL-1999;   99US-0145918P.
PR   27-JUL-1999;   99US-0145919P.
PR   28-JUL-1999;   99US-0145951P.
PR   02-AUG-1999;   99US-0146386P.
PR   02-AUG-1999;   99US-0146388P.
PR   02-AUG-1999;   99US-0146389P.
PR   03-AUG-1999;   99US-0147038P.
PR   04-AUG-1999;   99US-0147204P.
PR   04-AUG-1999;   99US-0147302P.
PR   05-AUG-1999;   99US-0147192P.
PR   05-AUG-1999;   99US-0147260P.
PR   06-AUG-1999;   99US-0147303P.
PR   06-AUG-1999;   99US-0147416P.
PR   09-AUG-1999;   99US-0147493P.
PR   09-AUG-1999;   99US-0147935P.
PR   10-AUG-1999;   99US-0148171P.
PR   11-AUG-1999;   99US-0148319P.
PR   12-AUG-1999;   99US-0148341P.
PR   13-AUG-1999;   99US-0148565P.
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
PR   17-AUG-1999;   99US-0149175P.
PR   18-AUG-1999;   99US-0149426P.
PR   20-AUG-1999;   99US-0149722P.
```

```
PR   20-AUG-1999;   99US-0149723P.
PR   20-AUG-1999;   99US-0149929P.
PR   23-AUG-1999;   99US-0149902P.
PR   23-AUG-1999;   99US-0149930P.
PR   25-AUG-1999;   99US-0150566P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

Alignment Scores:
Pred. No.:              3.72e+03        Length:         1503
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AAC38030 (1-1503)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        705 TCTTTGCTTACATGGAACATT 725

RESULT 46
```

```
ADT17920
ID   ADT17920 standard; cDNA; 1630 BP.
XX
AC   ADT17920;
XX
DT   13-JAN-2005  (first entry)
XX
DE   Plant cDNA, Seq ID 3246.
XX
KW   Plant; ss; gene; transgenic; cold tolerance; growth rate;
KW   drought tolerance; disease resistance; galactomannan production;
KW   plant growth regulator; heat tolerance; herbicide tolerance;
KW   lignin production; extreme osmotic condition tolerance;
KW   pathogens resistance; pest resistance; yield improvement; seed oil yield;
KW   seed protein yield.
XX
OS   Viridiplantae.
XX
PN   US2004216190-A1.
XX
PD   28-OCT-2004.
XX
PF   18-DEC-2003; 2003US-00739930.
XX
PR   28-APR-2003; 2003US-00424599.
PR   28-APR-2003; 2003US-00425115.
XX
PA   (KOVA/) KOVALIC D K.
XX
PI   Kovalic DK;
XX
DR   WPI; 2004-757369/74.
XX
PT   New recombinant DNA constructs useful in the field of biochemistry and
PT   genetics, and in particular for producing transgenic plants with improved
PT   biological characteristics.
XX
PS   Claim 1; SEQ ID NO 3246; 14pp; English.
XX
CC   The invention relates a recombinant DNA construct comprising a
CC   polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC   1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC   (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC   Arabidopsis, wheat and rape but the specification does not indicate which
CC   sequences is derived from which organism. Also included is a method of
CC   producing a plant having an improved property, comprising transforming a
CC   plant with a recombinant DNA construct comprising a promoter region
CC   functional in a plant cell operably joined to a polynucleotide encoding a
CC   polypeptide associated with the property, and growing the transformed
CC   plant. The property is selected from improving plant cold tolerance, for
CC   manipulating growth rate in plant cells by modification of the cell cycle
CC   pathway, for improving plant drought tolerance, for providing increased
CC   resistance to plant disease, for galactomannan production, for production
CC   of plant growth regulators, for improving plant heat tolerance, for
CC   improving plant tolerance to herbicides, for increasing the rate of
CC   homologous recombination in plants, for lignin production, for improving
CC   plant tolerance to extreme osmotic conditions, for improving plant
CC   tolerance to pathogens or pests, for yield improvement by modification of
CC   photosynthesis, for modifying seed oil yield and/or content, for
CC   modifying seed protein yield and/or content, for yield improvement by
CC   modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC   and for yield improvement by providing improved plant growth and
CC   development under at least one stress condition. The polynucleotide may
CC   also encode a plant transcription factor. The methods and compositions of
CC   the present invention are useful in the field of biochemistry and
CC   genetics, in particular for producing transgenic plants with improved
CC   biological characteristics such as increased yield, improved nitrogen
CC   flow, increasing plant tolerance to cold or heat, improving plant
CC   tolerance to extreme osmotic and drought conditions, and improving plant
CC   tolerance to plant pests or pathogens. They can also be used in physical
CC   arrays of molecules, plant breeding markers, computer-based storage and
CC   analysis systems. The present sequence is one of the 5544 plant cDNA
CC   sequences of the invention. Note: The sequence data for this patent did
```

```
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
SQ   Sequence 1630 BP; 403 A; 382 C; 359 G; 486 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.08e+03        Length:          1630
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       13              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x ADT17920 (1-1630)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        485 AGCCTCTTAACCTGGAACATT 505

RESULT 47
ADM77255/c
ID   ADM77255 standard; cDNA; 1661 BP.
XX
AC   ADM77255;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Human macroprotein 24.97 encoding cDNA, seq id 1.
XX
KW   Human; macroprotein; 24.97; dementia; facial paralysis; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             484. .1167
FT                   /*tag=  a
FT                   /product= "macroprotein 24.97"
XX
PN   CN1376704-A.
XX
PD   30-OCT-2002.
XX
PF   22-MAR-2001; 2001CN-00105758.
XX
PR   22-MAR-2001; 2001CN-00105758.
XX
PA   (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2003-184961/19.
DR   P-PSDB; ADM77256.
XX
PT   A novel human macroprotein -24.97 polypeptide, and the polynucleotide
PT   encoding it, useful for treating e.g. dementia, facial paralysis.
XX
PS   Claim 6; SEQ ID NO 1; 31pp; Chinese.
XX
CC   The invention relates to a human macroprotein designated 24.97. Further
CC   disclosed is the polynucleotide encoding the, a process for preparing the
CC   polypeptide by DNA recombination and the application of the polypeptide
CC   in treating diseases e.g. dementia, facial paralysis. The antagonist of
CC   the polypeptide and its medical action, and the application of the
CC   polynucleotide are also disclosed. The current sequence represents the
CC   human macroprotein 24.97 encoding cDNA.
XX
SQ   Sequence 1661 BP; 554 A; 282 C; 340 G; 485 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.17e+03        Length:          1661
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       11              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x ADM77255 (1-1661)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        973 TCATTACTAACTTGGAATATT 953

RESULT 48
ADB61930/c
ID   ADB61930 standard; cDNA; 2087 BP.
XX
AC   ADB61930;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human cDNA encoding clone BRACE20007330.
XX
KW   Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
KW   tissue regeneration; cell regeneration; membrane protein;
KW   signal transduction-related protein; transcription-related protein;
KW   osteoporosis; neurological disease; cancer; tumour.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             339. .2015
FT                   /*tag=  a
FT                   /product= "Clone BRACE20007330 protein"
XX
PN   EP1308459-A2.
XX
PD   07-MAY-2003.
XX
PF   28-MAR-2002; 2002EP-00007401.
XX
PR   05-NOV-2001; 2001JP-00379298.
PR   25-JAN-2002; 2002US-00350978.
XX
PA   (HELI-) HELIX RES INST.
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Sugiyama T,  Otsuki T,  Wakamatsu A,  Sato H,  Ishii S;
PI   Yamamoto J,  Isono Y,  Hio Y,  Otsuka K,  Nagai K,  Irie R,  Tamechika I;
PI   Seki N,  Yoshikawa T,  Otsuka M,  Nagahari K,  Masuho Y;
XX
DR   WPI; 2003-450961/43.
DR   P-PSDB; ADB63900.
XX
PT   New polynucleotides and polypeptides, useful for developing a diagnostic
PT   marker or medicines for regulation of their expression and activity, or
PT   as targets of gene therapy.
XX
PS   Claim 1; Page; 222pp; English.
XX
CC   The invention discloses a polynucleotide comprising a sequence selected
CC   from 1970 fully defined nucleotide sequences which encode novel
CC   polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC   or its partial peptide, an antibody binding to the polypeptide or peptide
CC   of the polynucleotide, immunologically assaying the polypeptide or
CC   peptide of the polynucleotide by contacting the polypeptide or peptide
CC   with the antibody of the encoded protein, and observing the binding
CC   between the two, a transformant carrying the polynucleotide in an
CC   expressible manner and an antisense polynucleotide. The oligonucleotide
CC   is useful as a primer for synthesising the polynucleotide, or as a probe
CC   for detecting the polynucleotide. The polynucleotides and encoded
CC   proteins are useful as pharmaceutical agents and many disease-related
CC   genes may be included in them, for developing a diagnostic marker or
CC   medicines for regulation of their expression and activity, or as targets
CC   of gene therapy. The genes are involved in tissue and/or cell
CC   regeneration. Membrane proteins, signal transduction-related proteins,
```

CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a cDNA of the invention. Note: Some of the sequence
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.
XX
SQ Sequence 2087 BP; 631 A; 383 C; 489 G; 584 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.38e+03	Length:	2087
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADB61930 (1-2087)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        164 AGCATTATTACGTGGAACATC 144
```

RESULT 49
AEB53889
ID AEB53889 standard; DNA; 2100 BP.
XX
AC AEB53889;
XX
DT 06-OCT-2005 (first entry)
XX
DE DNA encoding prostate cancer-specific hypothetical protein FLJ, seqid 36.
XX
KW screening; diagnosis; cancer; prostate tumor; cytostatic; drug screening;
KW antibody therapy; immunoconjugate; immuno-diagnosis; immunotherapy;
KW anion transport protein; ds; gene; splice variant.
XX
OS Homo sapiens.
XX
PN WO2005067629-A2.
XX
PD 28-JUL-2005.
XX
PF 04-JAN-2005; 2005WO-US000040.
XX
PR 06-JAN-2004; 2004US-0534419P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Weigle B, Ebner R;
XX
DR WPI; 2005-563837/57.
DR P-PSDB; AEB53890.
XX
PT Identifying agent that modulates activity of cancer-related gene, by
PT contacting compound with cell, to promote expression of gene, and
PT detecting difference in expression of gene relative to when compound is
PT not present, to identify agent.
XX
PS Claim 1; SEQ ID NO 36; 131pp; English.
XX
CC The present invention relates to methods of screening cancer-linked genes
CC and expression products for cancer diagnosis, and for screening potential
CC anti-cancer agents. Specifically claimed is a method of identifying (M1)
CC an agent that modulates activity of cancer-related genes, by contacting a
CC compound with a cell expressing a gene under conditions promoting the
CC expression of gene, and detecting difference in expression relative to
CC when the compound is not present. Also claimed are methods of identifying
CC (M2) an anti-neoplastic agent, by contacting a cell exhibiting neoplastic
CC activity with a compound first identified as a cancer related gene
CC modulator by (M1), and detecting a decrease in the neoplastic, or
CC administering agent identified by (M1) to an animal exhibiting a cancer
CC condition and detecting a decrease in the cancerous condition. Also given
CC is a protein (I), which has at least one immunogenic fragment; an
CC antibody that reacts with a protein; an immunoconjugate comprising the
CC antibody and a cytotoxic agent. The agent is useful for treating cancer,
CC by contacting a cancerous cell in vivo with an agent having activity
CC against an expression produced encoded by a gene sequence given in the
CC specification. The agent is also useful for treating cancer in an animal,
CC by eliciting the production of cytotoxic T lymphocytes specific for the
CC protein. Also described is a method of diagnosing prostate cancer. The
CC proteins are anion transport proteins specific for prostate cancer. The
CC present sequence is splice variant 1 DNA encoding a prostate cancer-
CC specific hypothetical protein FLJ.
XX
SQ Sequence 2100 BP; 730 A; 344 C; 426 G; 600 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.42e+03	Length:	2100
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	14	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AEB53889 (1-2100)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        241 AGCTTAGTAACTTGGAATGTT 261
```

RESULT 50
ADB63770
ID ADB63770 standard; cDNA; 2303 BP.
XX
AC ADB63770;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA encoding clone UTERU20040730.
XX
KW Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
KW tissue regeneration; cell regeneration; membrane protein;
KW signal transduction-related protein; transcription-related protein;
KW osteoporosis; neurological disease; cancer; tumour.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1009. .2303
FT /*tag= a
FT /partial
FT /product= "Clone UTERU20040730 protein"
XX
PN EP1308459-A2.
XX
PD 07-MAY-2003.
XX
PF 28-MAR-2002; 2002EP-00007401.
XX
PR 05-NOV-2001; 2001JP-00379298.
PR 25-JAN-2002; 2002US-00350978.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-450961/43.
DR P-PSDB; ADB65740.
XX
PT New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or

PT as targets of gene therapy.
XX
PS Claim 1; Page; 222pp; English.
XX
CC The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesising the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a cDNA of the invention. Note: Some of the sequence
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.
XX
SQ Sequence 2303 BP; 712 A; 421 C; 516 G; 654 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.01e+03	Length:	2303
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADB63770 (1-2303)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db        993 AGCTTAGTAACTTGGAATGTT 1013
```

RESULT 51
AAD54616/c
ID AAD54616 standard; DNA; 2397 BP.
XX
AC AAD54616;
XX
DT 26-JUN-2003 (first entry)
XX
DE Human diacylglycerol kinase (DGK) DNA #17.
XX
KW Diacylglycerol kinase; p53 pathway; diacylglycerol kinase; DAG; cancer;
KW DGK; gene therapy; cytostatic; enzyme; ds.
XX
OS Homo sapiens.
XX
PN WO200299060-A2.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017527.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 22-OCT-2001; 2001US-0338733P.
PR 15-FEB-2002; 2002US-0357253P.
PR 15-FEB-2002; 2002US-0357600P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
DR WPI; 2003-201301/19.
XX
PT Identifying a candidate p53 pathway modulating agent for diagnosing or
PT treating cancer by detecting a test agent-biased activity of the assay
PT system comprising a purified DGK (Diacylglycerol kinase) polypeptide or
PT nucleic acid.
XX
PS Disclosure; Page 74-75; 111pp; English.
XX
CC The present invention relates to a method for identifying a candidate p53
CC pathway modulating agent. The method involves providing an assay system
CC comprising a purified diacylglycerol (DAG) kinase (DGK) polypeptide or
CC nucleic acid, contacting the assay system with a test agent and detecting
CC the test agent-biased activity of the assay system where a difference
CC between the test agent-biased activity and reference activity identifies
CC the agent as a candidate p53 pathway modulating agent. The method is
CC useful for manufacturing a medicament for diagnosing or treating breast,
CC colon, lung or ovary cancer. The invention is also useful in gene
CC therapy. The present sequence is human DGK DNA
XX
SQ Sequence 2397 BP; 638 A; 605 C; 655 G; 499 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.29e+03	Length:	2397
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAD54616 (1-2397)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db       1810 TCTCTCATCACATGGAATCTG.1790
```

RESULT 52
ABQ61204/c
ID ABQ61204 standard; cDNA; 2560 BP.
XX
AC ABQ61204;
XX
DT 26-FEB-2003 (first entry)
XX
DE Protein expressed in tumour cells encoding sequence.
XX
KW Neuroprotective; immunomodulator; cancer; chromosome 7q21; cytostatic;
KW anti-inflammatory; gene therapy; nutritional supplement; wound; burn;
KW ulcer; Alzheimer's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
KW vulnerary; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231111-A2.
XX
PD 18-APR-2002.
XX
PF 11-OCT-2001; 2001WO-US027760.
XX
PR 12-OCT-2000; 2000US-00687527.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-426278/45.
DR N-PSDB; ABP43960.
XX
PT New polypeptides and their encoded proteins, useful as nutritional

```
PT   sources or supplements, or in gene therapy, particularly for treating
PT   wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT   inflammation.
XX
PS   Claim 1; SEQ ID # 417; 357pp + Sequence Listing; English.
XX
CC   The invention relates to 446 newly isolated polynucleotide sequences. The
CC   activity of polynucleotides of the invention may be described as,
CC   vulnerary, neuroprotective, immunomodulator, cytostatic and anti-
CC   inflammatory. Compositions comprising nucleic acids of the invention are
CC   useful for treating a mammalian subject, or as nutritional sources or
CC   supplements. These are useful in gene therapy, particularly for treating
CC   wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC   amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC   inflammation. The nucleic acids and polypeptides are also useful in
CC   diagnostic and research methods. The sequences given in records ABQ60788-
CC   ABQ61233 represent polynucleotides of the invention. NOTE: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2560 BP; 760 A; 544 C; 586 G; 670 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.77e+03        Length:         2560
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ABQ61204 (1-2560)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db       2336 TCATTACTAACTTGGAATATT 2316

RESULT 53
AEB53893
ID   AEB53893 standard; DNA; 2727 BP.
XX
AC   AEB53893;
XX
DT   06-OCT-2005  (first entry)
XX
DE   DNA encoding prostate cancer-specific hypothetical protein FLJ, seqid 40.
XX
KW   screening; diagnosis; cancer; prostate tumor; cytostatic; drug screening;
KW   antibody therapy; immunoconjugate; immuno-diagnosis; immunotherapy;
KW   anion transport protein; ds; gene; splice variant.
XX
OS   Homo sapiens.
XX
PN   WO2005067629-A2.
XX
PD   28-JUL-2005.
XX
PF   04-JAN-2005; 2005WO-US000040.
XX
PR   06-JAN-2004; 2004US-0534419P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Weigle B,  Ebner R;
XX
DR   WPI; 2005-563837/57.
DR   P-PSDB; AEB53894.
XX
PT   Identifying agent that modulates activity of cancer-related gene, by
PT   contacting compound with cell, to promote expression of gene, and
PT   detecting difference in expression of gene relative to when compound is
PT   not present, to identify agent.
XX
PS   Claim 1; SEQ ID NO 40; 131pp; English.
XX
CC   The present invention relates to methods of screening cancer-linked genes
CC   and expression products for cancer diagnosis, and for screening potential
CC   anti-cancer agents. Specifically claimed is a method of identifying (M1)
CC   an agent that modulates activity of cancer-related genes, by contacting a
CC   compound with a cell expressing a gene under conditions promoting the
CC   expression of gene, and detecting difference in expression relative to
CC   when the compound is not present. Also claimed are methods of identifying
CC   (M2) an anti-neoplastic agent, by contacting a cell exhibiting neoplastic
CC   activity with a compound first identified as a cancer related gene
CC   modulator by (M1), and detecting a decrease in the neoplastic, or
CC   administering agent identified by (M1) to an animal exhibiting a cancer
CC   condition and detecting a decrease in the cancerous condition. Also given
CC   is a protein (I), which has at least one immunogenic fragment; an
CC   antibody that reacts with a protein; an immunoconjugate comprising the
CC   antibody and a cytotoxic agent. The agent is useful for treating cancer,
CC   by contacting a cancerous cell in vivo with an agent having activity
CC   against an expression produced encoded by a gene sequence given in the
CC   specification. The agent is also useful for treating cancer in an animal,
CC   by eliciting the production of cytotoxic T lymphocytes specific for the
CC   protein. Also described is a method of diagnosing prostate cancer. The
CC   proteins are anion transport proteins specific for prostate cancer. The
CC   present sequence is splice variant 3 DNA encoding a prostate cancer-
CC   specific hypothetical protein FLJ.
XX
SQ   Sequence 2727 BP; 921 A; 460 C; 533 G; 813 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                7.27e+03        Length:         2727
Score:                    36.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AEB53893 (1-2727)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||:::
Db        241 AGCTTAGTAACTTGGAATGTT 261

RESULT 54
AEB53895
ID   AEB53895 standard; DNA; 2820 BP.
XX
AC   AEB53895;
XX
DT   06-OCT-2005  (first entry)
XX
DE   DNA encoding prostate cancer-specific hypothetical protein FLJ, seqid 42.
XX
KW   screening; diagnosis; cancer; prostate tumor; cytostatic; drug screening;
KW   antibody therapy; immunoconjugate; immuno-diagnosis; immunotherapy;
KW   anion transport protein; ds; gene; splice variant.
XX
OS   Homo sapiens.
XX
PN   WO2005067629-A2.
XX
PD   28-JUL-2005.
XX
PF   04-JAN-2005; 2005WO-US000040.
XX
PR   06-JAN-2004; 2004US-0534419P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Weigle B,  Ebner R;
XX
DR   WPI; 2005-563837/57.
DR   P-PSDB; AEB53896.
XX
```

PT Identifying agent that modulates activity of cancer-related gene, by
PT contacting compound with cell, to promote expression of gene, and
PT detecting difference in expression of gene relative to when compound is
PT not present, to identify agent.
XX
PS Claim 1; SEQ ID NO 42; 131pp; English.
XX
CC The present invention relates to methods of screening cancer-linked genes
CC and expression products for cancer diagnosis, and for screening potential
CC anti-cancer agents. Specifically claimed is a method of identifying (M1)
CC an agent that modulates activity of cancer-related genes, by contacting a
CC compound with a cell expressing a gene under conditions promoting the
CC expression of gene, and detecting difference in expression relative to
CC when the compound is not present. Also claimed are methods of identifying
CC (M2) an anti-neoplastic agent, by contacting a cell exhibiting neoplastic
CC activity with a compound first identified as a cancer related gene
CC modulator by (M1), and detecting a decrease in the neoplastic, or
CC administering agent identified by (M1) to an animal exhibiting a cancer
CC condition and detecting a decrease in the cancerous condition. Also given
CC is a protein (I), which has at least one immunogenic fragment; an
CC antibody that reacts with a protein; an immunoconjugate comprising the
CC antibody and a cytotoxic agent. The agent is useful for treating cancer,
CC by contacting a cancerous cell in vivo with an agent having activity
CC against an expression produced encoded by a gene sequence given in the
CC specification. The agent is also useful for treating cancer in an animal,
CC by eliciting the production of cytotoxic T lymphocytes specific for the
CC protein. Also described is a method of diagnosing prostate cancer. The
CC proteins are anion transport proteins specific for prostate cancer. The
CC present sequence is splice variant 4 DNA encoding a prostate cancer-
CC specific hypothetical protein FLJ.
XX
SQ Sequence 2820 BP; 965 A; 460 C; 556 G; 839 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.55e+03	Length:	2820
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	14	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AEB53895 (1-2820)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||||:::
Db        241 AGCTTAGTAACTTGGAATGTT 261
```

RESULT 55
AEB53891
ID AEB53891 standard; DNA; 3063 BP.
XX
AC AEB53891;
XX
DT 06-OCT-2005 (first entry)
XX
DE DNA encoding prostate cancer-specific hypothetical protein FLJ, seqid 38.
XX
KW screening; diagnosis; cancer; prostate tumor; cytostatic; drug screening;
KW antibody therapy; immunoconjugate; immuno-diagnosis; immunotherapy;
KW anion transport protein; ds; gene; splice variant.
XX
OS Homo sapiens.
XX
PN WO2005067629-A2.
XX
PD 28-JUL-2005.
XX
PF 04-JAN-2005; 2005WO-US000040.
XX
PR 06-JAN-2004; 2004US-0534419P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Weigle B, Ebner R;
XX
DR WPI; 2005-563837/57.
DR P-PSDB; AEB53892.
XX
PT Identifying agent that modulates activity of cancer-related gene, by
PT contacting compound with cell, to promote expression of gene, and
PT detecting difference in expression of gene relative to when compound is
PT not present, to identify agent.
XX
PS Claim 1; SEQ ID NO 38; 131pp; English.
XX
CC The present invention relates to methods of screening cancer-linked genes
CC and expression products for cancer diagnosis, and for screening potential
CC anti-cancer agents. Specifically claimed is a method of identifying (M1)
CC an agent that modulates activity of cancer-related genes, by contacting a
CC compound with a cell expressing a gene under conditions promoting the
CC expression of gene, and detecting difference in expression relative to
CC when the compound is not present. Also claimed are methods of identifying
CC (M2) an anti-neoplastic agent, by contacting a cell exhibiting neoplastic
CC activity with a compound first identified as a cancer related gene
CC modulator by (M1), and detecting a decrease in the neoplastic, or
CC administering agent identified by (M1) to an animal exhibiting a cancer
CC condition and detecting a decrease in the cancerous condition. Also given
CC is a protein (I), which has at least one immunogenic fragment; an
CC antibody that reacts with a protein; an immunoconjugate comprising the
CC antibody and a cytotoxic agent. The agent is useful for treating cancer,
CC by contacting a cancerous cell in vivo with an agent having activity
CC against an expression produced encoded by a gene sequence given in the
CC specification. The agent is also useful for treating cancer in an animal,
CC by eliciting the production of cytotoxic T lymphocytes specific for the
CC protein. Also described is a method of diagnosing prostate cancer. The
CC proteins are anion transport proteins specific for prostate cancer. The
CC present sequence is splice variant 2 DNA encoding a prostate cancer-
CC specific hypothetical protein FLJ.
XX
SQ Sequence 3063 BP; 1064 A; 511 C; 591 G; 897 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8.28e+03	Length:	3063
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	14	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AEB53891 (1-3063)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        241 AGCTTAGTAACTTGGAATGTT 261
```

RESULT 56
AAS91627
ID AAS91627 standard; cDNA; 3185 BP.
XX
AC AAS91627;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #27431.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX

```
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG27440.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 27431; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3185 BP; 1019 A; 632 C; 716 G; 818 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                8.65e+03        Length:         3185
Score:                    36.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AAS91627 (1-3185)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db       1454 AGCTTAGTAACTTGGAATGTT 1474

RESULT 57
ADA53068/c
ID   ADA53068 standard; cDNA; 3185 BP.
XX
AC   ADA53068;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Human coding sequence, SEQ ID 636.
XX
KW   Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW   Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW   inflammatory disease; osteoporosis; neurological disease; gene; ss.
XX
OS   Homo sapiens.
XX
PN   EP1293569-A2.
XX
PD   19-MAR-2003.
XX
PF   21-MAR-2002; 2002EP-00006586.
XX
PR   14-SEP-2001; 2001JP-00328381.
PR   24-JAN-2002; 2002US-0350435P.
XX
PA   (HELI-) HELIX RES INST.
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Sugiyama T,  Otsuki T,  Wakamatsu A,  Sato H,  Ishii S;
PI   Yamamoto J,  Isono Y,  Hio Y,  Otsuka K,  Nagai K,  Irie R,  Tamechika I;
PI   Seki N,  Yoshikawa T,  Otsuka M,  Nagahari K,  Masuho Y;
XX
DR   WPI; 2003-395539/38.
DR   P-PSDB; ADA54707.
XX
PT   New polynucleotides encoding full-length polypeptides, e.g. secretory
PT   and/or membrane proteins, useful for developing medicines for diseases in
PT   which the gene is involved, or as target molecules for gene therapy.
XX
PS   Claim 1; SEQ ID NO 636; 205pp; English.
XX
CC   The present invention relates to novel human secretory or membrane
CC   proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC   ADA54071). The coding sequences are useful in the gene therapy of
CC   diseases caused by abnormalities of the proteins, e.g. cancer,
CC   inflammatory diseases, osteoporosis or neurological disease.
XX
SQ   Sequence 3185 BP; 960 A; 660 C; 719 G; 846 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                8.65e+03        Length:         3185
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       10              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ADA53068 (1-3185)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db       2536 TCATTACTAACTTGGAATATT 2516

RESULT 58
ADP72747/c
ID   ADP72747 standard; DNA; 3466 BP.
XX
AC   ADP72747;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Renal toxin progression gene marker #1336.
XX
KW   ds; toxic effect; gene expression profile; kidney tissue;
KW   differential gene expression; toxicity progression; toxicity marker;
KW   drug screening; toxicity assay; kidney pathology; nephritis;
KW   kidney necrosis; glomerular injury; tubular injury;
KW   focal segmental glomerulosclerosis.
XX
OS   Rattus norvegicus.
XX
PN   WO2004048598-A2.
XX
PD   10-JUN-2004.
XX
PF   24-NOV-2003; 2003WO-US037556.
XX
PR   22-NOV-2002; 2002US-00301856.
XX
PA   (GENE-) GENE LOGIC INC.
XX
```

PI Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
PI Elashoff M;
XX
DR WPI; 2004-460771/43.
XX
PT Predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.
XX
PS Claim 11; SEQ ID NO 1336; 266pp; English.
XX
CC The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The
CC methods are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC kidney necrosis, glomerular and tubular injury, or focal segmental
CC glomerulosclerosis. The methods are useful for determining the similarity
CC of a toxic response to one or more individual compounds and for
CC predicting or elucidating the potential cellular pathways influenced,
CC induced or modulated by the compound or test agent. The kit is useful for
CC predicting or modelling the toxic response of a test compound, for
CC monitoring the progression of renal disease states, for identifying genes
CC that show promise as new drug targets and for screening known and newly
CC designed drugs. This sequence corresponds to a gene marker used in the
CC method of the invention. (Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 3466 BP; 941 A; 873 C; 912 G; 740 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9.51e+03	Length:	3466
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADP72747 (1-3466)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db       2097 TCTCTCATCACATGGAATCTG 2077
```

RESULT 59
ADP72511/c
ID ADP72511 standard; DNA; 3466 BP.
XX
AC ADP72511;
XX
DT 26-AUG-2004 (first entry)
XX
DE Renal toxin progression gene marker #1100.
XX
KW ds; toxic effect; gene expression profile; kidney tissue;
KW differential gene expression; toxicity progression; toxicity marker;
KW drug screening; toxicity assay; kidney pathology; nephritis;
KW kidney necrosis; glomerular injury; tubular injury;
KW focal segmental glomerulosclerosis.
XX
OS Rattus norvegicus.
XX
PN WO2004048598-A2.
XX
PD 10-JUN-2004.
XX
PF 24-NOV-2003; 2003WO-US037556.
XX
PR 22-NOV-2002; 2002US-00301856.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
PI Elashoff M;
XX
DR WPI; 2004-460771/43.
XX
PT Predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.
XX
PS Claim 11; SEQ ID NO 1100; 266pp; English.
XX
CC The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The
CC methods are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC kidney necrosis, glomerular and tubular injury, or focal segmental
CC glomerulosclerosis. The methods are useful for determining the similarity
CC of a toxic response to one or more individual compounds and for
CC predicting or elucidating the potential cellular pathways influenced,
CC induced or modulated by the compound or test agent. The kit is useful for
CC predicting or modelling the toxic response of a test compound, for
CC monitoring the progression of renal disease states, for identifying genes
CC that show promise as new drug targets and for screening known and newly
CC designed drugs. This sequence corresponds to a gene marker used in the
CC method of the invention. (Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 3466 BP; 941 A; 873 C; 912 G; 740 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9.51e+03	Length:	3466
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x ADP72511 (1-3466)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db       2097 TCTCTCATCACATGGAATCTG 2077
```

RESULT 60
AAD54615/c
ID AAD54615 standard; DNA; 3492 BP.
XX
AC AAD54615;
XX
DT 26-JUN-2003 (first entry)
XX
DE Human diacylglycerol kinase (DGK) DNA #16.
XX
KW Diacylglycerol kinase; p53 pathway; diacylglycerol kinase; DAG; cancer;
KW DGK; gene therapy; cytostatic; enzyme; ds.
XX
OS Homo sapiens.

```
XX
PN   WO200299060-A2.
XX
PD   12-DEC-2002.
XX
PF   03-JUN-2002; 2002WO-US017527.
XX
PR   05-JUN-2001; 2001US-0296076P.
PR   10-OCT-2001; 2001US-0328605P.
PR   22-OCT-2001; 2001US-0338733P.
PR   15-FEB-2002; 2002US-0357253P.
PR   15-FEB-2002; 2002US-0357600P.
XX
PA   (EXEL-) EXELIXIS INC.
XX
PI   Friedman L,  Plowman GD,  Belvin M,  Francis-Lang H,  Li D,  Funke RP;
XX
DR   WPI; 2003-201301/19.
XX
PT   Identifying a candidate p53 pathway modulating agent for diagnosing or
PT   treating cancer by detecting a test agent-biased activity of the assay
PT   system comprising a purified DGK (Diacylglycerol kinase) polypeptide or
PT   nucleic acid.
XX
PS   Disclosure; Page 72-74; 111pp; English.
XX
CC   The present invention relates to a method for identifying a candidate p53
CC   pathway modulating agent. The method involves providing an assay system
CC   comprising a purified diacylglycerol (DAG) kinase (DGK) polypeptide or
CC   nucleic acid, contacting the assay system with a test agent and detecting
CC   the test agent-biased activity of the assay system where a difference
CC   between the test agent-biased activity and reference activity identifies
CC   the agent as a candidate p53 pathway modulating agent. The method is
CC   useful for manufacturing a medicament for diagnosing or treating breast,
CC   colon, lung or ovary cancer. The invention is also useful in gene
CC   therapy. The present sequence is human DGK DNA
XX
SQ   Sequence 3492 BP; 931 A; 888 C; 910 G; 763 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.59e+03        Length:         3492
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AAD54615 (1-3492)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db       2049 TCTCTCATCACATGGAATCTG 2029

RESULT 61
AAF75055/c
ID   AAF75055 standard; DNA; 3556 BP.
XX
AC   AAF75055;
XX
DT   10-MAY-2001  (first entry)
XX
DE   TR16-long receptor DNA.
XX
KW   TR16 recptor; tumour necrosis factor receptor superfamily; apoptosis;
KW   inflammatory; cancer; immune; neurodegenerative; ds.
XX
OS   Unidentified.
XX
PN   WO200112671-A1.
XX
PD   22-FEB-2001.
XX
PF   10-AUG-2000; 2000WO-US021885.
XX
PR   12-AUG-1999;   99US-0148348P.
PR   13-AUG-1999;   99US-0148683P.
PR   13-AUG-1999;   99US-0148870P.
PR   16-AUG-1999;   99US-0148758P.
PR   17-AUG-1999;   99US-0149181P.
PR   18-AUG-1999;   99US-0149453P.
PR   19-AUG-1999;   99US-0149498P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Young PE,  Baker KP;
XX
DR   WPI; 2001-138754/14.
XX
PT   New nucleic acid molecule encoding a TR16 tumor necrosis factor receptor
PT   polypeptide, useful for the diagnosis and treatment of cancer, autoimmune
PT   disorders and cardiovascular diseases.
XX
PS   Disclosure; Fig 4; 286pp; English.
XX
CC   The present invention relates to a TR16 receptor (tumour necrosis factor
CC   receptor superfamily). The invention is useful treating diseases and
CC   disorders associated with the inhibited or increased apoptosis. In
CC   particular inflammatory diseases, cancers, immune and neurodegenerative
CC   disorders may be treated
XX
SQ   Sequence 3556 BP; 1043 A; 751 C; 836 G; 926 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.79e+03        Length:         3556
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AAF75055 (1-3556)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||||||
Db       2890 TCATTACTAACTTGGAATATT 2870

RESULT 62
AAD62893/c
ID   AAD62893 standard; cDNA; 3556 BP.
XX
AC   AAD62893;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human TR16-long receptor cDNA.
XX
KW   Human; TR16; tumour necrosis factor; TNF; antibacterial; virucide;
KW   autoimmune disease; graft-versus-host disease; inflammatory disease;
KW   infection; immunodeficiency; lymphadenopathy; cardiovascular disorder;
KW   gene therapy; TNF-related apoptosis-inducing ligand; TRAIL; cytostatic;
KW   tumour; metastasis; antiparasitic; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .3084
FT                   /*tag=  a
FT                   /product= "Human TR16 receptor protein"
FT   sig_peptide     1. .141
FT                   /*tag=  b
FT   mat_peptide     142. .3081
FT                   /*tag=  c
FT                   /product= "Human mature TR16 receptor protein"
XX
PN   US2003134788-A1.
XX
```

```
PD   17-JUL-2003.
XX
PF   13-FEB-2002; 2002US-00073333.
XX
PR   12-AUG-1999;   99US-0148348P.
PR   13-AUG-1999;   99US-0148683P.
PR   13-AUG-1999;   99US-0148870P.
PR   16-AUG-1999;   99US-0148758P.
PR   17-AUG-1999;   99US-0149181P.
PR   18-AUG-1999;   99US-0149453P.
PR   19-AUG-1999;   99US-0149498P.
PR   10-AUG-2000; 2000US-00637856.
PR   14-FEB-2001; 2001US-0268364P.
XX
PA   (BAKE/) BAKER K P.
PA   (YOUN/) YOUNG P E.
PA   (RUBE/) RUBEN S M.
XX
PI   Baker KP,  Young PE,  Ruben SM;
XX
DR   WPI; 2003-829644/77.
DR   P-PSDB; ABW01590.
XX
PT   New TR16 nucleic acid molecule and polypeptide for diagnosing, preventing
PT   or treating diseases such as cancer, infections by bacteria, viruses and
PT   other parasites, immunodeficiencies, inflammation or cardiovascular
PT   diseases.
XX
PS   Claim 1; Fig 4; 162pp; English.
XX
CC   The invention relates to TR16 protein, a member of tumour necrosis factor
CC   (TNF) receptor superfamily and the TNF-related apoptosis-inducing ligand
CC   (TRAIL) receptor superfamily. TR16 DNA and protein sequences are useful
CC   in diagnosing, preventing or treating diseases and disorders such as
CC   tumours and tumour metastases, infections by bacteria, viruses and other
CC   parasites, autoimmune diseases, graft-versus-host disease, inflammatory
CC   diseases, immunodeficiencies, lymphadenopathy or cardiovascular
CC   disorders. TR16 DNA is also useful in gene therapy. The present sequence
CC   is human TR16-long receptor cDNA
XX
SQ   Sequence 3556 BP; 1043 A; 751 C; 836 G; 926 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.79e+03        Length:          3556
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       10              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AAD62893 (1-3556)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       2890 TCATTACTAACTTGGAATATT 2870

RESULT 63
ADT89650/c
ID   ADT89650 standard; cDNA; 3556 BP.
XX
AC   ADT89650;
XX
DT   13-JAN-2005  (first entry)
XX
DE   Human TR16-long receptor cDNA.
XX
KW   TR16; tumour necrosis factor; TNF; TNF-related apoptosis inducing ligand;
KW   TRAIL; cancer; autoimmune disease; viral infection; inflammation;
KW   graft versus host disease; acute graft rejection;
KW   chronic graft rejection; apoptosis; AIDS; neurodegenerative disorder;
KW   myelodysplastic syndrome; ischaemic injury; toxin-induced liver disease;
KW   septic shock; cachexia; anorexia; gene therapy; cytostatic;
KW   immunosuppressive; virucide; antiinflammatory; anti-HIV; neuroprotective;
KW   vasotropic; hepatotropic; antibacterial; immunomodulator;
KW   eating-disorder; anabolic; human; chromosome 7q21; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .3084
FT                   /*tag=  b
FT                   /product= "Human TR16-long receptor protein"
FT   sig_peptide     1. .141
FT                   /*tag=  a
FT   mat_peptide     142. .3081
FT                   /*tag=  c
FT                   /product= "Human mature TR16-long receptor protein"
XX
PN   US2004209328-A1.
XX
PD   21-OCT-2004.
XX
PF   10-MAY-2004; 2004US-00841471.
XX
PR   12-AUG-1999;   99US-0148348P.
PR   13-AUG-1999;   99US-0148683P.
PR   13-AUG-1999;   99US-0148870P.
PR   16-AUG-1999;   99US-0148758P.
PR   17-AUG-1999;   99US-0149181P.
PR   18-AUG-1999;   99US-0149453P.
PR   19-AUG-1999;   99US-0149498P.
PR   10-AUG-2000; 2000US-00637856.
PR   14-FEB-2001; 2001US-0268364P.
PR   13-FEB-2002; 2002US-00073333.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Baker KP,  Young PE,  Ruben SM;
XX
DR   WPI; 2004-747215/73.
DR   P-PSDB; ADT89651.
XX
PT   New isolated TR16 nucleic acid and polypeptide, useful for preventing or
PT   treating disorders associated with increased cell survival, unregulated
PT   cell proliferation and/or inhibition of apoptosis.
XX
PS   Claim 1; SEQ ID NO 3; 161pp; English.
XX
CC   The invention relates to TR16 protein which is a member of the tumour
CC   necrosis factor (TNF) receptor superfamily and the TNF-related apoptosis
CC   inducing ligand (TRAIL) receptor subfamily and its corresponding nucleic
CC   acid sequence. The sequences of the invention are useful for screening
CC   and identifying agonists and antagonists of TR16 activity which are
CC   useful for preventing or treating disorders associated with increased
CC   cell survival, unregulated cell proliferation or the inhibition of
CC   apoptosis, such as cancer, autoimmune diseases, viral infections,
CC   inflammation, graft versus host disease, acute graft rejection and
CC   chronic graft rejection and diseases associated with increased apoptosis
CC   such as AIDS, neurodegenerative disorders, myelodysplastic syndromes,
CC   ischaemic injury, toxin-induced liver disease, septic shock, cachexia and
CC   anorexia. TR16 DNA is useful in gene therapy. The present sequence is the
CC   human TR16-long receptor cDNA. The human TR16 gene is located at
CC   chromosome 7q21.
XX
SQ   Sequence 3556 BP; 1043 A; 751 C; 836 G; 926 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.79e+03        Length:          3556
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       13              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x ADT89650 (1-3556)
```

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db       2890 TCATTACTAACTTGGAATATT 2870

RESULT 64
ABL59504/c
ID   ABL59504 standard; cDNA; 3758 BP.
XX
AC   ABL59504;
XX
DT   16-JUL-2002  (first entry)
XX
DE   Human diacylglycerol kinase (gamma) cDNA SEQ ID NO:4.
XX
KW   Human; diacylglycerol kinase (gamma); diacylglycerol kinase; enzyme;
KW   tumour; lipid associated gene; lipid metabolism; lipid synthesis;
KW   chromosome 3q27-q28; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200227028-A1.
XX
PD   04-APR-2002.
XX
PF   27-SEP-2001; 2001WO-US030366.
XX
PR   28-SEP-2000; 2000US-00676052.
XX
PA   (ATAI-) ATAIRGIN TECHNOLOGIES INC.
XX
PI   Skinner MK,  Patton JL,  Chaudhary J;
XX
DR   WPI; 2002-405056/43.
XX
PT   Identifying tumor characteristics in a tissue sample taken from a
PT   patient, involves determining the copy number or expression level of
PT   genes associated with lipid metabolism, synthesis or action.
XX
PS   Example 1; Page 65-66; 113pp; English.
XX
CC   The present invention describes a method for identifying tumour
CC   characteristics, comprising measuring a copy number or expression level
CC   of at least two genes associated with lipid metabolism, synthesis, or
CC   action in cells from a patient tissue sample, and comparing the results
CC   with a copy number or expression level of the genes in a normal cell.
CC   Also described is an array of nucleic acid polymers immobilised on a
CC   solid support, comprising a solid support, at least two different nucleic
CC   acid polymers which are each specific for a different gene associated
CC   with lipid metabolism, synthesis or action, where each nucleic acid
CC   polymer is located at a predetermined position on the solid support, and
CC   the array comprises nucleic acid polymers which are specific for less
CC   than 100 genes other than the selected genes. The method is useful for
CC   determining tumour characteristics in a tissue sample taken from a
CC   patient. The present sequence represents a human lipid-associated gene
CC   related cDNA sequence, which is used in the exemplification of the
CC   present invention
XX
SQ   Sequence 3758 BP; 983 A; 980 C; 991 G; 804 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.04e+04      Length:         3758
Score:                  36.00         Matches:        6
Percent Similarity:     100.0%        Conservative:   1
Best Local Similarity:  85.7%         Mismatches:     0
Query Match:            94.7%         Indels:         0
DB:                     6             Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ABL59504 (1-3758)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db       2315 TCTCTCATCACATGGAATCTG 2295

RESULT 65
AAD54613/c
ID   AAD54613 standard; DNA; 3758 BP.
XX
AC   AAD54613;
XX
DT   26-JUN-2003  (first entry)
XX
DE   Human diacylglycerol kinase (DGK) DNA #14.
XX
KW   Diacylglycerol kinase; p53 pathway; diacylglycerol kinase; DAG; cancer;
KW   DGK; gene therapy; cytostatic; enzyme; ds.
XX
OS   Homo sapiens.
XX
PN   WO200299060-A2.
XX
PD   12-DEC-2002.
XX
PF   03-JUN-2002; 2002WO-US017527.
XX
PR   05-JUN-2001; 2001US-0296076P.
PR   10-OCT-2001; 2001US-0328605P.
PR   22-OCT-2001; 2001US-0338733P.
PR   15-FEB-2002; 2002US-0357253P.
PR   15-FEB-2002; 2002US-0357600P.
XX
PA   (EXEL-) EXELIXIS INC.
XX
PI   Friedman L,  Plowman GD,  Belvin M,  Francis-Lang H,  Li D,  Funke RP;
XX
DR   WPI; 2003-201301/19.
XX
PT   Identifying a candidate p53 pathway modulating agent for diagnosing or
PT   treating cancer by detecting a test agent-biased activity of the assay
PT   system comprising a purified DGK (Diacylglycerol kinase) polypeptide or
PT   nucleic acid.
XX
PS   Disclosure; Page 67-69; 111pp; English.
XX
CC   The present invention relates to a method for identifying a candidate p53
CC   pathway modulating agent. The method involves providing an assay system
CC   comprising a purified diacylglycerol (DAG) kinase (DGK) polypeptide or
CC   nucleic acid, contacting the assay system with a test agent and detecting
CC   the test agent-biased activity of the assay system where a difference
CC   between the test agent-biased activity and reference activity identifies
CC   the agent as a candidate p53 pathway modulating agent. The method is
CC   useful for manufacturing a medicament for diagnosing or treating breast,
CC   colon, lung or ovary cancer. The invention is also useful in gene
CC   therapy. The present sequence is human DGK DNA
XX
SQ   Sequence 3758 BP; 983 A; 980 C; 991 G; 804 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.04e+04      Length:         3758
Score:                  36.00         Matches:        6
Percent Similarity:     100.0%        Conservative:   1
Best Local Similarity:  85.7%         Mismatches:     0
Query Match:            94.7%         Indels:         0
DB:                     8             Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AAD54613 (1-3758)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db       2315 TCTCTCATCACATGGAATCTG 2295

RESULT 66
AAD54614/c
ID   AAD54614 standard; DNA; 3758 BP.
XX
AC   AAD54614;
XX
```

DT 26-JUN-2003 (first entry)
XX
DE Human diacylglycerol kinase (DGK) DNA #15.
XX
KW Diacylglycerol kinase; p53 pathway; diacylglycerol kinase; DAG; cancer;
KW DGK; gene therapy; cytostatic; enzyme; ds.
XX
OS Homo sapiens.
XX
PN WO200299060-A2.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017527.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 22-OCT-2001; 2001US-0338733P.
PR 15-FEB-2002; 2002US-0357253P.
PR 15-FEB-2002; 2002US-0357600P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
DR WPI; 2003-201301/19.
XX
PT Identifying a candidate p53 pathway modulating agent for diagnosing or
PT treating cancer by detecting a test agent-biased activity of the assay
PT system comprising a purified DGK (Diacylglycerol kinase) polypeptide or
PT nucleic acid.
XX
PS Disclosure; Page 70-72; 111pp; English.
XX
CC The present invention relates to a method for identifying a candidate p53
CC pathway modulating agent. The method involves providing an assay system
CC comprising a purified diacylglycerol (DAG) kinase (DGK) polypeptide or
CC nucleic acid, contacting the assay system with a test agent and detecting
CC the test agent-biased activity of the assay system where a difference
CC between the test agent-biased activity and reference activity identifies
CC the agent as a candidate p53 pathway modulating agent. The method is
CC useful for manufacturing a medicament for diagnosing or treating breast,
CC colon, lung or ovary cancer. The invention is also useful in gene
CC therapy. The present sequence is human DGK DNA
XX
SQ Sequence 3758 BP; 983 A; 980 C; 991 G; 804 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.04e+04	Length:	3758
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAD54614 (1-3758)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db       2315 TCTCTCATCACATGGAATCTG 2295
```

RESULT 67
ADE31700/c
ID ADE31700 standard; DNA; 3758 BP.
XX
AC ADE31700;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human 3484 gene #SEQ ID 57.
XX
KW Antiarteriosclerotic; cardiant; vasotropic; antiinflammatory;
KW thrombolytic; antiarrhythmic; antianginal; hypotensive; gene therapy;
KW cardiovascular; disorder; ischaemia; aortic bending;
KW vascular heart disease; endocarditis; atrial fibrillation; heart failure;
KW angina; cardiomyopathy; cardiac death; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003065984-A2.
XX
PD 14-AUG-2003.
XX
PF 29-JAN-2003; 2003WO-US002571.
XX
PR 01-FEB-2002; 2002US-0353224P.
PR 15-MAR-2002; 2002US-0364529P.
PR 19-APR-2002; 2002US-0373861P.
PR 29-APR-2002; 2002US-0376287P.
PR 12-JUN-2002; 2002US-0388080P.
PR 24-JUN-2002; 2002US-0390971P.
PR 03-JUL-2002; 2002US-0394130P.
PR 10-JUL-2002; 2002US-0394797P.
PR 21-AUG-2002; 2002US-0404904P.
PR 23-AUG-2002; 2002US-0405450P.
PR 04-SEP-2002; 2002US-0408070P.
PR 06-NOV-2002; 2002US-0424300P.
PR 05-DEC-2002; 2002US-0431042P.
PR 05-DEC-2002; 2002US-0431079P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Logan TJ, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M;
PI Stagliano N, Perodin J, Rodrigue-Way A;
XX
DR WPI; 2003-731468/69.
DR P-PSDB; ADE31701.
XX
PT Identifying a compound capable of treating a cardiovascular disorder
PT (e.g. atherosclerosis) comprises assaying the ability of the compound to
PT modulate the expression or activity of e.g. 1682, 6169 or 6193
PT polypeptide or nucleic acid.
XX
PS Disclosure; SEQ ID NO 57; 328pp; English.
XX
CC The invention relates to a method for identifying a compound capable of
CC treating a cardiovascular disorder. The present invention identifies the
CC differential expression of 1682, 6169, 6193, 7771, 14395, 29002, 33216,
CC 43726, 69292, 21656, 32427, 2402, 7747, 1720, 9151, 60491, 1371, 7077,
CC 33207, 1419, 18036, 16105, 38650, 14245, 58848, 1870, 25856, 32394, 3484,
CC 345, 9252, 9135, 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912,
CC 2868, 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914, 51130,
CC 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474, 9792, 15400, 1452 or
CC 6585 genes in cardiovascular disease states. The methods are useful in
CC diagnosing, preventing and treating cardiovascular disorders, such as
CC atherosclerosis, cardiac hypertrophy, ischaemia, reperfusion injury,
CC restenosis, arterial inflammation, vascular wall remodeling, coronary
CC microembolism, tachycardia, bradycardia, pressure overload, aortic
CC bending, coronary artery ligation, vascular heart disease, valvular
CC disease, including but not limited to, valvular degeneration caused by
CC calcification, rheumatic heart disease, endocarditis, or complications of
CC artificial valves; atrial fibrillation, long-QT syndrome, congestive
CC heart failure, sinus node dysfunction, angina, heart failure,
CC hypertension, atrial fibrillation, atrial flutter, pericardial disease,
CC including but not limited to, pericardial effusion and pericarditis;
CC cardiomyopathies, e.g. dilated cardiomyopathy or idiopathic
CC cardiomyopathy, myocardial infarction, coronary artery disease, coronary
CC artery spasm, ischaemic disease, arrhythmia, sudden cardiac death and
CC cardiovascular developmental disorders. The methods may also be used for
CC identifying compounds that modulate cardiovascular disorders. Sequences
CC given in ADE31644-ADE31769 represent the genes and proteins that may be
CC regulated by a compound of the invention.
XX
SQ Sequence 3758 BP; 983 A; 980 C; 991 G; 804 T; 0 U; 0 Other;

Alignment Scores:

```
Pred. No.:              1.04e+04        Length:         3758
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     10              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ADE31700 (1-3758)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2315 TCTCTCATCACATGGAATCTG 2295

RESULT 68
ABX74401/c
ID   ABX74401 standard; cDNA; 3759 BP.
XX
AC   ABX74401;
XX
DT   21-MAR-2003  (first entry)
XX
DE   Human cDNA sequence #12 up-regulated in non-aggressive CC-RCC.
XX
KW   Human; microarray; solid surface; immobilised probe; CC-RCC;
KW   differential expression profile; aggressive CC-RCC tumour type;
KW   non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
KW   gene expression profiling; tumour tissue; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200279411-A2.
XX
PD   10-OCT-2002.
XX
PF   29-MAR-2002; 2002WO-US009576.
XX
PR   29-MAR-2001; 2001US-0279411P.
XX
PA   (VAND-) VAN ANDEL INST.
XX
PI   Haab B,  Rhodes D,  Teh BT,  Takashi M;
XX
DR   WPI; 2003-040679/03.
XX
PT   New microarray, comprising a matrix of cDNA probe from a set of probes
PT   immobilized to a solid surface in predetermined order, useful in the
PT   prognosis of patients with clear cell renal carcinoma.
XX
PS   Claim 1; Page 85-86; 179pp; English.
XX
CC   The present invention relates to a microarray comprising a matrix of at
CC   least one cDNA probe from a set of probes immobilised to a solid surface
CC   in a predetermined order, where a row of pixels corresponds to replicates
CC   of one distinct probe from the set. The probes are complementary to
CC   nucleic acid sequences that are expressed differentially in aggressive as
CC   compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)
CC   and that hybridise to the probes under high stringency conditions. The
CC   microarray is useful for the prognosis of patients with CC-RCC, wherein
CC   aggressive and non-aggressive CC-RCC tumour types are characterised by
CC   differential expression profiles of genes that hybridise with one or more
CC   probes immobilised on the microarray. The arrays are useful for gene
CC   expression profiling of tumour and normal tissues. The present sequence
CC   represents a human cDNA sequence up-regulated in non-aggressive CC-RCC
CC   phenotypes
XX
SQ   Sequence 3759 BP; 983 A; 983 C; 990 G; 803 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.04e+04        Length:         3759
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     10              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ABX74401 (1-3759)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2316 TCTCTCATCACATGGAATCTG 2296

RESULT 69
AAS64300
ID   AAS64300 standard; cDNA; 3900 BP.
XX
AC   AAS64300;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #104.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG00113.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 104; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3900 BP; 1161 A; 798 C; 892 G; 1049 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.09e+04        Length:         3900
Score:                  36.00           Matches:        5
```

```
Percent Similarity:     100.0%          Conservative:  2
Best Local Similarity:  71.4%           Mismatches:    0
Query Match:            94.7%           Indels:        0
DB:                     5               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x AAS64300 (1-3900)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||:::
Db       2351 AGCTTAGTAACTTGGAATGTT 2371

RESULT 70
ADS98356
ID   ADS98356 standard; DNA; 3900 BP.
XX
AC   ADS98356;
XX
DT   30-DEC-2004  (first entry)
XX
DE   Protein factor discovery related isolated human contig DNA, SEQ ID 620.
XX
KW   antiinflammatory; cytostatic; antimicrobial; gene therapy; inflammation;
KW   leukaemia; nervous system disorder; infection; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004087874-A2.
XX
PD   14-OCT-2004.
XX
PF   24-MAR-2004; 2004WO-US009202.
XX
PR   28-MAR-2003; 2003US-0458824P.
XX
PA   (NUVE-) NUVELO INC.
PA   (DRMA/) DRMANAC R T.
XX
PI   Tang YT,  Zhou P,  Wang J,  Wang ZW,  Hu T;
XX
DR   WPI; 2004-737686/72.
DR   P-PSDB; ADS98696.
XX
PT   New polynucleotides encoding a polypeptide with biological activity,
PT   useful for treating inflammation, leukemias, nervous system disorders, or
PT   infections.
XX
PS   Example 2; SEQ ID NO 620; 253pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide comprising any
CC   of the 235 nucleotide sequences described in the specification. The
CC   invention further comprises: an isolated polynucleotide encoding a
CC   polypeptide with biological activity, where the polynucleotide hybridizes
CC   to one of the 235 novel polynucleotides under stringent hybridization
CC   conditions, or having greater than about 99% sequence identity with the
CC   novel polynucleotide; a vector comprising a novel polynucleotide; an
CC   expression vector comprising the novel polynucleotide; a host cell
CC   genetically engineered to comprise the novel polynucleotide, which can be
CC   operatively associated with a regulatory sequence that modulates
CC   expression of the polynucleotide in the host cell; an isolated
CC   polypeptide encoded by the novel polynucleotide, or a polynucleotide
CC   hybridizing under stringent conditions to the novel polynucleotide; a
CC   composition comprising the polypeptide and a carrier; an antibody
CC   directed against the polypeptide; a method for detecting the novel
CC   polynucleotide in a sample; a method for detecting the polypeptide in a
CC   sample; a method for identifying a compound that binds to the polypeptide
CC   ; a method for producing the polypeptide; an isolated polypeptide
CC   comprising any of the 235 amino acid sequences described in the
CC   specification; and a collection of polynucleotides comprising of at least
CC   one of the polynucleotides cited above. The polypeptides and
CC   polynucleotides of the invention have antiinflammatory, cytostatic, and
CC   antimicrobial activities. The novel polynucleotide may be used to treat
CC   disorders by gene therapy. The polypeptides and polynucleotides are
CC   useful for treating inflammation, leukaemias, nervous system disorders,
CC   or infections. This polynucleotide sequence represents contiguous DNA
CC   derived from one of the 235 novel isolated polynucleotides of the
CC   invention.
XX
SQ   Sequence 3900 BP; 1161 A; 798 C; 892 G; 1049 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.09e+04        Length:        3900
Score:                  36.00           Matches:       5
Percent Similarity:     100.0%          Conservative:  2
Best Local Similarity:  71.4%           Mismatches:    0
Query Match:            94.7%           Indels:        0
DB:                     13              Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x ADS98356 (1-3900)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||:::
Db       2351 AGCTTAGTAACTTGGAATGTT 2371

RESULT 71
ABQ93362
ID   ABQ93362 standard; cDNA; 3962 BP.
XX
AC   ABQ93362;
XX
DT   14-OCT-2002  (first entry)
XX
DE   Human cDNA SEQ ID NO 75.
XX
KW   Human; vulnerary; dermatological; neuroprotective; nootropic; cancer;
KW   antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
KW   antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
KW   burn; central nervous system disorder; Alzheimer's disease;
KW   Parkinson's disease; Huntington's disease; immune disorder;
KW   autoimmune disorder; multiple sclerosis; diabetes; allergy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200218424-A2.
XX
PD   07-MAR-2002.
XX
PF   31-AUG-2001; 2001WO-US027093.
XX
PR   01-SEP-2000; 2000US-00654935.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Asundi V,  Zhou P,  Xue AJ,  Ren F,  Zhang J,  Wang J;
PI   Zhao QA,  Wang D,  Liu C,  Drmanac RT,  Wehrman T;
XX
DR   WPI; 2002-583321/62.
DR   P-PSDB; ABP62883.
XX
PT   New polynucleotide and polypeptides, useful for treatment and diagnosis
PT   of Alzheimer's, Parkinson's, Huntington's, amylotrophic lateral
PT   sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
PT   sclerosis, diabetes and allergies.
XX
PS   Claim 22; SEQ ID NO 75; 284pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (I) comprising one of
CC   245 sequences (ABQ93288-ABQ93532). Treating a condition comprising
CC   administering to a mammalian subject a composition comprising the protein
CC   (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
CC   (I), (II) and (III) are useful for diagnostic evaluation of disorders.
CC   (I) is useful for gene therapy of diseases and (II) can be used for
CC   therapeutic treatment. Diseases that may be treated include wound healing
CC   and tissue repair, burns, central nervous system disorders (e.g.
CC   Alzheimer's, Parkinson's, Huntington's and amylotrophic lateral
CC   sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
CC   sclerosis, diabetes and allergies. Note: The sequence data for this
```

```
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3962 BP; 1191 A; 801 C; 899 G; 1071 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.11e+04        Length:         3962
Score:                  36.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ABQ93362 (1-3962)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db       2351 AGCTTAGTAACTTGGAATGTT 2371

RESULT 72
ADS98424
ID   ADS98424 standard; DNA; 3962 BP.
XX
AC   ADS98424;
XX
DT   30-DEC-2004  (first entry)
XX
DE   Protein factor discovery related isolated human contig DNA, SEQ ID 688.
XX
KW   antiinflammatory; cytostatic; antimicrobial; gene therapy; inflammation;
KW   leukaemia; nervous system disorder; infection; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004087874-A2.
XX
PD   14-OCT-2004.
XX
PF   24-MAR-2004; 2004WO-US009202.
XX
PR   28-MAR-2003; 2003US-0458824P.
XX
PA   (NUVE-) NUVELO INC.
PA   (DRMA/) DRMANAC R T.
XX
PI   Tang YT,  Zhou P,  Wang J,  Wang ZW,  Hu T;
XX
DR   WPI; 2004-737686/72.
DR   P-PSDB; ADS98764.
XX
PT   New polynucleotides encoding a polypeptide with biological activity,
PT   useful for treating inflammation, leukemias, nervous system disorders, or
PT   infections.
XX
PS   Example 2; SEQ ID NO 688; 253pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide comprising any
CC   of the 235 nucleotide sequences described in the specification. The
CC   invention further comprises: an isolated polynucleotide encoding a
CC   polypeptide with biological activity, where the polynucleotide hybridizes
CC   to one of the 235 novel polynucleotides under stringent hybridization
CC   conditions, or having greater than about 99% sequence identity with the
CC   novel polynucleotide; a vector comprising a novel polynucleotide; an
CC   expression vector comprising the novel polynucleotide; a host cell
CC   genetically engineered to comprise the novel polynucleotide, which can be
CC   operatively associated with a regulatory sequence that modulates
CC   expression of the polynucleotide in the host cell; an isolated
CC   polypeptide encoded by the novel polynucleotide, or a polynucleotide
CC   hybridizing under stringent conditions to the novel polynucleotide; a
CC   composition comprising the polypeptide and a carrier; an antibody
CC   directed against the polypeptide; a method for detecting the novel
CC   polynucleotide in a sample; a method for detecting the polypeptide in a
CC   sample; a method for identifying a compound that binds to the polypeptide
CC   ; a method for producing the polypeptide; an isolated polypeptide
CC   comprising any of the 235 amino acid sequences described in the
CC   specification; and a collection of polynucleotides comprising of at least
CC   one of the polynucleotides cited above. The polypeptides and
CC   polynucleotides of the invention have antiinflammatory, cytostatic, and
CC   antimicrobial activities. The novel polynucleotide may be used to treat
CC   disorders by gene therapy. The polypeptides and polynucleotides are
CC   useful for treating inflammation, leukaemias, nervous system disorders,
CC   or infections. This polynucleotide sequence represents contiguous DNA
CC   derived from one of the 235 novel isolated polynucleotides of the
CC   invention.
XX
SQ   Sequence 3962 BP; 1191 A; 801 C; 899 G; 1071 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.11e+04        Length:         3962
Score:                  36.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     13              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x ADS98424 (1-3962)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db       2351 AGCTTAGTAACTTGGAATGTT 2371

RESULT 73
AAT93393/c
ID   AAT93393 standard; cDNA; 4438 BP.
XX
AC   AAT93393;
XX
DT   27-JAN-1998  (first entry)
XX
DE   Bloom's syndrome BLM mutated gene (6 bp deletion/7 bp insertion).
XX
KW   BLM; Bloom's syndrome; BS; mutant; mutation; deletion; insertion;
KW   frameshift; truncation; therapy; diagnosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             75. .2294
FT                   /*tag=  a
FT                   /note= "encodes a truncated protein"
FT   mutation        2281. .2287
FT                   /*tag=  b
FT                   /note= "Deletion of ATCTGA from the wild type H1-5' BLM
FT                   gene and insertion of TAGATTC results in the insertion of
FT                   the novel codons for LDSR after amino acid position 735
FT                   followed by a stop codon"
XX
PN   WO9717979-A1.
XX
PD   22-MAY-1997.
XX
PF   15-NOV-1996;   96WO-US019046.
XX
PR   15-NOV-1995;   95US-00559303.
XX
PA   (NYBL-) NEW YORK BLOOD CENT INC.
XX
PI   Ellis N,  German J,  Groden J;
XX
DR   WPI; 1997-289051/26.
DR   P-PSDB; AAW31549.
XX
PT   Diagnosing Bloom's syndrome, and carriers, by detecting mutant BLM genes
PT   - for gene therapy with nucleic acid encoding active BLM protein to treat
PT   Bloom's syndrome and cancer in general.
```

XX
PS Disclosure; Page; 51pp; English.
XX
CC This is a mutated BLM gene sequence isolated from four Ashkenazi Jewish
CC individuals homozygous for Bloom's Syndrome. The deletion of ATCTGA from
CC the wild type H1-5' BLM gene and insertion of TAGATTC (nucleotide
CC positions 2281 to 2287 of the present sequence) results in the insertion
CC of the novel codons for LDSR after amino acid position 735 followed by a
CC stop codon. This is one of the seven unique mutations which were
CC identified in a study of 10 people with Bloom's Syndrome. Based on the
CC various mutations, diagnostic tests for Bloom's Syndrome have been
CC developed which use standard sequence analysis techniques to detect the
CC presence of 2 mutated BLM genes or the absence of a wild-type BLM gene.
CC Note: The present sequence does not appear in the specification; it has
CC been made by modifying the H1-5' wild type BLM sequence which is provided
CC in Figure 2 (AAT67013)
XX
SQ Sequence 4438 BP; 1465 A; 867 C; 911 G; 1195 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.26e+04	Length:	4438
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AAT93393 (1-4438)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2301 AGTCTTATCACCTGGAATCTA 2281
```

RESULT 74
AAK91262/c
ID AAK91262 standard; DNA; 6499 BP.
XX
AC AAK91262;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 4838.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001324.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.

```
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-502630/55.
XX
PT   Polynucleotides encoding digestive system antigens, useful for
PT   diagnosing, treating, preventing and/or prognozing disorders of the
PT   digestive system, particularly cancer and cancer metastases.
XX
PS   Disclosure; SEQ ID NO 4838; 986pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human digestive system antigens. These can be used in the
CC   diagnosis, treatment and prevention of digestive system disorders,
CC   including cancer, Meckel's diverticulum, bacterial or parasitic
CC   infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC   ulcerative colitis. The present sequence is a genomic DNA fragment
CC   encoding a digestive system antigen of the invention
XX
SQ   Sequence 6499 BP; 1933 A; 1117 C; 1190 G; 2259 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.93e+04        Length:          6499
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       4               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AAK91262 (1-6499)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db         33 AGCATTATTACGTGGAACATC 13

RESULT 75
AEA15412/c
ID   AEA15412 standard; cDNA; 11529 BP.
XX
AC   AEA15412;
XX
DT   28-JUL-2005  (first entry)
XX
DE   Human polynucleotide #109.
XX
KW   Diagnosis; prognosis; cancer; breast tumor; ovary tumor; stomach tumor;
KW   colon tumor; esophagus tumor; bladder tumor; non-small-cell lung cancer;
KW   cytostatic; neoplasm; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2005047534-A2.
XX
PD   26-MAY-2005.
XX
PF   15-OCT-2004; 2004WO-EP011599.
XX
PR   28-OCT-2003; 2003EP-00024565.
XX
PA   (FARB ) BAYER HEALTHCARE AG.
XX
PI   Wirtz R,  Munnes M;
XX
DR   WPI; 2005-372393/38.
DR   P-PSDB; AEA15490.
XX
PT   Predicting a response to cancer treatment by detecting at least 2
PT   markers, which are genes or genomic nucleic acid sequences that are
PT   located on one chromosomal region, which is altered in malignant
PT   neoplasia.
XX
PS   Claim 7; SEQ ID NO 374; 464pp; English.
XX
CC   The invention relates to a method of predicting response to cancer
CC   treatment comprising detection of at least 2 markers, where the markers
CC   are genes and fragments or genomic nucleic acid sequences that are
CC   located on one chromosomal region, which is altered in malignant
CC   neoplasia. The invention also relates to a method for the prediction,
CC   diagnosis or prognosis of malignant neoplasia, methods for detecting
CC   deregulations in malignant neoplasia and breast cancer, a method of
CC   determining the phenotype of a cell or tissue, a method for identifying
CC   genomic regions which are altered on the chromosomal level and encode
CC   genes that are linked by function and are differentially expressed in
CC   malignant neoplasia and breast cancer, methods of screening for agents
CC   which regulate the activity of a polypeptide or a polynucleotide and
CC   antibodies that specifically bind to a full length or partial
CC   polypeptide. The method is useful for predicting response to cancer
CC   treatment. The methods and compositions are useful for predicting,
CC   diagnosing, prognosing, preventing or treating malignant neoplasia
CC   including breast cancer, ovarian cancer, gastric cancer, colon cancer,
CC   esophageal cancer, mesenchymal cancer, bladder cancer or non-small-cell
CC   lung cancer. This sequence represents a human polynucleotide used in the
CC   scope of the invention.
XX
SQ   Sequence 11529 BP; 3180 A; 2801 C; 2936 G; 2612 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.66e+04        Length:          11529
Score:                    36.00           Matches:         5
```

Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	14	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AEA15412 (1-11529)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||:::
Db       9172 TCCTTAGTGACATGGAATGTG 9152
```

Search completed: March 10, 2006, 15:26:24
Job time : 44.087 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 20:06:44 ; Search time 7.9606 Seconds
(without alignments)
1563.065 Million cell updates/sec

Title: US-10-757-745-2_COPY_115_121
Perfect score: 38
Sequence: 1 SLITWNI 7

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170331_4111/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_417_@runat_09032006_170331_4111 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	38	100.0	1920	3	US-09-697-863A-1	Sequence 1, Appli
	2	38	100.0	1948	3	US-09-620-312D-889	Sequence 889, App
c	3	36	94.7	326	3	US-09-513-999C-35251	Sequence 35251, A
	4	36	94.7	601	3	US-09-949-016-67704	Sequence 67704, A
c	5	36	94.7	601	3	US-09-949-016-86285	Sequence 86285, A
c	6	36	94.7	601	3	US-09-949-016-86286	Sequence 86286, A
c	7	36	94.7	601	3	US-09-949-016-86287	Sequence 86287, A
c	8	36	94.7	601	3	US-09-949-016-86288	Sequence 86288, A
c	9	36	94.7	601	3	US-09-949-016-86289	Sequence 86289, A
c	10	36	94.7	601	3	US-09-949-016-86290	Sequence 86290, A
	11	36	94.7	601	3	US-09-949-016-177996	Sequence 177996,
c	12	36	94.7	2087	3	US-10-104-047-84	Sequence 84, Appl
	13	36	94.7	2303	3	US-10-104-047-1924	Sequence 1924, Ap
c	14	36	94.7	39344	3	US-09-973-278-834	Sequence 834, App
	15	36	94.7	70088	3	US-09-949-016-16845	Sequence 16845, A
	16	36	94.7	83178	3	US-09-949-016-14606	Sequence 14606, A
c	17	36	94.7	163181	3	US-09-949-016-13730	Sequence 13730, A
	18	36	94.7	422592	3	US-09-949-016-14182	Sequence 14182, A
c	19	35	92.1	235	3	US-09-513-999C-14281	Sequence 14281, A
	20	35	92.1	601	3	US-09-949-016-54230	Sequence 54230, A
	21	35	92.1	601	3	US-09-949-016-54231	Sequence 54231, A
	22	35	92.1	1297	2	US-08-370-975B-12	Sequence 12, Appl
	23	35	92.1	26764	2	US-08-370-975B-1	Sequence 1, Appli
c	24	35	92.1	68452	3	US-09-949-016-13305	Sequence 13305, A
c	25	35	92.1	95890	3	US-09-949-016-16412	Sequence 16412, A
	26	35	92.1	246240	2	US-08-724-394A-20	Sequence 20, Appl
	27	35	92.1	246240	2	US-08-724-394A-21	Sequence 21, Appl
	28	35	92.1	246240	2	US-08-724-394A-22	Sequence 22, Appl
c	29	35	92.1	325791	3	US-09-768-185A-1	Sequence 1, Appli
c	30	35	92.1	1664976	3	US-08-916-421B-1	Sequence 1, Appli
c	31	35	92.1	1664976	3	US-09-692-570-1	Sequence 1, Appli
	32	34	89.5	343	3	US-09-614-221A-23	Sequence 23, Appl
	33	34	89.5	601	3	US-09-949-016-20188	Sequence 20188, A
	34	34	89.5	601	3	US-09-949-016-20189	Sequence 20189, A
	35	34	89.5	601	3	US-09-949-016-20190	Sequence 20190, A
c	36	34	89.5	601	3	US-09-949-016-44629	Sequence 44629, A
	37	34	89.5	601	3	US-09-949-016-108625	Sequence 108625,
	38	34	89.5	601	3	US-09-949-016-108626	Sequence 108626,
	39	34	89.5	601	3	US-09-949-016-129014	Sequence 129014,
	40	34	89.5	601	3	US-09-949-016-165434	Sequence 165434,
	41	34	89.5	601	3	US-09-949-016-165435	Sequence 165435,
	42	34	89.5	601	3	US-09-949-016-165436	Sequence 165436,
	43	34	89.5	601	3	US-09-949-016-165491	Sequence 165491,
	44	34	89.5	601	3	US-09-949-016-165492	Sequence 165492,
	45	34	89.5	601	3	US-09-949-016-165493	Sequence 165493,
	46	34	89.5	601	3	US-09-949-016-196355	Sequence 196355,
	47	34	89.5	601	3	US-09-949-016-196356	Sequence 196356,
c	48	34	89.5	912	3	US-09-328-352-1038	Sequence 1038, Ap
	49	34	89.5	1312	3	US-09-697-863A-3	Sequence 3, Appli
c	50	34	89.5	1476	3	US-09-583-110-2462	Sequence 2462, Ap
c	51	34	89.5	1488	3	US-09-107-433-2350	Sequence 2350, Ap
c	52	34	89.5	1518	3	US-09-583-110-518	Sequence 518, App
c	53	34	89.5	1518	3	US-09-769-787-334	Sequence 334, App
c	54	34	89.5	1536	3	US-09-107-433-1911	Sequence 1911, Ap
	55	34	89.5	1766	3	US-09-182-145-9	Sequence 9, Appli
c	56	34	89.5	1766	3	US-09-182-145-10	Sequence 10, Appl
c	57	34	89.5	1908	3	US-09-107-532A-3481	Sequence 3481, Ap
c	58	34	89.5	3422	3	US-09-586-719-26	Sequence 26, Appl
c	59	34	89.5	5293	3	US-09-900-708-1	Sequence 1, Appli
c	60	34	89.5	10698	3	US-09-949-016-12780	Sequence 12780, A
c	61	34	89.5	14872	3	US-08-961-527-72	Sequence 72, Appl
c	62	34	89.5	26385	3	US-08-961-527-3	Sequence 3, Appli
c	63	34	89.5	29653	3	US-09-949-016-13442	Sequence 13442, A
c	64	34	89.5	33155	3	US-09-949-016-12704	Sequence 12704, A
c	65	34	89.5	33157	3	US-09-949-016-15756	Sequence 15756, A
c	66	34	89.5	36643	3	US-09-949-016-11860	Sequence 11860, A
c	67	34	89.5	36821	3	US-09-949-016-16403	Sequence 16403, A
c	68	34	89.5	36821	3	US-09-949-016-16404	Sequence 16404, A
c	69	34	89.5	37304	3	US-09-949-016-11862	Sequence 11862, A
c	70	34	89.5	51273	3	US-09-949-016-13018	Sequence 13018, A
c	71	34	89.5	66164	3	US-09-949-016-13574	Sequence 13574, A
	72	34	89.5	92139	3	US-09-918-686-1	Sequence 1, Appli
	73	34	89.5	126029	3	US-09-949-016-14731	Sequence 14731, A
c	74	34	89.5	152070	3	US-09-949-016-15402	Sequence 15402, A
	75	34	89.5	451924	3	US-09-949-016-12896	Sequence 12896, A

ALIGNMENTS

RESULT 1
US-09-697-863A-1
; Sequence 1, Application US/09697863A

; Patent No. 6812203
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/09/697,863A
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1627)..(1627)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1108)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-09-697-863A-1

Alignment Scores:

Pred. No.:	493	Length:	1920
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-697-863A-1 (1-1920)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        362 TCTCTCATTACCTGGAATATT 382
```

RESULT 2
US-09-620-312D-889
; Sequence 889, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-09-620-312D-889

Alignment Scores:

Pred. No.:	500	Length:	1948
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-620-312D-889 (1-1948)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        381 TCTCTCATTACCTGGAATATT 401
```

RESULT 3
US-09-513-999C-35251/c
; Sequence 35251, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 35251
; LENGTH: 326
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 233
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 285
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-35251

Alignment Scores:

Pred. No.:	172	Length:	326
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-513-999C-35251 (1-326)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        140 TCTCTCATCACATGGAATCTG 120
```

RESULT 4
US-09-949-016-67704
; Sequence 67704, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 67704
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-67704

Alignment Scores:

Pred. No.:	336	Length:	601
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-67704 (1-601)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||||:::
Db        313 TCATTAGTGACTTGGAATGTT 333
```

RESULT 5
US-09-949-016-86285/c
; Sequence 86285, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 86285
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-86285

Alignment Scores:

Pred. No.:	336	Length:	601
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-86285 (1-601)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db         27 AGTTTAATTACCTGGAACTTA 7
```

RESULT 6
US-09-949-016-86286/c
; Sequence 86286, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 86286
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-86286

Alignment Scores:

Pred. No.:	336	Length:	601
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-86286 (1-601)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        384 AGTTTAATTACCTGGAACTTA 364
```

RESULT 7
US-09-949-016-86287/c
; Sequence 86287, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 86287
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-86287

Alignment Scores:

Pred. No.:	336	Length:	601
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1

```
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            94.7%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-86287 (1-601)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db        493 AGTTTAATTACCTGGAACTTA 473

RESULT 8
US-09-949-016-86288/c
; Sequence 86288, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 86288
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-86288

Alignment Scores:
Pred. No.:              336            Length:         601
Score:                  36.00          Matches:        6
Percent Similarity:     100.0%         Conservative:   1
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            94.7%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-86288 (1-601)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db        570 AGTTTAATTACCTGGAACTTA 550

RESULT 9
US-09-949-016-86289/c
; Sequence 86289, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 86289
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-86289

Alignment Scores:
Pred. No.:              336            Length:         601
Score:                  36.00          Matches:        6
Percent Similarity:     100.0%         Conservative:   1
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            94.7%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-86289 (1-601)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db        573 AGTTTAATTACCTGGAACTTA 553

RESULT 10
US-09-949-016-86290/c
; Sequence 86290, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 86290
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-86290

Alignment Scores:
Pred. No.:              336            Length:         601
Score:                  36.00          Matches:        6
Percent Similarity:     100.0%         Conservative:   1
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            94.7%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-86290 (1-601)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db        575 AGTTTAATTACCTGGAACTTA 555

RESULT 11
US-09-949-016-177996
; Sequence 177996, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

; SEQ ID NO 177996
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-177996

Alignment Scores:

Pred. No.:	336	Length:	601
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-177996 (1-601)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db         20 TCTTTAGTGACATGGAATGTA 40
```

RESULT 12
US-10-104-047-84/c
; Sequence 84, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-84

Alignment Scores:

Pred. No.:	1.3e+03	Length:	2087
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-104-047-84 (1-2087)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||||
Db        164 AGCATTATTACGTGGAACATC 144
```

RESULT 13
US-10-104-047-1924
; Sequence 1924, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1924
; LENGTH: 2303
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1924

Alignment Scores:

Pred. No.:	1.45e+03	Length:	2303
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-104-047-1924 (1-2303)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        993 AGCTTAGTAACTTGGAATGTT 1013
```

RESULT 14
US-09-973-278-834/c
; Sequence 834, Application US/09973278
; Patent No. 6924354
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: PZ010P2
; CURRENT APPLICATION NUMBER: US/09/973,278
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 60/239,899
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/227,357
; PRIOR FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: PCT/US98/13684
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/051,926
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,793
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,925
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,929
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,803
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,732
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,931
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,932
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,916
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,930
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,918
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,920
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,733
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,795
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,928
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,722
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,723
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,948
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,949
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,953
; PRIOR FILING DATE: 1997-08-18

```
;  PRIOR APPLICATION NUMBER: 60/055,950
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,947
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,964
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/056,360
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,684
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,984
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,954
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/058,785
;  PRIOR FILING DATE: 1997-09-12
;  PRIOR APPLICATION NUMBER: 60/058,664
;  PRIOR FILING DATE: 1997-09-12
;  PRIOR APPLICATION NUMBER: 60/058,660
;  PRIOR FILING DATE: 1997-09-12
;  PRIOR APPLICATION NUMBER: 60/058,661
;  PRIOR FILING DATE: 1997-09-12
;  NUMBER OF SEQ ID NOS: 947
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 834
;    LENGTH: 39344
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-973-278-834

Alignment Scores:
Pred. No.:                3.17e+04        Length:         39344
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-973-278-834 (1-39344)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||
Db      27408 TCATTACTAACTTGGAATATT 27388

RESULT 15
US-09-949-016-16845
; Sequence 16845, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16845
;    LENGTH: 70088
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-16845

Alignment Scores:
Pred. No.:                5.92e+04        Length:         70088
Score:                    36.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-16845 (1-70088)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db      39183 TCTTTAGTGACATGGAATGTA 39203

RESULT 16
US-09-949-016-14606
; Sequence 14606, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14606
;    LENGTH: 83178
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-14606

Alignment Scores:
Pred. No.:                7.12e+04        Length:         83178
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-14606 (1-83178)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db      11378 AGTTTAATAACCTGGAATTTA 11398

RESULT 17
US-09-949-016-13730/c
; Sequence 13730, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13730
;    LENGTH: 163181
;    TYPE: DNA
;    ORGANISM: Human
;    FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(163181)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13730

Alignment Scores:
Pred. No.:              1.46e+05        Length:         163181
Score:                  36.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-13730 (1-163181)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db      93030 TCATTAGTGACTTGGAATGTT 93010

RESULT 18
US-09-949-016-14182
; Sequence 14182, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14182
;   LENGTH: 422592
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(422592)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14182

Alignment Scores:
Pred. No.:              3.91e+05        Length:         422592
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-14182 (1-422592)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db     317895 AGTTTAATTACCTGGAACTTA 317915

RESULT 19
US-09-513-999C-14281/c
; Sequence 14281, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 14281
;   LENGTH: 235
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-513-999C-14281

Alignment Scores:
Pred. No.:              187             Length:         235
Score:                  35.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-513-999C-14281 (1-235)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||
Db        223 TCTCTGATGACATGGAACATC 203

RESULT 20
US-09-949-016-54230
; Sequence 54230, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 54230
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-54230

Alignment Scores:
Pred. No.:              521             Length:         601
Score:                  35.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-54230 (1-601)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        441 TCAGTAATAACATGGAATATC 461

RESULT 21
US-09-949-016-54231
; Sequence 54231, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 54231
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-54231

Alignment Scores:
Pred. No.:                521             Length:         601
Score:                    35.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-54231 (1-601)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||||
Db        446 TCAGTAATAACATGGAATATC 466

RESULT 22
US-08-370-975B-12
; Sequence 12, Application US/08370975B
; Patent No. 5622851
;  GENERAL INFORMATION:
;    APPLICANT:  Maley, Frank
;    APPLICANT:  Maley, Gladys F.
;    APPLICANT:  Weiner, Karen X.B.
;    TITLE OF INVENTION:  Human Deoxycytidylate Deaminase Gene
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon, Hargrave, Devans & Doyle
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/370,975B
;      FILING DATE:  10-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Timian, Susan J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  20894/80
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716)263-1636
;      TELEFAX:  (716)263-1600
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1297 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  4q35
;
US-08-370-975B-12

Alignment Scores:
Pred. No.:                1.21e+03        Length:         1297
Score:                    35.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       2               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-08-370-975B-12 (1-1297)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        158 TCTCTTGTCACATGGAATCTA 178

RESULT 23
US-08-370-975B-1
; Sequence 1, Application US/08370975B
; Patent No. 5622851
;  GENERAL INFORMATION:
;    APPLICANT:  Maley, Frank
;    APPLICANT:  Maley, Gladys F.
;    APPLICANT:  Weiner, Karen X.B.
;    TITLE OF INVENTION:  Human Deoxycytidylate Deaminase Gene
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon, Hargrave, Devans & Doyle
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/370,975B
;      FILING DATE:  10-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Timian, Susan J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  20894/80
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716)263-1636
;      TELEFAX:  (716)263-1600
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26764 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  4q35
;
US-08-370-975B-1

Alignment Scores:
Pred. No.:                3.24e+04        Length:         26764
Score:                    35.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       2               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-08-370-975B-1 (1-26764)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db      25625 TCTCTTGTCACATGGAATCTA 25645
```

```
RESULT 24
US-09-949-016-13305/c
; Sequence 13305, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13305
;   LENGTH: 68452
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13305

Alignment Scores:
Pred. No.:                8.91e+04        Length:         68452
Score:                    35.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-13305 (1-68452)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db      47144 TCAGTAATAACATGGAATATC 47124

RESULT 25
US-09-949-016-16412/c
; Sequence 16412, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16412
;   LENGTH: 95890
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16412

Alignment Scores:
Pred. No.:                1.28e+05        Length:         95890
Score:                    35.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-16412 (1-95890)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db      94376 TCTCTGATGACATGGAACATC 94356

RESULT 26
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-20

Alignment Scores:
Pred. No.:                3.41e+05        Length:         246240
Score:                    35.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       2               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-08-724-394A-20 (1-246240)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db     118345 GCACTCATAACATGGAATATT 118365

RESULT 27
US-08-724-394A-21
```

```
; Sequence 21, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-21

Alignment Scores:
Pred. No.:                3.41e+05        Length:          246240
Score:                    35.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              92.1%           Indels:          0
DB:                       2               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-08-724-394A-21 (1-246240)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db     118345 GCACTCATAACATGGAATATT 118365

RESULT 28
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-22

Alignment Scores:
Pred. No.:                3.41e+05        Length:          246240
Score:                    35.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              92.1%           Indels:          0
DB:                       2               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-08-724-394A-22 (1-246240)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db     118345 GCACTCATAACATGGAATATT 118365

RESULT 29
US-09-768-185A-1/c
; Sequence 1, Application US/09768185A
; Patent No. 6818758
; GENERAL INFORMATION:
;  APPLICANT: Cassel, Michael et al
;  TITLE OF INVENTION: Estrogen receptor beta variants and
;  TITLE OF INVENTION:  methods of detection thereof
;  FILE REFERENCE: CL000280
;  CURRENT APPLICATION NUMBER: US/09/768,185A
;  CURRENT FILING DATE:  2001-01-24
;  PRIOR APPLICATION NUMBER: 09768185
;  PRIOR FILING DATE: 2001-01-24
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 325791
;   TYPE: DNA
```

; ORGANISM: HUMAN
US-09-768-185A-1

Alignment Scores:

Pred. No.:	4.52e+05	Length:	325791
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-768-185A-1 (1-325791)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db     135087 TCAGTAATAACATGGAATATC 135067
```

RESULT 30
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g

```
;   NAME/KEY: misc_feature
;   LOCATION: (855539)..(855539)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (871619)..(871619)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1084830)..(1084830)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1096846)..(1096846)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1119881)..(1119881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1130881)..(1130881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1310988)..(1310988)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1313224)..(1313224)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349473)..(1349473)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349491)..(1349491)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1470091)..(1470091)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1569020)..(1569020)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1602912)..(1602912)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1603734)..(1603734)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1637998)..(1637998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1664854)..(1664855)
;   OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Alignment Scores:
Pred. No.:                1.76e+06        Length:          1664976
Score:                    35.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              92.1%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-08-916-421B-1 (1-1664976)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db     776754 AATCTGATTACATGGAACATT 776734

RESULT 31
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
;  APPLICANT: Bult et al.
;  TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6797466
;  TITLE OF INVENTION:  jannaschii
;  FILE REFERENCE: PB275C1
;  CURRENT APPLICATION NUMBER: US/09/692,570
;  CURRENT FILING DATE:  2003-01-14
;  PRIOR APPLICATION NUMBER: US 60/024,428
;  PRIOR FILING DATE: 1996-08-22
;  PRIOR APPLICATION NUMBER: US 08/916,421
;  PRIOR FILING DATE: 1997-08-22
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1664976
;   TYPE: DNA
;   ORGANISM: Methanococcus jannaschii
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28222)..(28222)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28257)..(28258)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (84773)..(84773)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (84808)..(84808)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (84812)..(84812)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (98120)..(98120)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (98159)..(98159)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (98239)..(98239)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (98266)..(98266)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (98343)..(98343)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (103998)..(103998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (148948)..(148948)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (163385)..(163385)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (191989)..(191989)
;   OTHER INFORMATION: n equals a, t, c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (191995)..(191995)
;   OTHER INFORMATION: n equals a, t, c, or g
```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

Alignment Scores:

Pred. No.:	1.76e+06	Length:	1664976
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-692-570-1 (1-1664976)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db     776754 AATCTGATTACATGGAACATT 776734
```

RESULT 32
US-09-614-221A-23
; Sequence 23, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516.075

```
;  CURRENT APPLICATION NUMBER: US/09/614,221A
;  CURRENT FILING DATE:  2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/142,981
;  PRIOR FILING DATE: 1999-07-12
;  NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 23
;    LENGTH: 343
;    TYPE: DNA
;    ORGANISM: Arabidopsis thaliana
US-09-614-221A-23

Alignment Scores:
Pred. No.:                440             Length:          343
Score:                    34.00           Matches:         5
Percent Similarity:       100.0%          Conservative:    2
Best Local Similarity:    71.4%           Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-614-221A-23 (1-343)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::::|||||||||||||
Db        110 TCAGTGGTGACGTGGAACATC 130

RESULT 33
US-09-949-016-20188
; Sequence 20188, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20188
;    LENGTH: 601
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-20188

Alignment Scores:
Pred. No.:                810             Length:          601
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-20188 (1-601)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         17 AGTCTGATCACTTGGAAT 34

RESULT 34
US-09-949-016-20189
; Sequence 20189, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20189
;    LENGTH: 601
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-20189

Alignment Scores:
Pred. No.:                810             Length:          601
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-20189 (1-601)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         26 AGTCTGATCACTTGGAAT 43

RESULT 35
US-09-949-016-20190
; Sequence 20190, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20190
;    LENGTH: 601
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-20190

Alignment Scores:
Pred. No.:                810             Length:          601
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-20190 (1-601)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         92 AGTCTGATCACTTGGAAT 109

RESULT 36
US-09-949-016-44629/c
; Sequence 44629, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44629
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-44629

Alignment Scores:
Pred. No.:                810             Length:         601
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-44629 (1-601)

Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db        104 TTAATTACATGGAACATA 87

RESULT 37
US-09-949-016-108625
; Sequence 108625, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 108625
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-108625

Alignment Scores:
Pred. No.:                810             Length:         601
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-108625 (1-601)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        212 TCTCTTCTCACATGGAACTTG 232

RESULT 38
US-09-949-016-108626
; Sequence 108626, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 108626
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-108626

Alignment Scores:
Pred. No.:                810             Length:         601
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-108626 (1-601)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        188 TCTCTTCTCACATGGAACTTG 208

RESULT 39
US-09-949-016-129014
; Sequence 129014, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 129014
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-129014

Alignment Scores:
Pred. No.:                810             Length:         601
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0
```

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-129014 (1-601)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||:::
Db        120 TCAATAATTACCTGGAATCTA 140
```

RESULT 40
US-09-949-016-165434
; Sequence 165434, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 165434
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-165434

Alignment Scores:

Pred. No.:	810	Length:	601
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-165434 (1-601)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         17 AGTCTGATCACTTGGAAT 34
```

RESULT 41
US-09-949-016-165435
; Sequence 165435, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 165435
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-165435

Alignment Scores:

Pred. No.:	810	Length:	601
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-165435 (1-601)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         26 AGTCTGATCACTTGGAAT 43
```

RESULT 42
US-09-949-016-165436
; Sequence 165436, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 165436
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-165436

Alignment Scores:

Pred. No.:	810	Length:	601
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-165436 (1-601)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         92 AGTCTGATCACTTGGAAT 109
```

RESULT 43
US-09-949-016-165491
; Sequence 165491, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 165491
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human

US-09-949-016-165491

Alignment Scores:

Pred. No.:	810	Length:	601
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-165491 (1-601)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         17 AGTCTGATCACTTGGAAT 34
```

RESULT 44
US-09-949-016-165492
; Sequence 165492, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 165492
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-165492

Alignment Scores:

Pred. No.:	810	Length:	601
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-165492 (1-601)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         26 AGTCTGATCACTTGGAAT 43
```

RESULT 45
US-09-949-016-165493
; Sequence 165493, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 165493
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-165493

Alignment Scores:

Pred. No.:	810	Length:	601
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-165493 (1-601)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         92 AGTCTGATCACTTGGAAT 109
```

RESULT 46
US-09-949-016-196355
; Sequence 196355, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 196355
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-196355

Alignment Scores:

Pred. No.:	810	Length:	601
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-196355 (1-601)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||:::|||||||||
Db        133 TCACTTATTAGCTGGAATATT 153
```

RESULT 47
US-09-949-016-196356
; Sequence 196356, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 196356
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-196356

Alignment Scores:

Pred. No.:	810	Length:	601
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-196356 (1-601)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||:::|||||||||
Db         52 TCACTTATTAGCTGGAATATT 72
```

RESULT 48
US-09-328-352-1038/c
; Sequence 1038, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1038
; LENGTH: 912
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1038

Alignment Scores:

Pred. No.:	1.28e+03	Length:	912
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-328-352-1038 (1-912)

```
Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db         60 CTGATCACATGGAACATA 43
```

RESULT 49
US-09-697-863A-3
; Sequence 3, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/09/697,863A
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (122)..(1234)
; OTHER INFORMATION:
US-09-697-863A-3

Alignment Scores:

Pred. No.:	1.9e+03	Length:	1312
Score:	34.00	Matches:	6
Percent Similarity:	85.7%	Conservative:	0
Best Local Similarity:	85.7%	Mismatches:	1
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-697-863A-3 (1-1312)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||   ||||||||||||||||
Db        494 TCTTTCATTACCTGGAATATT 514
```

RESULT 50
US-09-583-110-2462/c
; Sequence 2462, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2462
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2462

Alignment Scores:

Pred. No.:	2.16e+03	Length:	1476
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-583-110-2462 (1-1476)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||||
Db         99 ACGCTTGTGACCTGGAACATC 79
```

RESULT 51
US-09-107-433-2350/c
; Sequence 2350, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206

```
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;            STREET: 100 Beaver Street
;            CITY: Waltham
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02354
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: CD/ROM ISO9660
;            COMPUTER: <Unknown>
;            OPERATING SYSTEM: <Unknown>
;            SOFTWARE: <Unknown>
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/107,433
;            FILING DATE: 30-Jun-1998
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/ 085131
;            FILING DATE: May 12, 1998
;            APPLICATION NUMBER: 60/051553
;            FILING DATE: July 2, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ariniello, Pamela Deneke
;            REGISTRATION NUMBER: 40,489
;            REFERENCE/DOCKET NUMBER: GTC-011
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (781)893-5007
;            TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 2350:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1488 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: circular
;       MOLECULE TYPE: DNA (genomic)
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       ORIGINAL SOURCE:
;           ORGANISM: Streptococcus pneumoniae
;       FEATURE:
;            NAME/KEY:  misc_feature
;            LOCATION: (B) LOCATION 1...1488
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2350:
US-09-107-433-2350

Alignment Scores:
Pred. No.:                2.18e+03        Length:         1488
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-107-433-2350 (1-1488)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||||
Db        111 ACGCTTGTGACCTGGAACATC 91

RESULT 52
US-09-583-110-518/c
; Sequence 518, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
;  APPLICANT:  Lynn Doucette-Stamm et al.
;  TITLE OF INVENTION:  Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
;  TITLE OF INVENTION:  Pneumoniae for Diagnostics and Therapeutics
;  FILE REFERENCE:  PATH00-07A
;  CURRENT APPLICATION NUMBER: US/09/583,110
;  CURRENT FILING DATE:  2000-05-26
;  PRIOR APPLICATION NUMBER:  US 09/107,433
;  PRIOR FILING DATE:  1998-06-30
;  PRIOR APPLICATION NUMBER:  US 60/085,131
;  PRIOR FILING DATE:  1998-05-12
;  PRIOR APPLICATION NUMBER:  US 60/051,553
;  PRIOR FILING DATE:  1997-07-02
;  NUMBER OF SEQ ID NOS:  5322
; SEQ ID NO 518
;   LENGTH: 1518
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
US-09-583-110-518

Alignment Scores:
Pred. No.:                2.22e+03        Length:         1518
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-583-110-518 (1-1518)

Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db        170 CTCATTACTTGGAATATT 153

RESULT 53
US-09-769-787-334/c
; Sequence 334, Application US/09769787
; Patent No. 6936252
; GENERAL INFORMATION:
;  APPLICANT: Microbial Technics Limited
;  APPLICANT:  Gilbert, Christophe FG
;  APPLICANT:  Hansbro, Philip M
;  TITLE OF INVENTION: Proteins
;  FILE REFERENCE: PWC/P21129WO
;  CURRENT APPLICATION NUMBER: US/09/769,787
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER: GB 9816337.1
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: US 60/125164
;  PRIOR FILING DATE: 1999-03-19
;  NUMBER OF SEQ ID NOS: 388
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 334
;   LENGTH: 1518
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
US-09-769-787-334

Alignment Scores:
Pred. No.:                2.22e+03        Length:         1518
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-769-787-334 (1-1518)

Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db        170 CTCATTACTTGGAATATT 153

RESULT 54
US-09-107-433-1911/c
; Sequence 1911, Application US/09107433
; Patent No. 6800744
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
;                            SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE   FOR DIAGNO
;                            THERAPEUTICS
;        NUMBER OF SEQUENCES: 5206
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
```

; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1911:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1536 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...1536
; SEQUENCE DESCRIPTION: SEQ ID NO: 1911:
US-09-107-433-1911

Alignment Scores:

Pred. No.:	2.25e+03	Length:	1536
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-107-433-1911 (1-1536)

```
Qy          2 LeuIleThrTrpAsnIle 7
              |||||||||||||||||
Db        188 CTCATTACTTGGAATATT 171
```

RESULT 55
US-09-182-145-9
; Sequence 9, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 9
; LENGTH: 1766
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 10
; OTHER INFORMATION: Unknown base
US-09-182-145-9

Alignment Scores:

Pred. No.:	2.62e+03	Length:	1766
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-182-145-9 (1-1766)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db       1133 TCTTTGCTGACTTGGAATCTT 1153
```

RESULT 56
US-09-182-145-10/c
; Sequence 10, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 10
; LENGTH: 1766
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 1757
; OTHER INFORMATION: Unknown base.
US-09-182-145-10

Alignment Scores:

Pred. No.:	2.62e+03	Length:	1766
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2

```
Best Local Similarity:  71.4%          Mismatches:     0
Query Match:            89.5%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-182-145-10 (1-1766)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        634 TCTTTGCTGACTTGGAATCTT 614

RESULT 57
US-09-107-532A-3481/c
; Sequence 3481, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 3481:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1908 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: circular
;        MOLECULE TYPE: DNA (genomic)
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION: (B) LOCATION 1...1908
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3481:
US-09-107-532A-3481

Alignment Scores:
Pred. No.:              2.85e+03       Length:         1908
Score:                  34.00          Matches:        5
Percent Similarity:     100.0%         Conservative:   2
Best Local Similarity:  71.4%          Mismatches:     0
Query Match:            89.5%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-107-532A-3481 (1-1908)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||
Db        219 ACGCTGGTAACTTGGAATATA 199

RESULT 58
US-09-586-719-26/c
; Sequence 26, Application US/09586719
; Patent No. 6294345
; GENERAL INFORMATION:
;  APPLICANT:  Zheng, Feng
;  APPLICANT:  Levin, Joshua Z.
;  APPLICANT:  Bauer, Michael W.
;  TITLE OF INVENTION:  Herbicide Target Genes and Methods
;  FILE REFERENCE:  PB/5-30909A
;  CURRENT APPLICATION NUMBER: US/09/586,719
;  CURRENT FILING DATE:  2000-06-05
;  NUMBER OF SEQ ID NOS:  34
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 26
;   LENGTH: 3422
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-586-719-26

Alignment Scores:
Pred. No.:              5.39e+03       Length:         3422
Score:                  34.00          Matches:        5
Percent Similarity:     100.0%         Conservative:   2
Best Local Similarity:  71.4%          Mismatches:     0
Query Match:            89.5%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-586-719-26 (1-3422)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||
Db        841 AATTTGGTGACTTGGAATATT 821

RESULT 59
US-09-900-708-1/c
; Sequence 1, Application US/09900708
; Patent No. 6762342
; GENERAL INFORMATION:
;  APPLICANT: Allen, Keith D.
;  TITLE OF INVENTION: TRANSGENIC MICE CONTAINING INTESTINAL
;  TITLE OF INVENTION:  ALKALINE PHOSPHATASE GENE DISRUPTIONS
;  FILE REFERENCE: R-733
;  CURRENT APPLICATION NUMBER: US/09/900,708
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: US 60/216,476
;  PRIOR FILING DATE: 2000-07-06
;  PRIOR APPLICATION NUMBER: US 60/221,489
;  PRIOR FILING DATE: 2000-07-27
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 5293
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-900-708-1

Alignment Scores:
Pred. No.:              8.67e+03       Length:         5293
Score:                  34.00          Matches:        6
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            89.5%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-900-708-1 (1-5293)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
```

```
Db        588 AGCCTCATTACCTGGAAT 571

RESULT 60
US-09-949-016-12780/c
; Sequence 12780, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12780
;   LENGTH: 10698
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12780

Alignment Scores:
Pred. No.:                1.87e+04        Length:         10698
Score:                    34.00           Matches:        6
Percent Similarity:       85.7%           Conservative:   0
Best Local Similarity:    85.7%           Mismatches:     1
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-12780 (1-10698)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||   ||||||||||||
Db       9734 AGTCTCTTCACTTGGAACATT 9714

RESULT 61
US-08-961-527-72/c
; Sequence 72, Application US/08961527
; Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,527
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-527-72

Alignment Scores:
Pred. No.:                2.66e+04        Length:         14872
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-08-961-527-72 (1-14872)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||||
Db       5352 ACGCTTGTGACCTGGAACATC 5332

RESULT 62
US-08-961-527-3/c
; Sequence 3, Application US/08961527
; Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,527
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26385 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-527-3

Alignment Scores:
Pred. No.:                4.94e+04        Length:         26385
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0
```

US-10-757-745-2_COPY_115_121 (1-7) x US-08-961-527-3 (1-26385)

```
Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db      19457 CTCATTACTTGGAATATT 19440
```

RESULT 63
US-09-949-016-13442/c
; Sequence 13442, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13442
; LENGTH: 29653
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(29653)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13442

Alignment Scores:

Pred. No.:	5.6e+04	Length:	29653
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-13442 (1-29653)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db       9307 TCATTGATAACGTGGAAT 9290
```

RESULT 64
US-09-949-016-12704/c
; Sequence 12704, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12704
; LENGTH: 33155
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12704

Alignment Scores:

Pred. No.:	6.32e+04	Length:	33155
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-12704 (1-33155)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db       7420 TCCCTGATCACCTGGAAT 7403
```

RESULT 65
US-09-949-016-15756/c
; Sequence 15756, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15756
; LENGTH: 33157
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15756

Alignment Scores:

Pred. No.:	6.32e+04	Length:	33157
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-15756 (1-33157)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db       7420 TCCCTGATCACCTGGAAT 7403
```

RESULT 66
US-09-949-016-11860/c
; Sequence 11860, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11860
;   LENGTH: 36643
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-11860

Alignment Scores:
Pred. No.:                7.04e+04        Length:          36643
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-11860 (1-36643)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db      29523 AGTCTGATCACTTGGAAT 29506

RESULT 67
US-09-949-016-16403/c
; Sequence 16403, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16403
;   LENGTH: 36821
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16403

Alignment Scores:
Pred. No.:                7.07e+04        Length:          36821
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-16403 (1-36821)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db      29697 AGTCTGATCACTTGGAAT 29680

RESULT 68
US-09-949-016-16404/c
; Sequence 16404, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16404
;   LENGTH: 36821
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16404

Alignment Scores:
Pred. No.:                7.07e+04        Length:          36821
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-16404 (1-36821)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db      29697 AGTCTGATCACTTGGAAT 29680

RESULT 69
US-09-949-016-11862/c
; Sequence 11862, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11862
;   LENGTH: 37304
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(37304)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11862

Alignment Scores:
Pred. No.:                7.17e+04        Length:          37304
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-11862 (1-37304)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db       9307 TCATTGATAACGTGGAAT 9290

RESULT 70
US-09-949-016-13018/c
; Sequence 13018, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13018
;   LENGTH: 51273
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(51273)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13018

Alignment Scores:
Pred. No.:                1.01e+05        Length:         51273
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-13018 (1-51273)

Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db       7461 TTAATTACATGGAACATA 7444

RESULT 71
US-09-949-016-13574/c
; Sequence 13574, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13574
;   LENGTH: 66164
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13574

Alignment Scores:
Pred. No.:                1.32e+05        Length:         66164
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-13574 (1-66164)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db      48620 TCTTTAATCACATGGAAT 48603

RESULT 72
US-09-918-686-1
; Sequence 1, Application US/09918686
; Patent No. 6475739
; GENERAL INFORMATION:
;  APPLICANT: Brunkow, Mary
;  APPLICANT:  Proll, Sean
;  APPLICANT:  Paeper, Bryan
;  APPLICANT:  Staehling-Hampton, Karen
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING
;  TITLE OF INVENTION:  GENOMIC DELETIONS
;  FILE REFERENCE: 240083.515
;  CURRENT APPLICATION NUMBER: US/09/918,686
;  CURRENT FILING DATE:  2001-07-30
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 92139
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 7043, 8369, 8401
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1

Alignment Scores:
Pred. No.:                1.88e+05        Length:         92139
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-918-686-1 (1-92139)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db      69863 TCCCTAATCACATGGAAC 69880

RESULT 73
US-09-949-016-14731
; Sequence 14731, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14731
;   LENGTH: 126029
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(126029)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14731
```

Alignment Scores:

Pred. No.:	2.6e+05	Length:	126029
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-14731 (1-126029)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db      17811 TCTCTTCTCACATGGAACTTG 17831
```

RESULT 74
US-09-949-016-15402/c
; Sequence 15402, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15402
; LENGTH: 152070
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15402

Alignment Scores:

Pred. No.:	3.15e+05	Length:	152070
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-15402 (1-152070)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||:::
Db      42376 TCAATAATTACCTGGAATCTA 42356
```

RESULT 75
US-09-949-016-12896
; Sequence 12896, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12896
; LENGTH: 451924
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12896

Alignment Scores:

Pred. No.:	8.97e+05	Length:	451924
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-949-016-12896 (1-451924)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||:::|||||||||
Db     131996 TCACTTATTAGCTGGAATATT 132016
```

Search completed: March 10, 2006, 21:40:26
Job time : 72.9606 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 03:44:14 ; Search time 20.2582 Seconds
(without alignments)
797.263 Million cell updates/sec

Title: US-10-757-745-2_COPY_115_121
Perfect score: 38
Sequence: 1 SLITWNI 7

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7673375 seqs, 1153648444 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170339_4330/app_query.fasta_1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1676_@runat_09032006_170339_4330 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	606	6	US-09-925-065A-783537	Sequence 783537,
2	38	100.0	634	6	US-09-925-065A-783536	Sequence 783536,
3	38	100.0	1498	6	US-09-925-065A-548649	Sequence 548649,
4	38	100.0	1498	6	US-09-925-065A-548650	Sequence 548650,
5	38	100.0	1498	6	US-09-925-065A-548651	Sequence 548651,
6	37	97.4	563	6	US-09-925-065A-134078	Sequence 134078,
c 7	37	97.4	1126	6	US-09-925-065A-78968	Sequence 78968, A
8	36	94.7	542	6	US-09-925-065A-767316	Sequence 767316,
9	36	94.7	542	6	US-09-925-065A-767317	Sequence 767317,
c 10	36	94.7	614	6	US-09-925-065A-953237	Sequence 953237,
c 11	36	94.7	842	6	US-09-925-065A-64641	Sequence 64641, A
c 12	36	94.7	842	6	US-09-925-065A-64642	Sequence 64642, A
c 13	36	94.7	1155	9	US-11-096-568A-4427	Sequence 4427, Ap
c 14	36	94.7	1473	7	US-10-932-182A-82264	Sequence 82264, A
c 15	36	94.7	1473	7	US-10-932-182A-82264	Sequence 82264, A
c 16	36	94.7	2087	9	US-11-072-512-84	Sequence 84, Appl
17	36	94.7	2303	9	US-11-072-512-1924	Sequence 1924, Ap
18	36	94.7	156735	12	US-11-121-086-93	Sequence 93, Appl
c 19	35	92.1	25	12	US-11-121-849-297571	Sequence 297571,
c 20	35	92.1	50	12	US-11-175-859-113577	Sequence 113577,
c 21	35	92.1	459	6	US-09-925-065A-533645	Sequence 533645,
c 22	35	92.1	570	6	US-09-925-065A-665681	Sequence 665681,
c 23	35	92.1	584	6	US-09-925-065A-370878	Sequence 370878,
24	35	92.1	607	6	US-09-925-065A-251509	Sequence 251509,
c 25	35	92.1	638	6	US-09-925-065A-889148	Sequence 889148,
c 26	35	92.1	652	6	US-09-925-065A-394687	Sequence 394687,
c 27	35	92.1	747	6	US-09-925-065A-16142	Sequence 16142, A
28	35	92.1	1261	8	US-10-750-185-33899	Sequence 33899, A
29	35	92.1	1261	8	US-10-750-623-33899	Sequence 33899, A
c 30	35	92.1	1418	8	US-10-750-185-53228	Sequence 53228, A
c 31	35	92.1	1418	8	US-10-750-623-53228	Sequence 53228, A
32	35	92.1	2046	9	US-11-096-568A-28383	Sequence 28383, A
33	35	92.1	5278	8	US-10-750-185-61034	Sequence 61034, A
34	35	92.1	5278	8	US-10-750-623-61034	Sequence 61034, A
c 35	35	92.1	235033	12	US-11-157-389-1	Sequence 1, Appli
c 36	35	92.1	237326	12	US-11-157-389-2	Sequence 2, Appli
37	35	92.1	440684	7	US-10-330-773-114	Sequence 114, App
38	34	89.5	19	10	US-11-101-244-984899	Sequence 984899,
39	34	89.5	19	11	US-11-083-784-984899	Sequence 984899,
40	34	89.5	78	8	US-10-310-914A-6342	Sequence 6342, Ap
c 41	34	89.5	398	6	US-09-925-065A-212315	Sequence 212315,
c 42	34	89.5	408	6	US-09-925-065A-237156	Sequence 237156,
c 43	34	89.5	419	6	US-09-925-065A-585692	Sequence 585692,
c 44	34	89.5	446	6	US-09-925-065A-212316	Sequence 212316,
c 45	34	89.5	452	6	US-09-925-065A-212314	Sequence 212314,
c 46	34	89.5	458	6	US-09-925-065A-150206	Sequence 150206,
c 47	34	89.5	464	6	US-09-925-065A-144172	Sequence 144172,
c 48	34	89.5	498	6	US-09-925-065A-344924	Sequence 344924,
c 49	34	89.5	498	6	US-09-925-065A-344925	Sequence 344925,
50	34	89.5	521	6	US-09-925-065A-207175	Sequence 207175,
51	34	89.5	542	6	US-09-925-065A-621687	Sequence 621687,
c 52	34	89.5	549	6	US-09-925-065A-227040	Sequence 227040,
c 53	34	89.5	554	6	US-09-925-065A-226154	Sequence 226154,
c 54	34	89.5	554	6	US-09-925-065A-226155	Sequence 226155,
55	34	89.5	555	6	US-09-925-065A-560599	Sequence 560599,
56	34	89.5	555	6	US-09-925-065A-560600	Sequence 560600,
57	34	89.5	555	6	US-09-925-065A-560601	Sequence 560601,
c 58	34	89.5	560	6	US-09-925-065A-331333	Sequence 331333,
59	34	89.5	561	6	US-09-925-065A-103567	Sequence 103567,
60	34	89.5	562	6	US-09-925-065A-731658	Sequence 731658,
61	34	89.5	567	6	US-09-925-065A-122265	Sequence 122265,
c 62	34	89.5	578	6	US-09-925-065A-454651	Sequence 454651,
c 63	34	89.5	578	6	US-09-925-065A-454652	Sequence 454652,
c 64	34	89.5	578	6	US-09-925-065A-454653	Sequence 454653,
c 65	34	89.5	579	6	US-09-925-065A-126922	Sequence 126922,
c 66	34	89.5	579	6	US-09-925-065A-126923	Sequence 126923,
c 67	34	89.5	580	6	US-09-925-065A-782074	Sequence 782074,
c 68	34	89.5	580	6	US-09-925-065A-782075	Sequence 782075,
c 69	34	89.5	580	6	US-09-925-065A-842988	Sequence 842988,
70	34	89.5	583	6	US-09-925-065A-322599	Sequence 322599,
c 71	34	89.5	585	6	US-09-925-065A-103578	Sequence 103578,
72	34	89.5	585	6	US-09-925-065A-460308	Sequence 460308,
c 73	34	89.5	587	6	US-09-925-065A-946652	Sequence 946652,
c 74	34	89.5	588	6	US-09-925-065A-215413	Sequence 215413,
c 75	34	89.5	588	6	US-09-925-065A-822492	Sequence 822492,

ALIGNMENTS

RESULT 1
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-783537

Alignment Scores:			
Pred. No.:	303	Length:	606
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-783537 (1-606)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        504 TCTCTCATTACCTGGAATATT 524
```

RESULT 2
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-783536

Alignment Scores:			
Pred. No.:	321	Length:	634
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-783536 (1-634)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        355 TCTCTCATTACCTGGAATATT 375
```

RESULT 3
US-09-925-065A-548649
; Sequence 548649, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548649
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-548649

Alignment Scores:			
Pred. No.:	950	Length:	1498
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-548649 (1-1498)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db       1402 TCATTAATAACTTGGAACATT 1422
```

RESULT 4
US-09-925-065A-548650
; Sequence 548650, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20

```
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548650
;   LENGTH: 1498
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-548650

Alignment Scores:
Pred. No.:                950             Length:         1498
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-548650 (1-1498)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db       1402 TCATTAATAACTTGGAACATT 1422

RESULT 5
US-09-925-065A-548651
; Sequence 548651, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548651
;   LENGTH: 1498
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-548651

Alignment Scores:
Pred. No.:                950             Length:         1498
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-548651 (1-1498)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db       1402 TCATTAATAACTTGGAACATT 1422

RESULT 6
US-09-925-065A-134078
; Sequence 134078, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134078
;   LENGTH: 563
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-134078

Alignment Scores:
Pred. No.:                453             Length:         563
Score:                    37.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              97.4%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-134078 (1-563)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        532 TCCCTAGTCACATGGAATATT 552

RESULT 7
US-09-925-065A-78968/c
; Sequence 78968, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78968
;   LENGTH: 1126
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-78968

Alignment Scores:
Pred. No.:                1.09e+03        Length:         1126
Score:                    37.00           Matches:        6
```

```
Percent Similarity:     100.0%          Conservative:  1
Best Local Similarity:  85.7%           Mismatches:    0
Query Match:            97.4%           Indels:        0
DB:                     6               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-78968 (1-1126)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        876 TCACTTGTAACTTGGAATATT 856

RESULT 8
US-09-925-065A-767316
; Sequence 767316, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 767316
;   LENGTH: 542
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-767316

Alignment Scores:
Pred. No.:              709             Length:        542
Score:                  36.00           Matches:       6
Percent Similarity:     100.0%          Conservative:  1
Best Local Similarity:  85.7%           Mismatches:    0
Query Match:            94.7%           Indels:        0
DB:                     6               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-767316 (1-542)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||||
Db        475 TCCATCATCACATGGAACATT 495

RESULT 9
US-09-925-065A-767317
; Sequence 767317, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 767317
;   LENGTH: 542
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-767317

Alignment Scores:
Pred. No.:              709             Length:        542
Score:                  36.00           Matches:       6
Percent Similarity:     100.0%          Conservative:  1
Best Local Similarity:  85.7%           Mismatches:    0
Query Match:            94.7%           Indels:        0
DB:                     6               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-767317 (1-542)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||||
Db        475 TCCATCATCACATGGAACATT 495

RESULT 10
US-09-925-065A-953237/c
; Sequence 953237, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 953237
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-953237

Alignment Scores:
Pred. No.:              830             Length:        614
Score:                  36.00           Matches:       6
Percent Similarity:     100.0%          Conservative:  1
Best Local Similarity:  85.7%           Mismatches:    0
Query Match:            94.7%           Indels:        0
DB:                     6               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-953237 (1-614)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||||
Db        515 TCCATCATCACATGGAACATT 495

RESULT 11
US-09-925-065A-64641/c
; Sequence 64641, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

```
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64641
;   LENGTH: 842
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-64641

Alignment Scores:
Pred. No.:              1.24e+03        Length:         842
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-64641 (1-842)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db        161 TCATTGATTACCTGGAATTTA 141

RESULT 12
US-09-925-065A-64642/c
; Sequence 64642, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64642
;   LENGTH: 842
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-64642

Alignment Scores:
Pred. No.:              1.24e+03        Length:         842
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-64642 (1-842)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db        161 TCATTGATTACCTGGAATTTA 141

RESULT 13
US-11-096-568A-4427/c
; Sequence 4427, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4427
;   LENGTH: 1155
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1155)
;   OTHER INFORMATION: Ceres Seq. ID no. 13637685
US-11-096-568A-4427

Alignment Scores:
Pred. No.:              1.84e+03        Length:         1155
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-096-568A-4427 (1-1155)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        731 AGCCTCTTAACCTGGAACATT 711

RESULT 14
US-10-932-182A-82264/c
; Sequence 82264, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82264
;   LENGTH: 1473
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1299)..(1384)
;   OTHER INFORMATION: a, c, g, t, unknown, or other
US-10-932-182A-82264

Alignment Scores:
Pred. No.:              2.5e+03         Length:         1473
Score:                  36.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
```

```
Best Local Similarity:  71.4%          Mismatches:     0
Query Match:            94.7%          Indels:         0
DB:                     7              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-932-182A-82264 (1-1473)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||:::|||||||||:::
Db        550 TCGCTTGTAACTTGGAACGTA 530

RESULT 15
US-10-932-182A-82264/c
; Sequence 82264, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82264
;   LENGTH: 1473
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1299)..(1384)
;   OTHER INFORMATION: a, c, g, t, unknown, or other
US-10-932-182A-82264

Alignment Scores:
Pred. No.:              2.5e+03        Length:         1473
Score:                  36.00          Matches:        5
Percent Similarity:     100.0%         Conservative:   2
Best Local Similarity:  71.4%          Mismatches:     0
Query Match:            94.7%          Indels:         0
DB:                     7              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-932-182A-82264 (1-1473)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||:::|||||||||:::
Db        550 TCGCTTGTAACTTGGAACGTA 530

RESULT 16
US-11-072-512-84/c
; Sequence 84, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
;   LENGTH: 2087
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-072-512-84

Alignment Scores:
Pred. No.:              3.89e+03       Length:         2087
Score:                  36.00          Matches:        6
Percent Similarity:     100.0%         Conservative:   1
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            94.7%          Indels:         0
DB:                     9              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-072-512-84 (1-2087)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        164 AGCATTATTACGTGGAACATC 144

RESULT 17
US-11-072-512-1924
; Sequence 1924, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1924
;   LENGTH: 2303
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-072-512-1924

Alignment Scores:
Pred. No.:              4.4e+03        Length:         2303
Score:                  36.00          Matches:        5
Percent Similarity:     100.0%         Conservative:   2
```

```
Best Local Similarity:  71.4%          Mismatches:     0
Query Match:            94.7%          Indels:         0
DB:                     9              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-072-512-1924 (1-2303)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        993 AGCTTAGTAACTTGGAATGTT 1013

RESULT 18
US-11-121-086-93
; Sequence 93, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 93
;   LENGTH: 156735
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-93

Alignment Scores:
Pred. No.:              8.8e+05        Length:         156735
Score:                  36.00          Matches:        6
Percent Similarity:     100.0%         Conservative:   1
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            94.7%          Indels:         0
DB:                     12             Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-121-086-93 (1-156735)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db     102478 TCTTTAATTACCTGGAATTTG 102498

RESULT 19
US-11-121-849-297571/c
; Sequence 297571, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
;  APPLICANT: John Palma
;  TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
;  TITLE OF INVENTION:  Microarrays
;  FILE REFERENCE: 3684.1
;  CURRENT APPLICATION NUMBER: US/11/121,849
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: 60/567,949
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 673904
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 297571
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-11-121-849-297571

Alignment Scores:
Pred. No.:              23.9           Length:         25
Score:                  35.00          Matches:        6
Percent Similarity:     100.0%         Conservative:   1
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            92.1%          Indels:         0
DB:                     12             Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-121-849-297571 (1-25)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db         25 TCTCTGATGACATGGAACATC 5

RESULT 20
US-11-175-859-113577/c
; Sequence 113577, Application US/11175859
; Publication No. US20060024715A1
; GENERAL INFORMATION:
;  APPLICANT: Affymetrix, Inc.
;  APPLICANT:  Liu, Guoying et al.
;  TITLE OF INVENTION: Method of Analysis of Human Polymorphism
;  FILE REFERENCE: 3690.1
;  CURRENT APPLICATION NUMBER: US/11/175,859
;  CURRENT FILING DATE:  2005-07-05
;  PRIOR APPLICATION NUMBER: US 60/585,352
;  PRIOR FILING DATE: 2004-07-02
;  NUMBER OF SEQ ID NOS: 116251
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 113577
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: homo sapien
US-11-175-859-113577

Alignment Scores:
Pred. No.:              57.4           Length:         50
Score:                  35.00          Matches:        6
Percent Similarity:     100.0%         Conservative:   1
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            92.1%          Indels:         0
DB:                     12             Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-175-859-113577 (1-50)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||
Db         28 ACRTTGATTACCTGGAACATT 8

RESULT 21
US-09-925-065A-533645/c
; Sequence 533645, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 533645
;   LENGTH: 459
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-533645

Alignment Scores:
Pred. No.:              943            Length:         459
```

Score:	35.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-533645 (1-459)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db         31 AGTTTAGTAACCTGGAATTTA 11
```

RESULT 22
US-09-925-065A-665681/c
; Sequence 665681, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 665681
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-665681

Alignment Scores:			
Pred. No.:	1.24e+03	Length:	570
Score:	35.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-665681 (1-570)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||:::
Db         87 TCAATAATCACTTGGAATGTC 67
```

RESULT 23
US-09-925-065A-370878/c
; Sequence 370878, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 370878
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-370878

Alignment Scores:			
Pred. No.:	1.28e+03	Length:	584
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-370878 (1-584)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::||||||||||||||||||
Db        419 GCCTTAATCACTTGGAATATT 399
```

RESULT 24
US-09-925-065A-251509
; Sequence 251509, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 251509
; LENGTH: 607
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-251509

Alignment Scores:			
Pred. No.:	1.34e+03	Length:	607
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-251509 (1-607)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::||||||||||||||||||
Db         48 ACATTAATCACATGGAATATT 68
```

RESULT 25
US-09-925-065A-889148/c
; Sequence 889148, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:

```
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 889148
;   LENGTH: 638
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-889148

Alignment Scores:
Pred. No.:                1.43e+03        Length:         638
Score:                    35.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-889148 (1-638)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||:::
Db        500 AGTCTCCTCACCTGGAATGTG 480

RESULT 26
US-09-925-065A-394687/c
; Sequence 394687, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 394687
;   LENGTH: 652
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-394687

Alignment Scores:
Pred. No.:                1.47e+03        Length:         652
Score:                    35.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-394687 (1-652)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::||||||||||:::
Db         58 TCAATAATTACCTGGAATGTA 38

RESULT 27
US-09-925-065A-16142/c
; Sequence 16142, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16142
;   LENGTH: 747
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-16142

Alignment Scores:
Pred. No.:                1.74e+03        Length:         747
Score:                    35.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-16142 (1-747)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::||||||||||:::
Db        542 AGTATCATTACGTGGAATGTT 522

RESULT 28
US-10-750-185-33899
; Sequence 33899, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 33899
;   LENGTH: 1261
```

```
;   TYPE: DNA
;   ORGANISM: Bovine  19866881273600
US-10-750-185-33899

Alignment Scores:
Pred. No.:                3.38e+03        Length:          1261
Score:                    35.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              92.1%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-750-185-33899 (1-1261)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db        486 GCTCTGATCACGTGGAATATT 506

RESULT 29
US-10-750-623-33899
; Sequence 33899, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 33899
;   LENGTH: 1261
;   TYPE: DNA
;   ORGANISM: Bovine  19866881273600
US-10-750-623-33899

Alignment Scores:
Pred. No.:                3.38e+03        Length:          1261
Score:                    35.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              92.1%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-750-623-33899 (1-1261)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db        486 GCTCTGATCACGTGGAATATT 506

RESULT 30
US-10-750-185-53228/c
; Sequence 53228, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53228
;   LENGTH: 1418
;   TYPE: DNA
;   ORGANISM: Bovine  19866881385101
US-10-750-185-53228

Alignment Scores:
Pred. No.:                3.91e+03        Length:          1418
Score:                    35.00           Matches:         5
Percent Similarity:       100.0%          Conservative:    2
Best Local Similarity:    71.4%           Mismatches:      0
Query Match:              92.1%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-750-185-53228 (1-1418)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::||||||||:::
Db        804 TCTCTACTAACCTGGAATGTT 784

RESULT 31
US-10-750-623-53228/c
; Sequence 53228, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53228
;   LENGTH: 1418
;   TYPE: DNA
;   ORGANISM: Bovine  19866881385101
US-10-750-623-53228

Alignment Scores:
Pred. No.:                3.91e+03        Length:          1418
Score:                    35.00           Matches:         5
Percent Similarity:       100.0%          Conservative:    2
Best Local Similarity:    71.4%           Mismatches:      0
Query Match:              92.1%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-750-623-53228 (1-1418)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::||||||||:::
Db        804 TCTCTACTAACCTGGAATGTT 784

RESULT 32
US-11-096-568A-28383
; Sequence 28383, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
;  TITLE OF INVENTION:  Therby
```

; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28383
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2046)
; OTHER INFORMATION: Ceres Seq. ID no. 2717056
US-11-096-568A-28383

Alignment Scores:

Pred. No.:	6.22e+03	Length:	2046
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-096-568A-28383 (1-2046)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||||
Db        730 AGTGTCATAACTTGGAATATT 750
```

RESULT 33
US-10-750-185-61034
; Sequence 61034, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 61034
; LENGTH: 5278
; TYPE: DNA
; ORGANISM: Bovine 19866880568555
US-10-750-185-61034

Alignment Scores:

Pred. No.:	2.06e+04	Length:	5278
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-750-185-61034 (1-5278)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||||
Db       2127 TCTGTTATCACCTGGAACATT 2147
```

RESULT 34
US-10-750-623-61034
; Sequence 61034, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 61034
; LENGTH: 5278
; TYPE: DNA
; ORGANISM: Bovine 19866880568555
US-10-750-623-61034

Alignment Scores:

Pred. No.:	2.06e+04	Length:	5278
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-750-623-61034 (1-5278)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||||
Db       2127 TCTGTTATCACCTGGAACATT 2147
```

RESULT 35
US-11-157-389-1/c
; Sequence 1, Application US/11157389
; Publication No. US20050266481A1
; GENERAL INFORMATION:
; APPLICANT: Ruddy, David A.
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: POLYMORPHISMS IN THE REGION OF THE HUMAN
; TITLE OF INVENTION: HEMOCHROMATOSIS GENE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/157,389
; FILING DATE: 20-june-2005
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,495
; FILING DATE: 07-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0057-999
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235033 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-11-157-389-1

Alignment Scores:

Pred. No.:	2.29e+06	Length:	235033
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-157-389-1 (1-235033)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db     121099 GCACTCATAACATGGAATATT 121079
```

RESULT 36
US-11-157-389-2/c
; Sequence 2, Application US/11157389
; Publication No. US20050266481A1
; GENERAL INFORMATION:
; APPLICANT: Ruddy, David A.
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: POLYMORPHISMS IN THE REGION OF THE HUMAN
; TITLE OF INVENTION: HEMOCHROMATOSIS GENE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/157,389
; FILING DATE: 20-june-2005
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,495
; FILING DATE: 07-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0057-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237326 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-11-157-389-2

Alignment Scores:

Pred. No.:	2.32e+06	Length:	237326
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-157-389-2 (1-237326)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db     123275 GCACTCATAACATGGAATATT 123255
```

RESULT 37
US-10-330-773-114
; Sequence 114, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 440684
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(440684)
; OTHER INFORMATION: n = A,T,C or G
US-10-330-773-114

Alignment Scores:

Pred. No.:	4.61e+06	Length:	440684
Score:	35.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-330-773-114 (1-440684)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||:::
Db     395255 TCAATAATTACATGGAATGTT 395275
```

RESULT 38
US-11-101-244-984899
; Sequence 984899, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary

; SEQ ID NO 984899
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-984899

Alignment Scores:

Pred. No.:	27.8	Length:	19
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-101-244-984899 (1-19)

```
Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db          1 CUCAUUACCUGGAAUAUU 18
```

RESULT 39
US-11-083-784-984899
; Sequence 984899, Application US/11083784
; Publication No. US20050245475A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/083,784
; CURRENT FILING DATE: 2005-03-18
; PRIOR APPLICATION NUMBER: US/10/714,333
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 984899
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-083-784-984899

Alignment Scores:

Pred. No.:	27.8	Length:	19
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-083-784-984899 (1-19)

```
Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db          1 CUCAUUACCUGGAAUAUU 18
```

RESULT 40
US-10-310-914A-6342
; Sequence 6342, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6342
; LENGTH: 78
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-6342

Alignment Scores:

Pred. No.:	165	Length:	78
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-310-914A-6342 (1-78)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db         13 UCUUUAUUGACAUGGAAUUUG 33
```

RESULT 41
US-09-925-065A-212315/c
; Sequence 212315, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 212315
; LENGTH: 398
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-212315

Alignment Scores:

Pred. No.:	1.29e+03	Length:	398
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-212315 (1-398)

```
Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db        103 CTAATTACCTGGAATATT 86
```

RESULT 42
US-09-925-065A-237156/c
; Sequence 237156, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 237156
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-237156

Alignment Scores:

Pred. No.:	1.33e+03	Length:	408
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-237156 (1-408)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db        382 TCTCTGATCACCTGGAAT 365
```

RESULT 43
US-09-925-065A-585692/c
; Sequence 585692, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 585692
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-585692

Alignment Scores:

Pred. No.:	1.38e+03	Length:	419
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-585692 (1-419)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||||
Db        122 GCATTAGTTACTTGGAATATA 102
```

RESULT 44
US-09-925-065A-212316/c
; Sequence 212316, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 212316
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-212316

Alignment Scores:

Pred. No.:	1.49e+03	Length:	446
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-212316 (1-446)

```
Qy          2 LeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        133 CTAATTACCTGGAATATT 116
```

RESULT 45
US-09-925-065A-212314/c
; Sequence 212314, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 212314
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-212314

Alignment Scores:

Pred. No.:	1.52e+03	Length:	452
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-212314 (1-452)

```
Qy          2 LeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        142 CTAATTACCTGGAATATT 125
```

RESULT 46
US-09-925-065A-150206/c
; Sequence 150206, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 150206
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-150206

Alignment Scores:

Pred. No.:	1.54e+03	Length:	458
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-150206 (1-458)

```
Qy          2 LeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        156 CTCATTACTTGGAACATT 139
```

RESULT 47
US-09-925-065A-144172/c
; Sequence 144172, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 144172
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-144172

Alignment Scores:

Pred. No.:	1.57e+03	Length:	464
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-144172 (1-464)

```
Qy          2 LeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        167 TTGATCACATGGAATATA 150
```

RESULT 48
US-09-925-065A-344924/c
; Sequence 344924, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 344924
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-344924

Alignment Scores:

Pred. No.:	1.71e+03	Length:	498
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-344924 (1-498)

```
Qy          2 LeuIleThrTrpAsnIle 7
```

```
                  ||||||||||||||||||
Db        139 CTAATTACATGGAATATT 122
```

RESULT 49
US-09-925-065A-344925/c
; Sequence 344925, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 344925
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-344925

Alignment Scores:

Pred. No.:	1.71e+03	Length:	498
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-344925 (1-498)

```
Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db        139 CTAATTACATGGAATATT 122
```

RESULT 50
US-09-925-065A-207175
; Sequence 207175, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 207175
; LENGTH: 521
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-207175

Alignment Scores:

Pred. No.:	1.81e+03	Length:	521
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-207175 (1-521)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||::::::|||||||||||||
Db        194 TCAATATTAACCTGGAATATA 214
```

RESULT 51
US-09-925-065A-621687
; Sequence 621687, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 621687
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-621687

Alignment Scores:

Pred. No.:	1.91e+03	Length:	542
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-621687 (1-542)

```
Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db        306 TTAATTACATGGAATATA 323
```

RESULT 52
US-09-925-065A-227040/c
; Sequence 227040, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147

```
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 227040
;   LENGTH: 549
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-227040

Alignment Scores:
Pred. No.:                1.94e+03        Length:         549
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-227040 (1-549)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||
Db        395 ACCCTGGTCACCTGGAATATT 375

RESULT 53
US-09-925-065A-226154/c
; Sequence 226154, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 226154
;   LENGTH: 554
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-226154

Alignment Scores:
Pred. No.:                1.96e+03        Length:         554
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-226154 (1-554)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db        349 TCTTTAATAACTTGGAAT 332

RESULT 54
US-09-925-065A-226155/c
; Sequence 226155, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 226155
;   LENGTH: 554
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-226155

Alignment Scores:
Pred. No.:                1.96e+03        Length:         554
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-226155 (1-554)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db        349 TCTTTAATAACTTGGAAT 332

RESULT 55
US-09-925-065A-560599
; Sequence 560599, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 560599
;   LENGTH: 555
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-560599

Alignment Scores:
Pred. No.:                1.97e+03        Length:         555
```

Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-560599 (1-555)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||
Db        295 GCATTAGTTACTTGGAATATA 315
```

RESULT 56
US-09-925-065A-560600
; Sequence 560600, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 560600
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-560600

Alignment Scores:			
Pred. No.:	1.97e+03	Length:	555
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-560600 (1-555)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||
Db        295 GCATTAGTTACTTGGAATATA 315
```

RESULT 57
US-09-925-065A-560601
; Sequence 560601, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 560601
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-560601

Alignment Scores:			
Pred. No.:	1.97e+03	Length:	555
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-560601 (1-555)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||
Db        295 GCATTAGTTACTTGGAATATA 315
```

RESULT 58
US-09-925-065A-331333/c
; Sequence 331333, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 331333
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-331333

Alignment Scores:			
Pred. No.:	1.99e+03	Length:	560
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-331333 (1-560)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db         91 TCATTACTGACCTGGAATCTT 71
```

RESULT 59
US-09-925-065A-103567
; Sequence 103567, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:

```
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103567
;   LENGTH: 561
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-103567

Alignment Scores:
Pred. No.:                1.99e+03        Length:          561
Score:                    34.00           Matches:         5
Percent Similarity:       100.0%          Conservative:    2
Best Local Similarity:    71.4%           Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-103567 (1-561)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||:::
Db        233 ACATTAATCACATGGAATGTT 253

RESULT 60
US-09-925-065A-731658
; Sequence 731658, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 731658
;   LENGTH: 562
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-731658

Alignment Scores:
Pred. No.:                2e+03           Length:          562
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-731658 (1-562)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||:::|||||||||
Db        487 TCTCTGATTTCCTGGAATATT 507

RESULT 61
US-09-925-065A-122265
; Sequence 122265, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 122265
;   LENGTH: 567
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-122265

Alignment Scores:
Pred. No.:                2.02e+03        Length:          567
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-122265 (1-567)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||:::|||||||||
Db        183 TCACTTATTAGCTGGAATATA 203

RESULT 62
US-09-925-065A-454651/c
; Sequence 454651, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
```

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 454651
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-454651

Alignment Scores:

Pred. No.:	2.07e+03	Length:	578
Score:	34.00	Matches:	4
Percent Similarity:	100.0%	Conservative:	3
Best Local Similarity:	57.1%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-454651 (1-578)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||::::::|||||||||:::
Db         62 TCCATAGTGACTTGGAATGTA 42
```

RESULT 63
US-09-925-065A-454652/c
; Sequence 454652, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 454652
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-454652

Alignment Scores:

Pred. No.:	2.07e+03	Length:	578
Score:	34.00	Matches:	4
Percent Similarity:	100.0%	Conservative:	3
Best Local Similarity:	57.1%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-454652 (1-578)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||::::::|||||||||:::
Db         62 TCCATAGTGACTTGGAATGTA 42
```

RESULT 64
US-09-925-065A-454653/c
; Sequence 454653, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 454653
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-454653

Alignment Scores:

Pred. No.:	2.07e+03	Length:	578
Score:	34.00	Matches:	4
Percent Similarity:	100.0%	Conservative:	3
Best Local Similarity:	57.1%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-454653 (1-578)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||::::::|||||||||:::
Db         62 TCCATAGTGACTTGGAATGTA 42
```

RESULT 65
US-09-925-065A-126922/c
; Sequence 126922, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 126922
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-126922

Alignment Scores:

Pred. No.:	2.07e+03	Length:	579
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-126922 (1-579)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db        572 TCTTTGATAACATGGAAC 555

RESULT 66
US-09-925-065A-126923/c
; Sequence 126923, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 126923
;   LENGTH: 579
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-126923

Alignment Scores:
Pred. No.:                2.07e+03        Length:         579
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-126923 (1-579)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db        572 TCTTTGATAACATGGAAC 555

RESULT 67
US-09-925-065A-782074/c
; Sequence 782074, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 782074
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-782074

Alignment Scores:
Pred. No.:                2.08e+03        Length:         580
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-782074 (1-580)

Qy          2 LeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        356 CTCATCACATGGAATATT 339

RESULT 68
US-09-925-065A-782075/c
; Sequence 782075, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 782075
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-782075

Alignment Scores:
Pred. No.:                2.08e+03        Length:         580
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-782075 (1-580)

Qy          2 LeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        356 CTCATCACATGGAATATT 339

RESULT 69
US-09-925-065A-842988/c
; Sequence 842988, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
```

```
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 842988
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-842988

Alignment Scores:
Pred. No.:                2.08e+03        Length:          580
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-842988 (1-580)

Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db        356 CTCATCACATGGAATATT 339

RESULT 70
US-09-925-065A-322599
; Sequence 322599, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 322599
;   LENGTH: 583
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-322599

Alignment Scores:
Pred. No.:                2.09e+03        Length:          583
Score:                    34.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-322599 (1-583)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         78 TCACTTATTACATGGAAC 95

RESULT 71
US-09-925-065A-103578/c
; Sequence 103578, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103578
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-103578

Alignment Scores:
Pred. No.:                2.1e+03         Length:          585
Score:                    34.00           Matches:         5
Percent Similarity:       100.0%          Conservative:    2
Best Local Similarity:    71.4%           Mismatches:      0
Query Match:              89.5%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-103578 (1-585)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::||||||||||||||::::
Db        567 ACATTAATAACTTGGAATGTA 547

RESULT 72
US-09-925-065A-460308
; Sequence 460308, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 460308
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-460308

Alignment Scores:
```

```
Pred. No.:              2.1e+03         Length:         585
Score:                  34.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            89.5%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-460308 (1-585)

Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db        195 CTGATTACCTGGAATATT 212

RESULT 73
US-09-925-065A-946652/c
; Sequence 946652, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 946652
;   LENGTH: 587
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-946652

Alignment Scores:
Pred. No.:              2.11e+03        Length:         587
Score:                  34.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            89.5%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-946652 (1-587)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db         76 TCTCTGATCACATGGAAT 59

RESULT 74
US-09-925-065A-215413/c
; Sequence 215413, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 215413
;   LENGTH: 588
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-215413

Alignment Scores:
Pred. No.:              2.11e+03        Length:         588
Score:                  34.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            89.5%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-215413 (1-588)

Qy          2 LeuIleThrTrpAsnIle 7
              ||||||||||||||||||
Db        418 TTGATTACATGGAATATC 401

RESULT 75
US-09-925-065A-822492/c
; Sequence 822492, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 822492
;   LENGTH: 588
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-822492

Alignment Scores:
Pred. No.:              2.11e+03        Length:         588
Score:                  34.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            89.5%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-822492 (1-588)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::||||||||:::
Db        454 TCCCTACTCACATGGAACCTA 434


Search completed: March 11, 2006, 12:42:01
Job time : 69.2581 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 00:16:23 ; Search time 36.4361 Seconds
(without alignments)
1588.688 Million cell updates/sec

Title: US-10-757-745-2_COPY_115_121
Perfect score: 38
Sequence: 1 SLITWNI 7

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170337_4279/app_query.fasta_1
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss08
-USER=US10757745_@CGN_1_1_2715_@runat_09032006_170337_4279 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	38	100.0	391	5	US-10-066-543-826	Sequence 826, App
	2	38	100.0	606	4	US-09-925-065A-783537	Sequence 783537,
	3	38	100.0	634	4	US-09-925-065A-783536	Sequence 783536,
	4	38	100.0	1296	3	US-09-925-299-170	Sequence 170, App
	5	38	100.0	1296	3	US-09-925-299-170	Sequence 170, App
	6	38	100.0	1498	4	US-09-925-065A-548649	Sequence 548649,
	7	38	100.0	1498	4	US-09-925-065A-548650	Sequence 548650,
	8	38	100.0	1498	4	US-09-925-065A-548651	Sequence 548651,
	9	38	100.0	1920	9	US-10-757-745-1	Sequence 1, Appli
	10	38	100.0	1940	9	US-10-783-271-28	Sequence 28, Appl
	11	38	100.0	1948	5	US-10-037-270-889	Sequence 889, App
	12	38	100.0	1948	6	US-10-117-722-889	Sequence 889, App
	13	38	100.0	1948	9	US-10-122-851-889	Sequence 889, App
	14	38	100.0	3152	3	US-09-981-353-88	Sequence 88, Appl
	15	38	100.0	166910	7	US-10-292-337-11	Sequence 11, Appl
	16	38	100.0	176930	7	US-10-741-601-5679	Sequence 5679, Ap
	17	38	100.0	176930	8	US-10-741-600-17725	Sequence 17725, A
	18	37	97.4	563	4	US-09-925-065A-134078	Sequence 134078,
	19	37	97.4	848	5	US-10-027-632-155656	Sequence 155656,
	20	37	97.4	848	5	US-10-027-632-155657	Sequence 155657,
	21	37	97.4	848	5	US-10-027-632-155658	Sequence 155658,
	22	37	97.4	848	5	US-10-027-632-155659	Sequence 155659,
	23	37	97.4	848	6	US-10-027-632-155656	Sequence 155656,
	24	37	97.4	848	6	US-10-027-632-155657	Sequence 155657,
	25	37	97.4	848	6	US-10-027-632-155658	Sequence 155658,
	26	37	97.4	848	6	US-10-027-632-155659	Sequence 155659,
c	27	37	97.4	1126	4	US-09-925-065A-78968	Sequence 78968, A
c	28	37	97.4	1168	6	US-10-242-355-976	Sequence 976, App
c	29	37	97.4	1168	6	US-10-242-355-977	Sequence 977, App
	30	37	97.4	2000	7	US-10-260-238-2437	Sequence 2437, Ap
	31	36	94.7	151	3	US-09-864-761-22834	Sequence 22834, A
c	32	36	94.7	194	3	US-09-783-590-9184	Sequence 9184, Ap
	33	36	94.7	201	8	US-10-719-993-23962	Sequence 23962, A
c	34	36	94.7	357	7	US-10-242-535A-20190	Sequence 20190, A
c	35	36	94.7	357	7	US-10-085-783A-20190	Sequence 20190, A
c	36	36	94.7	404	9	US-10-779-543-12112	Sequence 12112, A
	37	36	94.7	542	4	US-09-925-065A-767316	Sequence 767316,
	38	36	94.7	542	4	US-09-925-065A-767317	Sequence 767317,
c	39	36	94.7	582	9	US-10-779-543-20136	Sequence 20136, A
c	40	36	94.7	614	4	US-09-925-065A-953237	Sequence 953237,
c	41	36	94.7	842	4	US-09-925-065A-64641	Sequence 64641, A
c	42	36	94.7	842	4	US-09-925-065A-64642	Sequence 64642, A
	43	36	94.7	947	9	US-10-764-420-2753	Sequence 2753, Ap
	44	36	94.7	1271	3	US-09-263-959-257	Sequence 257, App
c	45	36	94.7	1387	8	US-10-425-115-117654	Sequence 117654,
c	46	36	94.7	1597	7	US-10-424-599-41163	Sequence 41163, A
	47	36	94.7	1630	8	US-10-739-930-3246	Sequence 3246, Ap
c	48	36	94.7	2087	6	US-10-104-047-84	Sequence 84, Appl
	49	36	94.7	2303	6	US-10-104-047-1924	Sequence 1924, Ap
	50	36	94.7	2395	5	US-10-027-632-112201	Sequence 112201,
	51	36	94.7	2395	6	US-10-027-632-112201	Sequence 112201,
c	52	36	94.7	2397	9	US-10-480-068-17	Sequence 17, Appl
c	53	36	94.7	3185	6	US-10-094-749-636	Sequence 636, App
	54	36	94.7	3185	9	US-10-450-763-27431	Sequence 27431, A
c	55	36	94.7	3492	9	US-10-480-068-16	Sequence 16, Appl
c	56	36	94.7	3556	5	US-10-140-164-3	Sequence 3, Appli
c	57	36	94.7	3556	6	US-10-073-333A-3	Sequence 3, Appli
c	58	36	94.7	3556	8	US-10-841-471-3	Sequence 3, Appli
c	59	36	94.7	3758	6	US-10-353-690-57	Sequence 57, Appl
c	60	36	94.7	3758	7	US-10-647-426-4	Sequence 4, Appli
c	61	36	94.7	3758	9	US-10-480-068-14	Sequence 14, Appl
c	62	36	94.7	3758	9	US-10-480-068-15	Sequence 15, Appl
	63	36	94.7	3900	9	US-10-450-763-104	Sequence 104, App
	64	36	94.7	3962	7	US-10-363-616-75	Sequence 75, Appl
c	65	36	94.7	16487	8	US-10-719-993-7014	Sequence 7014, Ap
c	66	36	94.7	23695	8	US-10-781-362-13	Sequence 13, Appl
c	67	36	94.7	23702	8	US-10-781-362-9	Sequence 9, Appli
c	68	36	94.7	23702	8	US-10-781-362-11	Sequence 11, Appl
c	69	36	94.7	23703	8	US-10-781-362-5	Sequence 5, Appli
c	70	36	94.7	23703	8	US-10-781-362-7	Sequence 7, Appli
c	71	36	94.7	23703	8	US-10-781-362-15	Sequence 15, Appl
c	72	36	94.7	23703	8	US-10-781-362-17	Sequence 17, Appl
c	73	36	94.7	23703	8	US-10-781-362-19	Sequence 19, Appl
c	74	36	94.7	30232	10	US-11-013-314-19	Sequence 19, Appl
c	75	36	94.7	39344	3	US-09-973-278-834	Sequence 834, App

ALIGNMENTS

RESULT 1

US-10-066-543-826/c
; Sequence 826, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 826
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 373
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-826

Alignment Scores:

Pred. No.:	701	Length:	391
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-066-543-826 (1-391)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        103 TCTCTCATTACCTGGAATATT 83
```

RESULT 2
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-783537

Alignment Scores:

Pred. No.:	1.12e+03	Length:	606
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-783537 (1-606)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        504 TCTCTCATTACCTGGAATATT 524
```

RESULT 3
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-783536

Alignment Scores:

Pred. No.:	1.18e+03	Length:	634
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-783536 (1-634)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        355 TCTCTCATTACCTGGAATATT 375
```

RESULT 4
US-09-925-299-170
; Sequence 170, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12

```
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;    LENGTH: 1296
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1261)
;    OTHER INFORMATION: n equals a,t,g, or c
;    NAME/KEY: misc_feature
;    LOCATION: (1276)
;    OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:              2.52e+03        Length:         1296
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-299-170 (1-1296)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        366 TCTCTCATTACCTGGAATATT 386

RESULT 5
US-09-925-299-170
; Sequence 170, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;    LENGTH: 1296
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1261)
;    OTHER INFORMATION: n equals a,t,g, or c
;    NAME/KEY: misc_feature
;    LOCATION: (1276)
;    OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:              2.52e+03        Length:         1296
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-299-170 (1-1296)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        366 TCTCTCATTACCTGGAATATT 386

RESULT 6
US-09-925-065A-548649
; Sequence 548649, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548649
;    LENGTH: 1498
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-548649

Alignment Scores:
Pred. No.:              2.94e+03        Length:         1498
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     4               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-548649 (1-1498)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db       1402 TCATTAATAACTTGGAACATT 1422

RESULT 7
US-09-925-065A-548650
; Sequence 548650, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548650
;    LENGTH: 1498
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-548650

Alignment Scores:
Pred. No.:              2.94e+03        Length:         1498
```

Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-548650 (1-1498)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db       1402 TCATTAATAACTTGGAACATT 1422
```

RESULT 8
US-09-925-065A-548651
; Sequence 548651, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548651
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-548651

Alignment Scores:			
Pred. No.:	2.94e+03	Length:	1498
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-548651 (1-1498)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db       1402 TCATTAATAACTTGGAACATT 1422
```

RESULT 9
US-10-757-745-1
; Sequence 1, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/10/757,745
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/697,863A
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1627)..(1627)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1108)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

Alignment Scores:			
Pred. No.:	3.84e+03	Length:	1920
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-757-745-1 (1-1920)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        362 TCTCTCATTACCTGGAATATT 382
```

RESULT 10
US-10-783-271-28
; Sequence 28, Application US/10783271
; Publication No. US20050186577A1
; GENERAL INFORMATION:
; APPLICANT: Veridex, LLC
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
; FILE REFERENCE: VDX-5003 USNP
; CURRENT APPLICATION NUMBER: US/10/783,271
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: human
US-10-783-271-28

Alignment Scores:			
Pred. No.:	3.88e+03	Length:	1940
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-783-271-28 (1-1940)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        368 TCTCTCATTACCTGGAATATT 388
```

RESULT 11
US-10-037-270-889
; Sequence 889, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-037-270-889

Alignment Scores:

Pred. No.:	3.9e+03	Length:	1948
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-037-270-889 (1-1948)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        381 TCTCTCATTACCTGGAATATT 401
```

RESULT 12
US-10-117-722-889
; Sequence 889, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-117-722-889

Alignment Scores:

Pred. No.:	3.9e+03	Length:	1948
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-117-722-889 (1-1948)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        381 TCTCTCATTACCTGGAATATT 401
```

RESULT 13
US-10-122-851-889
; Sequence 889, Application US/10122851
; Publication No. US20050239060A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BDV3
; CURRENT APPLICATION NUMBER: US/10/122,851
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-122-851-889

Alignment Scores:

Pred. No.:	3.9e+03	Length:	1948
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-122-851-889 (1-1948)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        381 TCTCTCATTACCTGGAATATT 401
```

RESULT 14
US-09-981-353-88
; Sequence 88, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.

; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 88
; LENGTH: 3152
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 232992.1
; NAME/KEY: unsure
; LOCATION: 1171
; OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

Alignment Scores:

Pred. No.:	6.52e+03	Length:	3152
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-981-353-88 (1-3152)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db       1594 TCTCTCATTACCTGGAATATT 1614
```

RESULT 15
US-10-292-337-11
; Sequence 11, Application US/10292337
; Publication No. US20040092462A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Nicholas M. Dean
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: MODULATION OF ENDOTHELIAL DIFFERENTIATION GENE 2 EXPRESSION
; FILE REFERENCE: PTS-0058
; CURRENT APPLICATION NUMBER: US/10/292,337
; CURRENT FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 139
; SEQ ID NO 11
; LENGTH: 166910
; TYPE: DNA
; ORGANISM: H. sapiens
US-10-292-337-11

Alignment Scores:

Pred. No.:	4.47e+05	Length:	166910
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-292-337-11 (1-166910)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      64005 TCTTTGATCACATGGAATATT 64025
```

RESULT 16
US-10-741-601-5679
; Sequence 5679, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5679
; LENGTH: 176930
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-5679

Alignment Scores:

Pred. No.:	4.76e+05	Length:	176930
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-741-601-5679 (1-176930)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      68596 TCTTTGATCACATGGAATATT 68616
```

RESULT 17
US-10-741-600-17725
; Sequence 17725, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17725
; LENGTH: 176930
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17725

Alignment Scores:

Pred. No.:	4.76e+05	Length:	176930
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-741-600-17725 (1-176930)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      68596 TCTTTGATCACATGGAATATT 68616
```

RESULT 18
US-09-925-065A-134078
; Sequence 134078, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147

```
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134078
;   LENGTH: 563
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-134078

Alignment Scores:
Pred. No.:              1.6e+03       Length:          563
Score:                  37.00         Matches:         6
Percent Similarity:     100.0%        Conservative:    1
Best Local Similarity:  85.7%         Mismatches:      0
Query Match:            97.4%         Indels:          0
DB:                     4             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-134078 (1-563)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        532 TCCCTAGTCACATGGAATATT 552

RESULT 19
US-10-027-632-155656
; Sequence 155656, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155656
;   LENGTH: 848
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-155656

Alignment Scores:
Pred. No.:              2.49e+03      Length:          848
Score:                  37.00         Matches:         6
Percent Similarity:     100.0%        Conservative:    1
Best Local Similarity:  85.7%         Mismatches:      0
Query Match:            97.4%         Indels:          0
DB:                     5             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-027-632-155656 (1-848)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        633 AGTCTAATCACATGGAATGTT 653

RESULT 20
US-10-027-632-155657
; Sequence 155657, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155657
;   LENGTH: 848
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-155657

Alignment Scores:
Pred. No.:              2.49e+03      Length:          848
Score:                  37.00         Matches:         6
Percent Similarity:     100.0%        Conservative:    1
Best Local Similarity:  85.7%         Mismatches:      0
Query Match:            97.4%         Indels:          0
DB:                     5             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-027-632-155657 (1-848)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        633 AGTCTAATCACATGGAATGTT 653

RESULT 21
US-10-027-632-155658
; Sequence 155658, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
```

```
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155658
;    LENGTH: 848
;    TYPE: DNA
;    ORGANISM: Human
US-10-027-632-155658

Alignment Scores:
Pred. No.:                2.49e+03        Length:          848
Score:                    37.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              97.4%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-027-632-155658 (1-848)

Qy           1 SerLeuIleThrTrpAsnIle 7
                ||||||||||||||||||:::
Db         633 AGTCTAATCACATGGAATGTT 653

RESULT 22
US-10-027-632-155659
; Sequence 155659, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155659
;    LENGTH: 848
;    TYPE: DNA
;    ORGANISM: Human
US-10-027-632-155659

Alignment Scores:
Pred. No.:                2.49e+03        Length:          848
Score:                    37.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              97.4%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-027-632-155659 (1-848)

Qy           1 SerLeuIleThrTrpAsnIle 7
                ||||||||||||||||||:::
Db         633 AGTCTAATCACATGGAATGTT 653

RESULT 23
US-10-027-632-155656
; Sequence 155656, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155656
;    LENGTH: 848
;    TYPE: DNA
;    ORGANISM: Human
US-10-027-632-155656

Alignment Scores:
Pred. No.:                2.49e+03        Length:          848
Score:                    37.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              97.4%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-027-632-155656 (1-848)

Qy           1 SerLeuIleThrTrpAsnIle 7
                ||||||||||||||||||:::
Db         633 AGTCTAATCACATGGAATGTT 653

RESULT 24
US-10-027-632-155657
; Sequence 155657, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155657
```

```
;   LENGTH: 848
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-155657

Alignment Scores:
Pred. No.:                2.49e+03        Length:         848
Score:                    37.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              97.4%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-027-632-155657 (1-848)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        633 AGTCTAATCACATGGAATGTT 653

RESULT 25
US-10-027-632-155658
; Sequence 155658, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155658
;   LENGTH: 848
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-155658

Alignment Scores:
Pred. No.:                2.49e+03        Length:         848
Score:                    37.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              97.4%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-027-632-155658 (1-848)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        633 AGTCTAATCACATGGAATGTT 653

RESULT 26
US-10-027-632-155659
; Sequence 155659, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155659
;   LENGTH: 848
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-155659

Alignment Scores:
Pred. No.:                2.49e+03        Length:         848
Score:                    37.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              97.4%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-027-632-155659 (1-848)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        633 AGTCTAATCACATGGAATGTT 653

RESULT 27
US-09-925-065A-78968/c
; Sequence 78968, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78968
;   LENGTH: 1126
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-78968

Alignment Scores:
Pred. No.:                3.36e+03        Length:         1126
Score:                    37.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
```

```
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            97.4%          Indels:         0
DB:                     4              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-78968 (1-1126)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        876 TCACTTGTAACTTGGAATATT 856

RESULT 28
US-10-242-355-976/c
; Sequence 976, Application US/10242355
; Publication No. US20030235831A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC003C1
;  CURRENT APPLICATION NUMBER: US/10/242,355
;  CURRENT FILING DATE:  2002-09-13
;  PRIOR APPLICATION NUMBER: 09/764,897
;  PRIOR FILING DATE: 2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1267
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 976
;    LENGTH: 1168
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-242-355-976

Alignment Scores:
Pred. No.:              3.5e+03        Length:         1168
Score:                  37.00          Matches:        6
Percent Similarity:     100.0%         Conservative:   1
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            97.4%          Indels:         0
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-242-355-976 (1-1168)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       1014 TCTCTTGTTACCTGGAATATT 994

RESULT 29
US-10-242-355-977/c
; Sequence 977, Application US/10242355
; Publication No. US20030235831A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC003C1
;  CURRENT APPLICATION NUMBER: US/10/242,355
;  CURRENT FILING DATE:  2002-09-13
;  PRIOR APPLICATION NUMBER: 09/764,897
;  PRIOR FILING DATE: 2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1267
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 977
;    LENGTH: 1168
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-242-355-977

Alignment Scores:
Pred. No.:              3.5e+03        Length:         1168
Score:                  37.00          Matches:        6
Percent Similarity:     100.0%         Conservative:   1
Best Local Similarity:  85.7%          Mismatches:     0
Query Match:            97.4%          Indels:         0
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-242-355-977 (1-1168)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       1014 TCTCTTGTTACCTGGAATATT 994

RESULT 30
US-10-260-238-2437
; Sequence 2437, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
;  APPLICANT: Budworth, Paul R.
;  APPLICANT:  Moughamer, Todd G.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Cooper, Bret
;  APPLICANT:  Glazebrook, Jane
;  APPLICANT:  Goff, Stephen A.
;  APPLICANT:  Katagiri, Fumiyaki
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Provart, Nicholas
;  APPLICANT:  Ricke, Darrell
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
;  FILE REFERENCE: 60111-NP
;  CURRENT APPLICATION NUMBER: US/10/260,238
;  CURRENT FILING DATE:  2002-09-26
;  PRIOR APPLICATION NUMBER: US 60/325,448
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: US 60/325,277
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: US 60/370,620
;  PRIOR FILING DATE: 2002-04-04
;  NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 2437
;    LENGTH: 2000
;    TYPE: DNA
```

; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (421)..(421)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (452)..(452)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (454)..(455)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N_region
; LOCATION: (465)..(465)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-2437

Alignment Scores:

Pred. No.:	6.21e+03	Length:	2000
Score:	37.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-260-238-2437 (1-2000)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db       1442 TCGTTAGTTACCTGGAACATT 1462
```

RESULT 31
US-09-864-761-22834
; Sequence 22834, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22834
; LENGTH: 151
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002081.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
; OTHER INFORMATION: SWISSPROT HIT: Q59295, EVALUE 2.30e+00
; OTHER INFORMATION: NT HIT: AL163252.2, EVALUE 1.40e+00
; OTHER INFORMATION: EST_HUMAN HIT: R80873.1, EVALUE 5.10e-01
US-09-864-761-22834

Alignment Scores:

Pred. No.:	610	Length:	151
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-864-761-22834 (1-151)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        120 TCATTACTAACTTGGAATATT 140
```

RESULT 32
US-09-783-590-9184/c
; Sequence 9184, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9184
; LENGTH: 194
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)
; OTHER INFORMATION: n equals a,t,g, or c

```
;   NAME/KEY: misc feature
;   LOCATION: (39)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (47)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (70)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (102)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (108)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (185)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9184

Alignment Scores:
Pred. No.:                797             Length:         194
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-783-590-9184 (1-194)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        157 AGCATTATTACGTGGAACATT 137

RESULT 33
US-10-719-993-23962
; Sequence 23962, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001496
;  CURRENT APPLICATION NUMBER: US/10/719,993
;  CURRENT FILING DATE:  2003-11-24
;  NUMBER OF SEQ ID NOS: 55342
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23962
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-719-993-23962

Alignment Scores:
Pred. No.:                827             Length:         201
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-719-993-23962 (1-201)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        116 TCATTACTAACATGGAATATA 136

RESULT 34
US-10-242-535A-20190/c
; Sequence 20190, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20190
;   LENGTH: 357
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (39)..(39)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (41)..(41)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (99)..(99)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (130)..(130)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (141)..(141)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (226)..(226)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (264)..(264)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (318)..(318)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-20190

Alignment Scores:
Pred. No.:                1.53e+03        Length:         357
Score:                    36.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-242-535A-20190 (1-357)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        301 AGTCTTGTGACCTGGAACGTG 281

RESULT 35
US-10-085-783A-20190/c
```

; Sequence 20190, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20190
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (39)..(39)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (99)..(99)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (130)..(130)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (141)..(141)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (226)..(226)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (264)..(264)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (318)..(318)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-20190

Alignment Scores:

Pred. No.:	1.53e+03	Length:	357
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-085-783A-20190 (1-357)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db        301 AGTCTTGTGACCTGGAACGTG 281
```

RESULT 36
US-10-779-543-12112/c
; Sequence 12112, Application US/10779543
; Publication No. US20050227917A1
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12112
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 404
; OTHER INFORMATION: n = A,T,C or G
US-10-779-543-12112

Alignment Scores:

Pred. No.:	1.74e+03	Length:	404
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-779-543-12112 (1-404)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        230 TCATTACTAACTTGGAATATT 210
```

RESULT 37
US-09-925-065A-767316
; Sequence 767316, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20

; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 767316
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-767316

Alignment Scores:

Pred. No.:	2.39e+03	Length:	542
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-767316 (1-542)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        475 TCCATCATCACATGGAACATT 495
```

RESULT 38
US-09-925-065A-767317
; Sequence 767317, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 767317
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-767317

Alignment Scores:

Pred. No.:	2.39e+03	Length:	542
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-767317 (1-542)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        475 TCCATCATCACATGGAACATT 495
```

RESULT 39
US-10-779-543-20136/c
; Sequence 20136, Application US/10779543
; Publication No. US20050227917A1
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20136
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-779-543-20136

Alignment Scores:

Pred. No.:	2.58e+03	Length:	582
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-779-543-20136 (1-582)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        436 TCTTTACTTACCTGGAATATT 416
```

RESULT 40
US-09-925-065A-953237/c
; Sequence 953237, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09

```
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 953237
;    LENGTH: 614
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-953237

Alignment Scores:
Pred. No.:                2.73e+03        Length:         614
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       4               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-953237 (1-614)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||||
Db        515 TCCATCATCACATGGAACATT 495

RESULT 41
US-09-925-065A-64641/c
; Sequence 64641, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64641
;    LENGTH: 842
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-64641

Alignment Scores:
Pred. No.:                3.82e+03        Length:         842
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       4               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-64641 (1-842)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        161 TCATTGATTACCTGGAATTTA 141

RESULT 42
US-09-925-065A-64642/c
; Sequence 64642, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64642
;    LENGTH: 842
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-64642

Alignment Scores:
Pred. No.:                3.82e+03        Length:         842
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       4               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-925-065A-64642 (1-842)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        161 TCATTGATTACCTGGAATTTA 141

RESULT 43
US-10-764-420-2753
; Sequence 2753, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
;  APPLICANT: Lum, Pek Yee
;  APPLICANT:  Tan, Yejun
;  APPLICANT:  Dai, Hongyue
;  TITLE OF INVENTION: Methods For Determining Whether An Agent
;  TITLE OF INVENTION:  Possesses A Defined Biological Activity
;  FILE REFERENCE: ROSA122057
;  CURRENT APPLICATION NUMBER: US/10/764,420
;  CURRENT FILING DATE:  2004-01-23
;  PRIOR APPLICATION NUMBER: US 60/442,797
;  PRIOR FILING DATE: 2003-01-24
;  PRIOR APPLICATION NUMBER: US 60/474,413
;  PRIOR FILING DATE: 2003-05-30
;  NUMBER OF SEQ ID NOS: 3683
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2753
;    LENGTH: 947
;    TYPE: DNA
;    ORGANISM: Mus musculus
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: 595, 748, 788, 817
;    OTHER INFORMATION: n = A,T,C or G
US-10-764-420-2753

Alignment Scores:
Pred. No.:                4.33e+03        Length:         947
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-764-420-2753 (1-947)
```

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||||||
Db        429 TCTTTACTCACTTGGAATATT 449

RESULT 44
US-09-263-959-257
; Sequence 257, Application US/09263959
; Patent No. US20020150891A1
;  GENERAL INFORMATION:
;    APPLICANT:  Hood, Leroy E.
;    APPLICANT:  Rowen, Lee
;    APPLICANT:  Koop, Ben F.
;    TITLE OF INVENTION:  DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI
;    NUMBER OF SEQUENCES:  1279
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/263,959
;      FILING DATE:  05-MAR-1999
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  920010.426C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  257:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1271 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-263-959-257

Alignment Scores:
Pred. No.:                5.93e+03        Length:          1271
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-263-959-257 (1-1271)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        266 TCTTTAATTACCTGGAATTTG 286

RESULT 45
US-10-425-115-117654/c
; Sequence 117654, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 117654
;   LENGTH: 1387
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1387)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_3878C.1
US-10-425-115-117654

Alignment Scores:
Pred. No.:                6.51e+03        Length:          1387
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-425-115-117654 (1-1387)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::||||||||||||||||
Db        758 TCGATAATAACATGGAACATT 738

RESULT 46
US-10-424-599-41163/c
; Sequence 41163, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 41163
;   LENGTH: 1597
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1597)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_137171C.1
US-10-424-599-41163

Alignment Scores:
Pred. No.:                7.57e+03        Length:          1597
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-424-599-41163 (1-1597)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       1199 AGCCTCTTAACCTGGAACATT 1179

RESULT 47
US-10-739-930-3246
; Sequence 3246, Application US/10739930
; Publication No. US20040216190A1
```

; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 3246
; LENGTH: 1630
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: GLYMA-23APR03-CLUSTER17438_1
US-10-739-930-3246

Alignment Scores:

Pred. No.:	7.74e+03	Length:	1630
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-739-930-3246 (1-1630)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||||||
Db        485 AGCCTCTTAACCTGGAACATT 505
```

RESULT 48
US-10-104-047-84/c
; Sequence 84, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-84

Alignment Scores:

Pred. No.:	1.01e+04	Length:	2087
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-104-047-84 (1-2087)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::||||||||||||||||
Db        164 AGCATTATTACGTGGAACATC 144
```

RESULT 49
US-10-104-047-1924
; Sequence 1924, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1924
; LENGTH: 2303
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1924

Alignment Scores:

Pred. No.:	1.12e+04	Length:	2303
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-104-047-1924 (1-2303)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::||||||||||:::
Db        993 AGCTTAGTAACTTGGAATGTT 1013
```

RESULT 50
US-10-027-632-112201
; Sequence 112201, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 112201
; LENGTH: 2395
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-112201

Alignment Scores:

Pred. No.:	1.17e+04	Length:	2395
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-027-632-112201 (1-2395)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::||||||||||||||||
Db        860 TCTATCATTACTTGGAATATA 880
```

RESULT 51

```
US-10-027-632-112201
; Sequence 112201, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 112201
;   LENGTH: 2395
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-112201

Alignment Scores:
Pred. No.:                1.17e+04        Length:          2395
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-027-632-112201 (1-2395)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        860 TCTATCATTACTTGGAATATA 880

RESULT 52
US-10-480-068-17/c
; Sequence 17, Application US/10480068
; Publication No. US20050112568A1
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: DGKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
;  FILE REFERENCE: EX02-079
;  CURRENT APPLICATION NUMBER: US/10/480,068
;  CURRENT FILING DATE:  2003-12-04
;  PRIOR APPLICATION NUMBER: US 60/296,076
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/328,605
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/338,733
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/357,253
;  PRIOR FILING DATE: 2002-02-15
;  PRIOR APPLICATION NUMBER: US 60/357,600
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
;   LENGTH: 2397
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-480-068-17

Alignment Scores:
Pred. No.:                1.17e+04        Length:          2397
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-480-068-17 (1-2397)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       1810 TCTCTCATCACATGGAATCTG 1790

RESULT 53
US-10-094-749-636/c
; Sequence 636, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;  FILE REFERENCE: 084335/0160
;  CURRENT APPLICATION NUMBER: US/10/094,749
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 60/350,435
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: JP 2001-328381
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 3381
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 636
;   LENGTH: 3185
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-094-749-636

Alignment Scores:
Pred. No.:                1.58e+04        Length:          3185
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-094-749-636 (1-3185)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db       2536 TCATTACTAACTTGGAATATT 2516

RESULT 54
US-10-450-763-27431
; Sequence 27431, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 27431
;   LENGTH: 3185
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1018)..(1125)
;   OTHER INFORMATION: 78% homologous to Homo sapiens Peptide sequence derived from
;   OTHER INFORMATION: a human secreted protein,accession number Y40039,Smith-Waterman
;   OTHER INFORMATION: Score=142.
US-10-450-763-27431

Alignment Scores:
Pred. No.:                1.58e+04        Length:         3185
Score:                    36.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-450-763-27431 (1-3185)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db       1454 AGCTTAGTAACTTGGAATGTT 1474

RESULT 55
US-10-480-068-16/c
; Sequence 16, Application US/10480068
; Publication No. US20050112568A1
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: DGKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
;  FILE REFERENCE: EX02-079
;  CURRENT APPLICATION NUMBER: US/10/480,068
;  CURRENT FILING DATE:  2003-12-04
;  PRIOR APPLICATION NUMBER: US 60/296,076
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/328,605
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/338,733
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/357,253
;  PRIOR FILING DATE: 2002-02-15
;  PRIOR APPLICATION NUMBER: US 60/357,600
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
;   LENGTH: 3492
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-480-068-16

Alignment Scores:
Pred. No.:                1.75e+04        Length:         3492
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-480-068-16 (1-3492)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db       2049 TCTCTCATCACATGGAATCTG 2029

RESULT 56
US-10-140-164-3/c
; Sequence 3, Application US/10140164
; Publication No. US20030072736A1
; GENERAL INFORMATION:
;  APPLICANT: Baker et al.
;  TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor TR16
;  FILE REFERENCE: PF514C1
;  CURRENT APPLICATION NUMBER: US/10/140,164
;  CURRENT FILING DATE:  2002-05-08
;  PRIOR APPLICATION NUMBER: 09/637,856
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/148,348
;  PRIOR FILING DATE: 1999-08-12
;  PRIOR APPLICATION NUMBER: 60/148,683
;  PRIOR FILING DATE: 1999-08-13
;  PRIOR APPLICATION NUMBER: 60/148,870
;  PRIOR FILING DATE: 1999-08-13
;  PRIOR APPLICATION NUMBER: 60/148,758
;  PRIOR FILING DATE: 1999-08-16
;  PRIOR APPLICATION NUMBER: 60/149,181
;  PRIOR FILING DATE: 1999-08-17
;  PRIOR APPLICATION NUMBER: 60/149,453
;  PRIOR FILING DATE: 1999-08-18
;  PRIOR APPLICATION NUMBER: 60/149,498
;  PRIOR FILING DATE: 1999-08-19
;  NUMBER OF SEQ ID NOS: 76
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 3556
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(3084)
US-10-140-164-3

Alignment Scores:
Pred. No.:                1.78e+04        Length:         3556
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-140-164-3 (1-3556)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||
Db       2890 TCATTACTAACTTGGAATATT 2870

RESULT 57
US-10-073-333A-3/c
; Sequence 3, Application US/10073333A
; Publication No. US20030134788A1
; GENERAL INFORMATION:
;  APPLICANT: Kevin Baker et al.
;  TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR RECEPTOR TR16
;  FILE REFERENCE: PF514P1
;  CURRENT APPLICATION NUMBER: US/10/073,333A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: 60/268,364
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 09/637,856
;  PRIOR FILING DATE: 2000-08-10
```

; PRIOR APPLICATION NUMBER: 60/148,348
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/148,683
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 60/148,758
; PRIOR FILING DATE: 1999-08-16
; PRIOR APPLICATION NUMBER: 60/148,870
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 60/149,181
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149,453
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/149,498
; PRIOR FILING DATE: 1999-08-18
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3556
; TYPE: DNA
; ORGANISM: human
US-10-073-333A-3

Alignment Scores:

Pred. No.:	1.78e+04	Length:	3556
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-073-333A-3 (1-3556)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       2890 TCATTACTAACTTGGAATATT 2870
```

RESULT 58
US-10-841-471-3/c
; Sequence 3, Application US/10841471
; Publication No. US20040209328A1
; GENERAL INFORMATION:
; APPLICANT: Baker et al.
; TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR RECEPTOR TR16
; FILE REFERENCE: PF514P1C1
; CURRENT APPLICATION NUMBER: US/10/841,471
; CURRENT FILING DATE: 2004-05-10
; PRIOR APPLICATION NUMBER: US 10/073,333
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/268,364
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,856
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 60/148,348
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: US 60/148,683
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: US 60/148,758
; PRIOR FILING DATE: 1999-08-16
; PRIOR APPLICATION NUMBER: US 60/148,870
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: US 60/149,181
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: US 60/149,453
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: US 60/149,498
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 3556
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-841-471-3

Alignment Scores:

Pred. No.:	1.78e+04	Length:	3556
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-841-471-3 (1-3556)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       2890 TCATTACTAACTTGGAATATT 2870
```

RESULT 59
US-10-353-690-57/c
; Sequence 57, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary
; APPLICANT: Stagliano, Nancy
; APPLICANT: Perodin, Jacquelin
; APPLICANT: Rodrigue-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
; FILE REFERENCE: MPI02-018P1RNOMNIM
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 3758
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-353-690-57

Alignment Scores:

Pred. No.:	1.89e+04	Length:	3758
Score:	36.00	Matches:	6

```
Percent Similarity:     100.0%          Conservative:  1
Best Local Similarity:  85.7%           Mismatches:    0
Query Match:            94.7%           Indels:        0
DB:                     6               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-353-690-57 (1-3758)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2315 TCTCTCATCACATGGAATCTG 2295

RESULT 60
US-10-647-426-4/c
; Sequence 4, Application US/10647426
; Publication No. US20040110197A1
; GENERAL INFORMATION:
;  APPLICANT: Skinner, Michael K.
;  APPLICANT:  Patton, Jodi L.
;  TITLE OF INVENTION: A METHOD OF DETERMINING TUMOR CHARACTERISTICS BY
;  TITLE OF INVENTION:  DETERMINING ABNORMAL COPY NUMBER OR EXPRESSION LEVEL OF
;  TITLE OF INVENTION:  LIPID-ASSOCIATED GENES
;  FILE REFERENCE: PATRICK EAGLEMAN: EMBOL-X 252/124
;  CURRENT APPLICATION NUMBER: US/10/647,426
;  CURRENT FILING DATE:  2003-08-26
;  PRIOR APPLICATION NUMBER: US/09/676,052
;  PRIOR FILING DATE: 2000-09-28
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 3758
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: gene
;   LOCATION: (1)..(3758)
;   OTHER INFORMATION: The sequence of the cDNA coding for Diacylglycerol
;   OTHER INFORMATION: kinase, gamma
US-10-647-426-4

Alignment Scores:
Pred. No.:              1.89e+04        Length:        3758
Score:                  36.00           Matches:       6
Percent Similarity:     100.0%          Conservative:  1
Best Local Similarity:  85.7%           Mismatches:    0
Query Match:            94.7%           Indels:        0
DB:                     7               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-647-426-4 (1-3758)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2315 TCTCTCATCACATGGAATCTG 2295

RESULT 61
US-10-480-068-14/c
; Sequence 14, Application US/10480068
; Publication No. US20050112568A1
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: DGKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
;  FILE REFERENCE: EX02-079
;  CURRENT APPLICATION NUMBER: US/10/480,068
;  CURRENT FILING DATE:  2003-12-04
;  PRIOR APPLICATION NUMBER: US 60/296,076
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/328,605
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/338,733
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/357,253
;  PRIOR FILING DATE: 2002-02-15
;  PRIOR APPLICATION NUMBER: US 60/357,600
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 3758
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-480-068-14

Alignment Scores:
Pred. No.:              1.89e+04        Length:        3758
Score:                  36.00           Matches:       6
Percent Similarity:     100.0%          Conservative:  1
Best Local Similarity:  85.7%           Mismatches:    0
Query Match:            94.7%           Indels:        0
DB:                     9               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-480-068-14 (1-3758)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2315 TCTCTCATCACATGGAATCTG 2295

RESULT 62
US-10-480-068-15/c
; Sequence 15, Application US/10480068
; Publication No. US20050112568A1
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: DGKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
;  FILE REFERENCE: EX02-079
;  CURRENT APPLICATION NUMBER: US/10/480,068
;  CURRENT FILING DATE:  2003-12-04
;  PRIOR APPLICATION NUMBER: US 60/296,076
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/328,605
;  PRIOR FILING DATE: 2001-10-10
;  PRIOR APPLICATION NUMBER: US 60/338,733
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/357,253
;  PRIOR FILING DATE: 2002-02-15
;  PRIOR APPLICATION NUMBER: US 60/357,600
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
;   LENGTH: 3758
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-480-068-15

Alignment Scores:
Pred. No.:              1.89e+04        Length:        3758
Score:                  36.00           Matches:       6
Percent Similarity:     100.0%          Conservative:  1
Best Local Similarity:  85.7%           Mismatches:    0
Query Match:            94.7%           Indels:        0
DB:                     9               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-480-068-15 (1-3758)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2315 TCTCTCATCACATGGAATCTG 2295

RESULT 63
US-10-450-763-104
; Sequence 104, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
```

```
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 104
;   LENGTH: 3900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (370)..(1314)
;   OTHER INFORMATION: 55% homologous to Rattus norvegicus tumor suppressor
;   OTHER INFORMATION: pHyde,accession number AF335281,Smith-Waterman Score=929.
US-10-450-763-104

Alignment Scores:
Pred. No.:                1.97e+04        Length:          3900
Score:                    36.00           Matches:         5
Percent Similarity:       100.0%          Conservative:    2
Best Local Similarity:    71.4%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-450-763-104 (1-3900)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db       2351 AGCTTAGTAACTTGGAATGTT 2371

RESULT 64
US-10-363-616-75
; Sequence 75, Application US/10363616
; Publication No. US20040044181A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-113 (793)
;  CURRENT APPLICATION NUMBER: US/10/363,616
;  CURRENT FILING DATE:  2003-03-03
;  PRIOR APPLICATION NUMBER: 09/654,935
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 490
; SEQ ID NO 75
;   LENGTH: 3962
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (652)..(3900)
US-10-363-616-75

Alignment Scores:
Pred. No.:                2e+04           Length:          3962
Score:                    36.00           Matches:         5
Percent Similarity:       100.0%          Conservative:    2
Best Local Similarity:    71.4%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-363-616-75 (1-3962)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db       2351 AGCTTAGTAACTTGGAATGTT 2371

RESULT 65
US-10-719-993-7014/c
; Sequence 7014, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001496
;  CURRENT APPLICATION NUMBER: US/10/719,993
;  CURRENT FILING DATE:  2003-11-24
;  NUMBER OF SEQ ID NOS: 55342
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7014
;   LENGTH: 16487
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-719-993-7014

Alignment Scores:
Pred. No.:                9.17e+04        Length:          16487
Score:                    36.00           Matches:         5
Percent Similarity:       100.0%          Conservative:    2
Best Local Similarity:    71.4%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-719-993-7014 (1-16487)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db      11304 TCCTTAGTGACATGGAATGTG 11284

RESULT 66
US-10-781-362-13/c
; Sequence 13, Application US/10781362
; Publication No. US20040259157A1
; GENERAL INFORMATION:
;  APPLICANT: Ginsburg, David
;  APPLICANT:  Zhang, Bin
;  APPLICANT:  Kaufman, Randal J.
;  TITLE OF INVENTION: MCFD2 Nucleic Acids and Proteins
;  FILE REFERENCE: UM-08901
;  CURRENT APPLICATION NUMBER: US/10/781,362
;  CURRENT FILING DATE:  2004-02-18
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
;   LENGTH: 23695
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-781-362-13

Alignment Scores:
Pred. No.:                1.35e+05        Length:          23695
Score:                    36.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              94.7%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-781-362-13 (1-23695)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db        263 AGTCTCATAACTTGGAACCTT 243

RESULT 67
US-10-781-362-9/c
; Sequence 9, Application US/10781362
; Publication No. US20040259157A1
; GENERAL INFORMATION:
;  APPLICANT: Ginsburg, David
;  APPLICANT:  Zhang, Bin
;  APPLICANT:  Kaufman, Randal J.
;  TITLE OF INVENTION: MCFD2 Nucleic Acids and Proteins
```

```
;  FILE REFERENCE: UM-08901
;  CURRENT APPLICATION NUMBER: US/10/781,362
;  CURRENT FILING DATE:  2004-02-18
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
;   LENGTH: 23702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-781-362-9

Alignment Scores:
Pred. No.:              1.35e+05        Length:         23702
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-781-362-9 (1-23702)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        263 AGTCTCATAACTTGGAACCTT 243

RESULT 68
US-10-781-362-11/c
; Sequence 11, Application US/10781362
; Publication No. US20040259157A1
; GENERAL INFORMATION:
;  APPLICANT: Ginsburg, David
;  APPLICANT:  Zhang, Bin
;  APPLICANT:  Kaufman, Randal J.
;  TITLE OF INVENTION: MCFD2 Nucleic Acids and Proteins
;  FILE REFERENCE: UM-08901
;  CURRENT APPLICATION NUMBER: US/10/781,362
;  CURRENT FILING DATE:  2004-02-18
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
;   LENGTH: 23702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-781-362-11

Alignment Scores:
Pred. No.:              1.35e+05        Length:         23702
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-781-362-11 (1-23702)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        263 AGTCTCATAACTTGGAACCTT 243

RESULT 69
US-10-781-362-5/c
; Sequence 5, Application US/10781362
; Publication No. US20040259157A1
; GENERAL INFORMATION:
;  APPLICANT: Ginsburg, David
;  APPLICANT:  Zhang, Bin
;  APPLICANT:  Kaufman, Randal J.
;  TITLE OF INVENTION: MCFD2 Nucleic Acids and Proteins
;  FILE REFERENCE: UM-08901
;  CURRENT APPLICATION NUMBER: US/10/781,362
;  CURRENT FILING DATE:  2004-02-18
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
;   LENGTH: 23703
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-781-362-5

Alignment Scores:
Pred. No.:              1.35e+05        Length:         23703
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-781-362-5 (1-23703)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        263 AGTCTCATAACTTGGAACCTT 243

RESULT 70
US-10-781-362-7/c
; Sequence 7, Application US/10781362
; Publication No. US20040259157A1
; GENERAL INFORMATION:
;  APPLICANT: Ginsburg, David
;  APPLICANT:  Zhang, Bin
;  APPLICANT:  Kaufman, Randal J.
;  TITLE OF INVENTION: MCFD2 Nucleic Acids and Proteins
;  FILE REFERENCE: UM-08901
;  CURRENT APPLICATION NUMBER: US/10/781,362
;  CURRENT FILING DATE:  2004-02-18
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
;   LENGTH: 23703
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-781-362-7

Alignment Scores:
Pred. No.:              1.35e+05        Length:         23703
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-781-362-7 (1-23703)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        263 AGTCTCATAACTTGGAACCTT 243

RESULT 71
US-10-781-362-15/c
; Sequence 15, Application US/10781362
; Publication No. US20040259157A1
; GENERAL INFORMATION:
;  APPLICANT: Ginsburg, David
;  APPLICANT:  Zhang, Bin
;  APPLICANT:  Kaufman, Randal J.
;  TITLE OF INVENTION: MCFD2 Nucleic Acids and Proteins
;  FILE REFERENCE: UM-08901
;  CURRENT APPLICATION NUMBER: US/10/781,362
;  CURRENT FILING DATE:  2004-02-18
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
;   LENGTH: 23703
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-781-362-15
```

Alignment Scores:
Pred. No.: 1.35e+05 Length: 23703
Score: 36.00 Matches: 6
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 85.7% Mismatches: 0
Query Match: 94.7% Indels: 0
DB: 8 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-781-362-15 (1-23703)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        263 AGTCTCATAACTTGGAACCTT 243
```

RESULT 72
US-10-781-362-17/c
; Sequence 17, Application US/10781362
; Publication No. US20040259157A1
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Zhang, Bin
; APPLICANT: Kaufman, Randal J.
; TITLE OF INVENTION: MCFD2 Nucleic Acids and Proteins
; FILE REFERENCE: UM-08901
; CURRENT APPLICATION NUMBER: US/10/781,362
; CURRENT FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 23703
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-781-362-17

Alignment Scores:
Pred. No.: 1.35e+05 Length: 23703
Score: 36.00 Matches: 6
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 85.7% Mismatches: 0
Query Match: 94.7% Indels: 0
DB: 8 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-781-362-17 (1-23703)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        263 AGTCTCATAACTTGGAACCTT 243
```

RESULT 73
US-10-781-362-19/c
; Sequence 19, Application US/10781362
; Publication No. US20040259157A1
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Zhang, Bin
; APPLICANT: Kaufman, Randal J.
; TITLE OF INVENTION: MCFD2 Nucleic Acids and Proteins
; FILE REFERENCE: UM-08901
; CURRENT APPLICATION NUMBER: US/10/781,362
; CURRENT FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 23703
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-781-362-19

Alignment Scores:
Pred. No.: 1.35e+05 Length: 23703
Score: 36.00 Matches: 6
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 85.7% Mismatches: 0
Query Match: 94.7% Indels: 0
DB: 8 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-781-362-19 (1-23703)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        263 AGTCTCATAACTTGGAACCTT 243
```

RESULT 74
US-11-013-314-19/c
; Sequence 19, Application US/11013314
; Publication No. US20050166269A1
; GENERAL INFORMATION:
; APPLICANT: TONONI, GIULIO
; APPLICANT: CIRELLI, CHIARA
; TITLE OF INVENTION: SLEEP GENES IN DROSOPHILA AND THEIR USE FOR THE
; TITLE OF INVENTION: SCREENING, DIAGNOSIS AND THERAPY OF SLEEP DISORDERS
; FILE REFERENCE: WARF:010US
; CURRENT APPLICATION NUMBER: US/11/013,314
; CURRENT FILING DATE: 2004-12-15
; PRIOR APPLICATION NUMBER: 60/563,858
; PRIOR FILING DATE: 2004-04-20
; PRIOR APPLICATION NUMBER: 60/529,536
; PRIOR FILING DATE: 2003-12-15
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 30232
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-11-013-314-19

Alignment Scores:
Pred. No.: 1.75e+05 Length: 30232
Score: 36.00 Matches: 6
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 85.7% Mismatches: 0
Query Match: 94.7% Indels: 0
DB: 10 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-013-314-19 (1-30232)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||
Db       5182 AGCTTGCTAACTTGGAATATA 5162
```

RESULT 75
US-09-973-278-834/c
; Sequence 834, Application US/09973278
; Publication No. US20040044191A1
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: PZ010P2
; CURRENT APPLICATION NUMBER: US/09/973,278
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 60/239,899
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/227,357
; PRIOR FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: PCT/US98/13684
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/051,926
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,793
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,925
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,929
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,803

; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,732
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,931
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,932
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,916
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,930
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,918
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,920
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,733
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,795
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,928
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,722
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,723
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,948
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,949
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,953
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,950
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,947
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,964
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/056,360
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,684
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,954
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,785
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,664
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,660
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,661
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 947
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 834
; LENGTH: 39344
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-973-278-834

Alignment Scores:

Pred. No.:	2.31e+05	Length:	39344
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-973-278-834 (1-39344)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db      27408 TCATTACTAACTTGGAATATT 27388
```

Search completed: March 11, 2006, 05:26:51
Job time : 47.4361 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:40:54 ; Search time 230.791 Seconds
(without alignments)
1724.091 Million cell updates/sec

Title: US-10-757-745-2_COPY_115_121
Perfect score: 38
Sequence: 1 SLITWNI 7

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170327_4011/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_5954_@runat_09032006_170327_4011 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	38	100.0	411	10	HS133N16S	AL110544 H.sapiens
	2	38	100.0	483	6	BD076937	BD076937 5' EST of
	3	38	100.0	488	5	PGR288432	AJ288432 Poeciliop
	4	38	100.0	752	6	BD150065	BD150065 Primer fo
	5	38	100.0	752	6	AX870003	AX870003 Sequence
	6	38	100.0	858	6	CQ726600	CQ726600 Sequence
	7	38	100.0	1261	8	AF201687	AF201687 Homo sapi
	8	38	100.0	1898	6	BD157138	BD157138 Primer fo
	9	38	100.0	1898	6	AX878304	AX878304 Sequence
	10	38	100.0	1898	8	AK002168	AK002168 Homo sapi
	11	38	100.0	1920	6	BD205490	BD205490 CD40-Inte
	12	38	100.0	1920	6	AR594294	AR594294 Sequence
	13	38	100.0	1920	6	AX011599	AX011599 Sequence
	14	38	100.0	1921	8	HSA269473	AJ269473 Homo sapi
	15	38	100.0	1936	8	AF223469	AF223469 Homo sapi
	16	38	100.0	1948	6	AR339398	AR339398 Sequence
	17	38	100.0	1952	8	BC017553	BC017553 Homo sapi
	18	38	100.0	2151	8	HSA420495	AJ420495 Homo sapi
	19	38	100.0	2499	6	E23195	E23195 Topoisomera
	20	38	100.0	4804	1	AF476102	AF476102 Secondary
	21	38	100.0	20938	8	AY613922	AY613922 Homo sapi
	22	38	100.0	68246	14	AC101400	AC101400 Mus muscu
	23	38	100.0	71665	14	AC101157	AC101157 Mus muscu
c	24	38	100.0	71665	14	AC101157	AC101157 Mus muscu
c	25	38	100.0	80891	15	AC005693	AC005693 Arabidops
	26	38	100.0	88989	15	AB026658	AB026658 Arabidops
	27	38	100.0	102200	8	HS30M3	AL031775 Human DNA
c	28	38	100.0	114468	8	AC093026	AC093026 Homo sapi
c	29	38	100.0	130323	9	AC135637	AC135637 Mus muscu
c	30	38	100.0	131005	14	AC150890	AC150890 Medicago
	31	38	100.0	135269	14	HSAC000383	AC000383 Homo sapi
c	32	38	100.0	138271	8	AC124017	AC124017 Homo sapi
c	33	38	100.0	149416	14	AC021985	AC021985 Homo sapi
c	34	38	100.0	150344	14	AC152027	AC152027 Dasypus n
c	35	38	100.0	152937	8	AC007157	AC007157 Homo sapi
c	36	38	100.0	158145	8	AC104664	AC104664 Homo sapi
	37	38	100.0	160177	8	AL445242	AL445242 Human DNA
c	38	38	100.0	164475	5	CR361565	CR361565 Zebrafish
c	39	38	100.0	164528	14	AC016800	AC016800 Homo sapi
c	40	38	100.0	165586	9	AC121567	AC121567 Mus muscu
	41	38	100.0	167847	8	CR925830	CR925830 Human DNA
c	42	38	100.0	168291	9	AC151746	AC151746 Mus muscu
c	43	38	100.0	169905	8	AL442064	AL442064 Human DNA
c	44	38	100.0	170896	9	AC140199	AC140199 Mus muscu
	45	38	100.0	176819	8	CR942205	CR942205 Human DNA
c	46	38	100.0	178252	14	AC164269	AC164269 Bos tauru
	47	38	100.0	183533	8	BS000021	BS000021 Pan trogl
c	48	38	100.0	186051	9	AC112957	AC112957 Mus muscu
	49	38	100.0	191676	14	AC011844	AC011844 Homo sapi
	50	38	100.0	193842	14	AC148903	AC148903 Otolemur
	51	38	100.0	198857	8	AC012404	AC012404 Homo sapi
	52	38	100.0	201511	14	AC154641	AC154641 Mus muscu
c	53	38	100.0	202887	9	AL928789	AL928789 Mouse DNA
c	54	38	100.0	205582	14	AC120951	AC120951 Rattus no
	55	38	100.0	214189	8	BS000022	BS000022 Pan trogl
	56	38	100.0	214287	14	AC152083	AC152083 Bos tauru
	57	38	100.0	218833	14	AC115541	AC115541 Rattus no
c	58	38	100.0	229363	14	AC094940	AC094940 Rattus no
c	59	38	100.0	235823	9	AC113013	AC113013 Mus muscu
	60	38	100.0	236814	14	AC131436	AC131436 Rattus no
c	61	38	100.0	252828	14	AL590646	AL590646 Homo sapi
	62	38	100.0	255090	14	AC073798	AC073798 Mus muscu
	63	38	100.0	255706	14	AC094186	AC094186 Rattus no
c	64	38	100.0	348479	14	AL445583	AL445583 Homo sapi
c	65	37	97.4	444	10	BV289706	BV289706 S232P6107
	66	37	97.4	601	10	BV197165	BV197165 sqnm19184
	67	37	97.4	601	10	BV197166	BV197166 sqnm19184
c	68	37	97.4	3185	9	BC087162	BC087162 Rattus no
c	69	37	97.4	3194	9	BC049792	BC049792 Mus muscu
c	70	37	97.4	53047	14	AC135376	AC135376 Rattus no
	71	37	97.4	79949	8	AL162613	AL162613 Human DNA
	72	37	97.4	110000	1	BX908798_00	BX908798 Parachlam
c	73	37	97.4	110000	14	AC138524_3	Continuation (4 of
	74	37	97.4	110000	15	AP008208_093	Continuation (94 o
c	75	37	97.4	110000	15	AP008209_123	Continuation (124

ALIGNMENTS

RESULT 1
HS133N16S/c
LOCUS HS133N16S 411 bp DNA linear STS 02-SEP-1999
DEFINITION H.sapiens STS from genomic clone 133N16, sequence tagged site.
ACCESSION AL110544
VERSION AL110544.1 GI:5830260
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 411)
AUTHORS Mungall,A.J., Huckle,E., Taylor,R. and Hunt,S.E.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) E-mail contact: humquery@sanger.ac.uk
COMMENT Marker stdJ133N16SP6 (Primer A : TGTTCCCTCTTTGCTGCAG; Primer B :
AACAGTTCAAGGTTGTGGGC; amplimer size : 127 bp) is from sequence
generated from the SP6 end of PAC 133N16. 133N16 is part of the
bacterial clone contigs constructed by the Chromosome 6 Mapping
Group.(http://www.sanger.ac.uk/HGP/Chr6/) 133N16 is from the
library constructed at the Roswell Park Cancer Institute by the
group of Pieter de Jong.
For further details see http://bacpac.med.buffalo.edu/.
FEATURES Location/Qualifiers
source 1. .411
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="133N16"
ORIGIN

Alignment Scores:
Pred. No.: 138 Length: 411
Score: 38.00 Matches: 7
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x HS133N16S (1-411)

Qy 1 SerLeuIleThrTrpAsnIle 7
||||||||||||||||||||
Db 333 TCTCTCATTACCTGGAATATT 313

RESULT 2
BD076937
LOCUS BD076937 483 bp DNA linear PAT 27-AUG-2002
DEFINITION 5' EST of secretory protein expressed in prostate.
ACCESSION BD076937
VERSION BD076937.1 GI:22622540
KEYWORDS JP 2001512013-A/184.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 483)
AUTHORS Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
TITLE 5' EST of secretory protein expressed in prostate
JOURNAL Patent: JP 2001512013-A 184 21-AUG-2001;
GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001512013-A/184
PD 21-AUG-2001
PF 31-JUL-1998 JP 2000505291
PR 01-AUG-1997 US 08/905144
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO PI
LACROIX
PC C12N15/09,C07K14/47,C12P21/02,C12Q1/02,C12Q1/68,C12N15/00 CC
blastn
CC identity 99
CC region 1. .310
CC id T26956
CC est
CC blastn
CC identity 98
CC region 1. .315
CC id T31666
CC est
CC blastn
CC identity 100
CC region 137. .267
CC id R14990
CC est
CC blastn
CC identity 100
CC region 63. .137
CC id R14990
CC est
CC blastn
CC identity 100
CC region 1. .50
CC id R14990
CC est
CC Von Heijne matrix
CC score 5.2
CC seq LCVEFASVASCDA/AV
FH Key Location/Qualifiers
FT misc_feature 105. .414
FT misc_feature 45. .359
FT misc_feature 202. .332
FT misc_feature 127. .201
FT misc_feature 65. .114
FT sig_peptide 1. .120.
FEATURES Location/Qualifiers
source 1. .483
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.: 160 Length: 483
Score: 38.00 Matches: 7
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x BD076937 (1-483)

Qy 1 SerLeuIleThrTrpAsnIle 7
||||||||||||||||||||
Db 343 TCTCTCATTACCTGGAATATT 363

RESULT 3
PGR288432
LOCUS PGR288432 488 bp DNA linear VRT 23-SEP-2003
DEFINITION Poeciliopsis gracilis transposon Rex1 partial rt gene for reverse
transcriptase, clone Rex1-Pog2.
ACCESSION AJ288432
VERSION AJ288432.1 GI:7529657
KEYWORDS reverse transcriptase; RT gene; transposon.
SOURCE Poeciliopsis gracilis
ORGANISM Poeciliopsis gracilis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Cyprinodontiformes; Poeciliidae; Poeciliopsis.
REFERENCE 1

```
  AUTHORS   Volff,J.N., Korting,C. and Schartl,M.
  TITLE     Multiple lineages of the non-LTR retrotransposon Rex1 with varying
            success in invading fish genomes
  JOURNAL   Mol. Biol. Evol. 17 (11), 1673-1684 (2000)
   PUBMED   11070055
REFERENCE   2  (bases 1 to 488)
  AUTHORS   Volff,J.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2000) Volff J.N., Physiological Chemistry I,
            Biocenter, University of Wuerzburg, Am Hubland, Wuerzburg D-97074,
            GERMANY
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Poeciliopsis gracilis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:119849"
                     /clone="Rex1-Pog2"
     gene            1. .488
                     /gene="rt"
                     /pseudo
     CDS             <1. .>488
                     /gene="rt"
                     /pseudo
                     /codon_start=1
                     /product="reverse transcriptase"
     repeat_region   1. .488
                     /transposon="Rex1"
ORIGIN

Alignment Scores:
Pred. No.:                161           Length:         488
Score:                    38.00         Matches:        7
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       5             Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x PGR288432 (1-488)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db         52 TCACTCATCACCTGGAACATC 72

RESULT 4
BD150065
LOCUS       BD150065                752 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD150065
VERSION     BD150065.1  GI:27855823
KEYWORDS    JP 2002191363-A/4908.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 4908 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/4908
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
```

```
            Primer for synthesizing full-length cDNA and use thereof FH   Key
                  Location/Qualifiers
            FT   source          1. .752
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                238           Length:         752
Score:                    38.00         Matches:        7
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       6             Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BD150065 (1-752)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363

RESULT 5
AX870003
LOCUS       AX870003                752 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 4908 from Patent EP1074617.
ACCESSION   AX870003
VERSION     AX870003.1  GI:40024866
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 4908 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                238           Length:         752
Score:                    38.00         Matches:        7
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       6             Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AX870003 (1-752)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363

RESULT 6
CQ726600
LOCUS       CQ726600                858 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 12534 from Patent WO02068579.
ACCESSION   CQ726600
VERSION     CQ726600.1  GI:42290170
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Hominidae; Homo.
REFERENCE  1
  AUTHORS  Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE    Kits, such as nucleic acid arrays, comprising a majority of
           humanexons or transcripts, for detecting expression and other uses
           thereof
  JOURNAL  Patent: WO 02068579-A 12534 06-SEP-2002;
           PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .858
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                 268             Length:         858
Score:                     38.00           Matches:        7
Percent Similarity:        100.0%          Conservative:   0
Best Local Similarity:     100.0%          Mismatches:     0
Query Match:               100.0%          Indels:         0
DB:                        6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x CQ726600 (1-858)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363

RESULT 7
AF201687
LOCUS       AF201687                1261 bp    mRNA    linear   PRI 16-MAY-2003
DEFINITION  Homo sapiens ETS1-associated protein 2 (EAP2) mRNA, complete cds.
ACCESSION   AF201687
VERSION     AF201687.1  GI:11493668
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1261)
  AUTHORS   Pei,H., Yordy,J.S., Leng,Q., Zhao,Q., Watson,D.K. and Li,R.
  TITLE     EAPII interacts with ETS1 and modulates its transcriptional
            function
  JOURNAL   Oncogene 22 (18), 2699-2709 (2003)
   PUBMED   12743594
REFERENCE   2  (bases 1 to 1261)
  AUTHORS   Li,R., Pei,H. and Papas,T.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1999) Center for Molecular and Structural
            Biology, Department of Medicine, and Hollings Cancer Center,
            Medical University of South Carolina, 86 Jonathan Lucas St.,
            Charleston, SC 29425, USA
FEATURES             Location/Qualifiers
     source          1. .1261
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22.1-p22.3"
                     /cell_line="Jurkat"
     gene            1. .1261
                     /gene="EAP2"
     CDS             26. .1114
                     /gene="EAP2"
                     /codon_start=1
                     /product="ETS1-associated protein 2"
                     /protein_id="AAG35600.1"
                     /db_xref="GI:11493669"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                 379             Length:         1261
Score:                     38.00           Matches:        7
Percent Similarity:        100.0%          Conservative:   0
Best Local Similarity:     100.0%          Mismatches:     0
Query Match:               100.0%          Indels:         0
DB:                        8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AF201687 (1-1261)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        368 TCTCTCATTACCTGGAATATT 388

RESULT 8
BD157138
LOCUS       BD157138                1898 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD157138
VERSION     BD157138.1  GI:27862896
KEYWORDS    JP 2002191363-A/11981.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1898)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 11981 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/11981
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                      Location/Qualifiers
            FT   CDS             (1). .(1086).
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                 548             Length:         1898
Score:                     38.00           Matches:        7
Percent Similarity:        100.0%          Conservative:   0
Best Local Similarity:     100.0%          Mismatches:     0
Query Match:               100.0%          Indels:         0
DB:                        6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BD157138 (1-1898)

Qy          1 SerLeuIleThrTrpAsnIle 7
```

```
                    ||||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363

RESULT 9
AX878304
LOCUS       AX878304               1898 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 13209 from Patent EP1074617.
ACCESSION   AX878304
VERSION     AX878304.1  GI:40033040
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 13209 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .1089
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE90095.1"
                     /db_xref="GI:40033041"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEPASVASCDAAVAQ
                     CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                548           Length:          1898
Score:                    38.00         Matches:         7
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AX878304 (1-1898)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363

RESULT 10
AK002168
LOCUS       AK002168               1898 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ11306 fis, clone PLACE1010031.
ACCESSION   AK002168
VERSION     AK002168.1  GI:7023882
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
            Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
            Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
            Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
            Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
            Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
            Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
  TITLE     Complete sequencing and characterization of 21,243 full-length
            human cDNAs
  JOURNAL   Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED   14702039
REFERENCE   2
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
            Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
            Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
            Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
            Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1898)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1010031"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="cloning vector: pME18SFL3"
     CDS             1. .1089
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAA92119.1"
                     /db_xref="GI:7023883"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEPASVASCDAAVAQ
                     CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                548             Length:         1898
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AK002168 (1-1898)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363

RESULT 11
BD205490
LOCUS       BD205490                1920 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  CD40-Interacting and TRAF-interacting protein.
ACCESSION   BD205490
VERSION     BD205490.1  GI:33015260
KEYWORDS    JP 2002512796-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1920)
  AUTHORS   Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting protein
  JOURNAL   Patent: JP 2002512796-A 1 08-MAY-2002;
            VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002512796-A/1
            PD   08-MAY-2002
            PF   28-APR-1999 JP 2000546003
            PR   29-APR-1998 EP   98201392.2
            PI   STEFAN MARIA CHRISTIAAN PYPE,
            PI   JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI   DANNY
            FRANCOIS EVELINE HUYLEBROECK
            PC   C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
              A61P37/02,
            PC   A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
            PC   G01N33/566,
            PC   C12N15/00,A61K37/02
            CC   CD40-Interacting and TRAF-interacting protein. FH   Key
                 Location/Qualifiers
            FT   CDS             (20). .(1108).
FEATURES                Location/Qualifiers
     source             1. .1920
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                553             Length:         1920
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BD205490 (1-1920)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        362 TCTCTCATTACCTGGAATATT 382

RESULT 12
AR594294
LOCUS       AR594294                1920 bp    DNA     linear   PAT 15-DEC-2004
DEFINITION  Sequence 1 from patent US 6812203.
ACCESSION   AR594294
VERSION     AR594294.1  GI:56643900
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1920)
  AUTHORS   Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting proteins
  JOURNAL   Patent: US 6812203-A 1 02-NOV-2004;
            Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
            Zwijnaarde;
            WOX;
FEATURES                Location/Qualifiers
     source             1. .1920
                        /organism="unknown"
                        /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                553             Length:         1920
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AR594294 (1-1920)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        362 TCTCTCATTACCTGGAATATT 382

RESULT 13
AX011599
LOCUS       AX011599                1920 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 1 from Patent WO9955859.
ACCESSION   AX011599
VERSION     AX011599.1  GI:9998123
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 1 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES                Location/Qualifiers
     source             1. .1920
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"
     CDS                20. .1108
                        /note="unnamed protein product"
                        /codon_start=1
                        /protein_id="CAC07555.1"
                        /db_xref="GI:9998124"
                        /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                        CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                        SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                        LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                        ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                        PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                        DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                553             Length:         1920
Score:                    38.00           Matches:        7
```

```
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            100.0%          Indels:        0
DB:                     6               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x AX011599 (1-1920)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        362 TCTCTCATTACCTGGAATATT 382

RESULT 14
HSA269473
LOCUS       HSA269473               1921 bp    mRNA    linear   PRI 27-MAY-2000
DEFINITION  Homo sapiens mRNA for TRAF and TNF receptor associated protein
            (ttrap gene).
ACCESSION   AJ269473
VERSION     AJ269473.1  GI:8247253
KEYWORDS    TRAF and TNF receptor associated protein; ttrap gene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van
            Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,
            Huylebroeck,D. and Remacle,J.E.
  TITLE     TTRAP, a novel protein that associates with CD40, tumor necrosis
            factor (TNF) receptor-75 and TNF receptor-associated factors
            (TRAFs), and that inhibits nuclear factor-kappa B activation
  JOURNAL   J. Biol. Chem. 275 (24), 18586-18593 (2000)
   PUBMED   10764746
REFERENCE   2  (bases 1 to 1921)
  AUTHORS   Pype,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1999) Pype S., VIB07, Dept. Cell Growth,
            Differentiation and Development, Flanders Interuniversity Institute
            for Biotechnology, CELGEN, K.U.Leuven, Gasthuisberg Campus,
            Herestraat 49, B-3000 Leuven, BELGIUM
COMMENT     Related sequence: AL031775.
FEATURES             Location/Qualifiers
     source          1. .1921
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .1921
                     /gene="ttrap"
     5'UTR           <1. .21
                     /gene="ttrap"
     CDS             22. .1110
                     /gene="ttrap"
                     /function="putative role in signal transduction"
                     /codon_start=1
                     /product="TRAF and TNF receptor associated protein"
                     /protein_id="CAB92966.1"
                     /db_xref="GI:8247254"
                     /db_xref="GOA:O95551"
                     /db_xref="UniProt/TrEMBL:O95551"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
     3'UTR           1111. .1921
                     /gene="ttrap"
     polyA_signal    1900. .1905
                     /gene="ttrap"
ORIGIN

Alignment Scores:

Pred. No.:              554             Length:        1921
Score:                  38.00           Matches:       7
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            100.0%          Indels:        0
DB:                     8               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x HSA269473 (1-1921)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        364 TCTCTCATTACCTGGAATATT 384

RESULT 15
AF223469
LOCUS       AF223469                1936 bp    mRNA    linear   PRI 17-APR-2000
DEFINITION  Homo sapiens AD022 protein (AD022) mRNA, complete cds.
ACCESSION   AF223469
VERSION     AF223469.1  GI:7578788
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1936)
  AUTHORS   Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
  TITLE     A novel gene expressed in human adrenal gland
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1936)
  AUTHORS   Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) Chinese National Human Genome Center at
            Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
            Shanghai 201203, China
FEATURES             Location/Qualifiers
     source          1. .1936
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="adrenal gland"
     gene            1. .1936
                     /gene="AD022"
     CDS             17. .1105
                     /gene="AD022"
                     /codon_start=1
                     /product="AD022 protein"
                     /protein_id="AAF64144.1"
                     /db_xref="GI:7578789"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVRFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:              558             Length:        1936
Score:                  38.00           Matches:       7
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            100.0%          Indels:        0
DB:                     8               Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x AF223469 (1-1936)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        359 TCTCTCATTACCTGGAATATT 379

RESULT 16
```

```
AR339398
LOCUS       AR339398                1948 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 889 from patent US 6569662.
ACCESSION   AR339398
VERSION     AR339398.1  GI:33726255
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Tang,Y.T., Zhou,P. and Drmanac,R.T.
  TITLE     Nucleic acids and polypeptides
  JOURNAL   Patent: US 6569662-A 889 27-MAY-2003;
            Hyseq, Inc.; Sunnyvale, CA
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                561             Length:         1948
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AR339398 (1-1948)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        381 TCTCTCATTACCTGGAATATT 401

RESULT 17
BC017553
LOCUS       BC017553                1952 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein, mRNA (cDNA
            clone MGC:9099 IMAGE:3920790), complete cds.
ACCESSION   BC017553
VERSION     BC017553.2  GI:34782842
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1952)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1952)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Sep 16, 2003 this sequence version replaced gi:17028464.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 15 Row: m Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 23510347.
FEATURES             Location/Qualifiers
     source          1. .1952
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:9099 IMAGE:3920790"
                     /tissue_type="Skin, melanotic melanoma."
                     /clone_lib="NIH_MGC_72"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1952
                     /gene="TTRAP"
                     /note="synonyms: AD022, EAP2, MGC9099, dJ30M3.3"
                     /db_xref="GeneID:51567"
                     /db_xref="MIM:605764"
     CDS             38. .1126
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TRAF and TNF receptor-associated protein"
                     /protein_id="AAH17553.1"
                     /db_xref="GI:17028465"
                     /db_xref="GeneID:51567"
                     /db_xref="MIM:605764"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                562             Length:         1952
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BC017553 (1-1952)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        380 TCTCTCATTACCTGGAATATT 400
```

```
RESULT 18
HSA420495
LOCUS       HSA420495               2151 bp    mRNA    linear   PRI 23-NOV-2001
DEFINITION  Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1761756.
ACCESSION   AJ420495
VERSION     AJ420495.1  GI:17066359
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Auffray,C., Ansorge,W., Ballabio,A., Estivill,X., Gibson,K.,
            Lehrach,H., Poustka,A. and Lundeberg,J.
  TITLE     The European IMAGE consortium for integrated Molecular analysis of
            human gene transcripts
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2151)
  AUTHORS   Persson,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2001) Persson A., Center for Molecular
            Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagsvagen
            30B, 106 91 Stockholm, SWEDEN
COMMENT     This clone is available royalty-free through IMAGE Consortium
            Distributors. IMPORTANT: This sequence represents the full insert
            of this IMAGE cDNA clone. No attempt has been made to verify
            whether this corresponds to the full-length of the original mRNA
            from which it was derived.
FEATURES             Location/Qualifiers
     source          1. .2151
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="EUROIMAGE 1761756"
                     /clone_lib="NCI_CGAP_Kid3"
     polyA_signal    2105. .2110
ORIGIN

Alignment Scores:
Pred. No.:              613          Length:         2151
Score:                  38.00        Matches:        7
Percent Similarity:     100.0%       Conservative:   0
Best Local Similarity:  100.0%       Mismatches:     0
Query Match:            100.0%       Indels:         0
DB:                     8            Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x HSA420495 (1-2151)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        569 TCTCTCATTACCTGGAATATT 589

RESULT 19
E23195
LOCUS       E23195                  2499 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Topoisomerase binding protein.
ACCESSION   E23195
VERSION     E23195.1  GI:13024277
KEYWORDS    JP 1999075856-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2499)
  AUTHORS   Takashi,T. and Kazuhiko,Y.
  TITLE     Topoisomerase binding protein
  JOURNAL   Patent: JP 1999075856-A 1 23-MAR-1999;
            TAKASHI TSURUO,CHUGAI PHARMACEUT CO LTD
COMMENT     OS   Unidentified
            PN   JP 1999075856-A/1
            PD   23-MAR-1999
            PF   17-SEP-1997 JP 1997251544
            PR
            PI   TAKASHI TSURUO,KAZUHIKO YAMANE
            PC   C12N15/09,C07K14/47,C12N9/90,C12N15/00
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .2499
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2499
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:              702          Length:         2499
Score:                  38.00        Matches:        7
Percent Similarity:     100.0%       Conservative:   0
Best Local Similarity:  100.0%       Mismatches:     0
Query Match:            100.0%       Indels:         0
DB:                     6            Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x E23195 (1-2499)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        824 TCTCTCATTACCTGGAATATT 844

RESULT 20
AF476102
LOCUS       AF476102                4804 bp    DNA     linear   BCT 10-JUL-2002
DEFINITION  Secondary endosymbiont of Cyphonococcus alpinus 16S ribosomal RNA
            and 23S ribosomal RNA genes, partial sequence.
ACCESSION   AF476102
VERSION     AF476102.1  GI:21717584
KEYWORDS    .
SOURCE      secondary endosymbiont of Cyphonococcus alpinus
  ORGANISM  secondary endosymbiont of Cyphonococcus alpinus
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae.
REFERENCE   1  (bases 1 to 4804)
  AUTHORS   Thao,M.L., Gullan,P.J. and Baumann,P.
  TITLE     Secondary (gamma-Proteobacteria) Endosymbionts Infect the Primary
            (beta-Proteobacteria) Endosymbionts of Mealybugs Multiple Times and
            Coevolve with Their Hosts
  JOURNAL   Appl. Environ. Microbiol. 68 (7), 3190-3197 (2002)
   PUBMED   12088994
REFERENCE   2  (bases 1 to 4804)
  AUTHORS   Thao,M.L., Gullan,P.J. and Baumann,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-2002) Section of Microbiology, University of
            California Davis, One Shields Avenue, Davis, CA 95616, USA
FEATURES             Location/Qualifiers
     source          1. .4804
                     /organism="secondary endosymbiont of Cyphonococcus
                     alpinus"
                     /mol_type="genomic DNA"
                     /specific_host="Cyphonococcus alpinus"
                     /db_xref="taxon:199907"
     rRNA            <1. .1516
                     /product="16S ribosomal RNA"
     rRNA            2177. .>4804
                     /product="23S ribosomal RNA"
ORIGIN

Alignment Scores:
Pred. No.:              1.26e+03     Length:         4804
Score:                  38.00        Matches:        7
Percent Similarity:     100.0%       Conservative:   0
Best Local Similarity:  100.0%       Mismatches:     0
Query Match:            100.0%       Indels:         0
```

```
DB:                    1             Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x AF476102 (1-4804)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||||
Db       1945 TCTTTAATAACCTGGAATATA 1965

RESULT 21
AY613922
LOCUS       AY613922               20938 bp    DNA     linear   PRI 11-MAY-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein (TTRAP) gene,
            complete cds.
ACCESSION   AY613922
VERSION     AY613922.1  GI:47060296
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 20938)
  AUTHORS   Rieder,M.J., Daniels,R.L., da Ponte,S.H., Hastings,N.C.,
            Ahearn,M.O., Rajkumar,N., Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2004) Genome Sciences, University of Washington,
            1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  SeattleSNPs. NHLBI HL66682 Program
            for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.gs.washington.edu).
FEATURES             Location/Qualifiers
     source          1. .20938
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     repeat_region   1. .32
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     variation       81
                     /frequency="0.67"
                     /replace=""
     repeat_region   87. .368
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     repeat_region   369. .673
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       449
                     /frequency="0.96"
                     /replace=""
     variation       540
                     /frequency="0.01"
                     /replace="t"
     variation       586
                     /frequency="0.01"
                     /replace="a"
     variation       650
                     /frequency="0.03"
                     /replace="g"
     variation       658. .660
                     /frequency="0.01"
                     /replace=""
     repeat_region   683. .989
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       695
                     /frequency="0.18"
                     /replace="t"
     variation       922
                     /frequency="0.02"
                     /replace="t"
     variation       1319
                     /frequency="0.15"
                     /replace="t"
     variation       1456
                     /frequency="0.19"
                     /replace="a"
     variation       1532
                     /frequency="0.01"
                     /replace="c"
     variation       1711. .1713
                     /frequency="0.04"
                     /replace=""
     variation       1795
                     /frequency="0.10"
                     /replace="a"
     variation       1821
                     /frequency="0.02"
                     /replace="t"
     variation       1830. .1862
                     /frequency="0.04"
                     /replace=""
     variation       1913. .1914
                     /frequency="0.13"
                     /replace=""
     gene            <2005. .18938
                     /gene="TTRAP"
     mRNA            join(<2005. .2194,2281. .2366,10172. .10345,11003. .11094,
                     14376. .14494,15753. .15923,17846. .18938)
                     /gene="TTRAP"
                     /product="TRAF and TNF receptor associated protein"
     CDS             join(2030. .2194,2281. .2366,10172. .10345,11003. .11094,
                     14376. .14494,15753. .15923,17846. .18127)
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TRAF and TNF receptor associated protein"
                     /protein_id="AAT09764.1"
                     /db_xref="GI:47060297"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPPPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
     variation       2381
                     /gene="TTRAP"
                     /frequency="0.08"
                     /replace="a"
     variation       2767
                     /gene="TTRAP"
                     /frequency="0.08"
                     /replace="g"
     variation       2800
                     /gene="TTRAP"
                     /frequency="0.79"
                     /replace=""
     repeat_region   2919. .3323
                     /rpt_family="L2"
                     /rpt_type=dispersed
     variation       2927
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="g"
     variation       3004
                     /gene="TTRAP"
                     /frequency="0.04"
                     /replace="t"
     variation       3061
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="t"
     variation       3062
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="t"
```

```
     variation       3114
                     /gene="TTRAP"
                     /frequency="0.21"
                     /replace="c"
     variation       3418
                     /gene="TTRAP"
                     /frequency="0.03"
                     /replace="c"
     variation       3781
                     /gene="TTRAP"
                     /frequency="0.15"
                     /replace="g"
     variation       3988
                     /gene="TTRAP"
                     /frequency="0.01"
                     /replace="a"
     variation       4191
                     /gene="TTRAP"
                     /frequency="0.01"
                     /replace="c"
     variation       4393
                     /gene="TTRAP"
                     /frequency="0.97"
                     /replace=""
     variation       4414
                     /gene="TTRAP"
                     /frequency="0.03"
                     /replace="c"
     variation       4437
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="g"
     repeat_region   4446. .4530
                     /rpt_family="MER1_type"
                     /rpt_type=dispersed
     repeat_region   4559. .4609
                     /rpt_family="MER1_type"
                     /rpt_type=dispersed
     variation       4565
                     /gene="TTRAP"
                     /frequency="0.52"
                     /replace=""
     repeat_region   4625. .4760
                     /rpt_family="MER1_type"
                     /rpt_type=dispersed
     variation       4658
                     /gene="TTRAP"
                     /frequency="0.01"
                     /replace="a"
     variation       4861
                     /gene="TTRAP"
                     /frequency="0.02"
                     /replace="a"
     variation       4941
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="g"
     variation       4983
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="t"
     variation       5024
                     /gene="TTRAP"
                     /frequency="0.07"

Alignment Scores:
Pred. No.:              4.77e+03        Length:         20938
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0
```

```
US-10-757-745-2_COPY_115_121 (1-7) x AY613922 (1-20938)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      10263 TCTCTCATTACCTGGAATATT 10283

RESULT 22
AC101400
LOCUS       AC101400               68246 bp    DNA     linear   HTG 23-NOV-2001
DEFINITION  Mus musculus clone RP23-120C5, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC101400
VERSION     AC101400.1  GI:17060175
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 68246)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP23-120C5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 68246)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L16627
                Center clone name: 120_C_5
            --------------.
            * NOTE: This record contains 83 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *         1      721: contig of 721 bp in length
            *       722      821: gap of 100 bp
```

```
*      822     1544: contig of 723 bp in length
*     1545     1644: gap of 100 bp
*     1645     2365: contig of 721 bp in length
*     2366     2465: gap of 100 bp
*     2466     3177: contig of 712 bp in length
*     3178     3277: gap of 100 bp
*     3278     4011: contig of 734 bp in length
*     4012     4111: gap of 100 bp
*     4112     4857: contig of 746 bp in length
*     4858     4957: gap of 100 bp
*     4958     5699: contig of 742 bp in length
*     5700     5799: gap of 100 bp
*     5800     6529: contig of 730 bp in length
*     6530     6629: gap of 100 bp
*     6630     7362: contig of 733 bp in length
*     7363     7462: gap of 100 bp
*     7463     8086: contig of 624 bp in length
*     8087     8186: gap of 100 bp
*     8187     8910: contig of 724 bp in length
*     8911     9010: gap of 100 bp
*     9011     9734: contig of 724 bp in length
*     9735     9834: gap of 100 bp
*     9835    10545: contig of 711 bp in length
*    10546    10645: gap of 100 bp
*    10646    11370: contig of 725 bp in length
*    11371    11470: gap of 100 bp
*    11471    12217: contig of 747 bp in length
*    12218    12317: gap of 100 bp
*    12318    13059: contig of 742 bp in length
*    13060    13159: gap of 100 bp
*    13160    13901: contig of 742 bp in length
*    13902    14001: gap of 100 bp
*    14002    14741: contig of 740 bp in length
*    14742    14841: gap of 100 bp
*    14842    15537: contig of 696 bp in length
*    15538    15637: gap of 100 bp
*    15638    16371: contig of 734 bp in length
*    16372    16471: gap of 100 bp
*    16472    17195: contig of 724 bp in length
*    17196    17295: gap of 100 bp
*    17296    18017: contig of 722 bp in length
*    18018    18117: gap of 100 bp
*    18118    18844: contig of 727 bp in length
*    18845    18944: gap of 100 bp
*    18945    19675: contig of 731 bp in length
*    19676    19775: gap of 100 bp
*    19776    20497: contig of 722 bp in length
*    20498    20597: gap of 100 bp
*    20598    21311: contig of 714 bp in length
*    21312    21411: gap of 100 bp
*    21412    22149: contig of 738 bp in length
*    22150    22249: gap of 100 bp
*    22250    22968: contig of 719 bp in length
*    22969    23068: gap of 100 bp
*    23069    23805: contig of 737 bp in length
*    23806    23905: gap of 100 bp
*    23906    24637: contig of 732 bp in length
*    24638    24737: gap of 100 bp
*    24738    25465: contig of 728 bp in length
*    25466    25565: gap of 100 bp
*    25566    26292: contig of 727 bp in length
*    26293    26392: gap of 100 bp
*    26393    27053: contig of 661 bp in length
*    27054    27153: gap of 100 bp
*    27154    27856: contig of 703 bp in length
*    27857    27956: gap of 100 bp
*    27957    28665: contig of 709 bp in length
*    28666    28765: gap of 100 bp
*    28766    29504: contig of 739 bp in length
*    29505    29604: gap of 100 bp
*    29605    30331: contig of 727 bp in length
*    30332    30431: gap of 100 bp
*    30432    31173: contig of 742 bp in length
*    31174    31273: gap of 100 bp
*    31274    32011: contig of 738 bp in length
*    32012    32111: gap of 100 bp
*    32112    32820: contig of 709 bp in length
*    32821    32920: gap of 100 bp
*    32921    33651: contig of 731 bp in length
*    33652    33751: gap of 100 bp
*    33752    34480: contig of 729 bp in length
*    34481    34580: gap of 100 bp
*    34581    35299: contig of 719 bp in length
*    35300    35399: gap of 100 bp
*    35400    36133: contig of 734 bp in length
*    36134    36233: gap of 100 bp
*    36234    36965: contig of 732 bp in length
*    36966    37065: gap of 100 bp
*    37066    37789: contig of 724 bp in length
*    37790    37889: gap of 100 bp
*    37890    38625: contig of 736 bp in length
*    38626    38725: gap of 100 bp
*    38726    39463: contig of 738 bp in length
*    39464    39563: gap of 100 bp
*    39564    40307: contig of 744 bp in length
*    40308    40407: gap of 100 bp
*    40408    41152: contig of 745 bp in length
*    41153    41252: gap of 100 bp
*    41253    41992: contig of 740 bp in length
*    41993    42092: gap of 100 bp
*    42093    42806: contig of 714 bp in length
*    42807    42906: gap of 100 bp
*    42907    43639: contig of 733 bp in length
*    43640    43739: gap of 100 bp
*    43740    44466: contig of 727 bp in length
*    44467    44566: gap of 100 bp
*    44567    45177: contig of 611 bp in length
*    45178    45277: gap of 100 bp
*    45278    45994: contig of 717 bp in length
*    45995    46094: gap of 100 bp
*    46095    46809: contig of 715 bp in length
*    46810    46909: gap of 100 bp
*    46910    47623: contig of 714 bp in length
*    47624    47723: gap of 100 bp
*    47724    48422: contig of 699 bp in length
*    48423    48522: gap of 100 bp
*    48523    49255: contig of 733 bp in length
*    49256    49355: gap of 100 bp
*    49356    50090: contig of 735 bp in length
*    50091    50190: gap of 100 bp
*    50191    50900: contig of 710 bp in length
*    50901    51000: gap of 100 bp
*    51001    51723: contig of 723 bp in length
*    51724    51823: gap of 100 bp
*    51824    52532: contig of 709 bp in length
*    52533    52632: gap of 100 bp
*    52633    53357: contig of 725 bp in length
*    53358    53457: gap of 100 bp
*    53458    54184: contig of 727 bp in length
*    54185    54284: gap of 100 bp
*    54285    55026: contig of 742 bp in length
*    55027    55126: gap of 100 bp
*    55127    55874: contig of 748 bp in length
*    55875    55974: gap of 100 bp
*    55975    56707: contig of 733 bp in length

Alignment Scores:
Pred. No.:                1.38e+04        Length:          68246
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       14              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC101400 (1-68246)
```

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||||
Db      50684 TCATTAATAACCTGGAATATC 50704

RESULT 23
AC101157
LOCUS       AC101157               71665 bp    DNA     linear   HTG 23-NOV-2001
DEFINITION  Mus musculus clone RP23-172J12, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC101157
VERSION     AC101157.1  GI:17059931
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 71665)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP23-172J12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 71665)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L16045
                Center clone name: 172_J_12
            --------------.
            * NOTE: This record contains 87 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      709: contig of 709 bp in length
            *      710      809: gap of 100 bp
            *      810     1532: contig of 723 bp in length
            *     1533     1632: gap of 100 bp
            *     1633     2369: contig of 737 bp in length
            *     2370     2469: gap of 100 bp
            *     2470     3180: contig of 711 bp in length
            *     3181     3280: gap of 100 bp
            *     3281     4009: contig of 729 bp in length
            *     4010     4109: gap of 100 bp
            *     4110     4846: contig of 737 bp in length
            *     4847     4946: gap of 100 bp
            *     4947     5674: contig of 728 bp in length
            *     5675     5774: gap of 100 bp
            *     5775     6501: contig of 727 bp in length
            *     6502     6601: gap of 100 bp
            *     6602     7318: contig of 717 bp in length
            *     7319     7418: gap of 100 bp
            *     7419     8108: contig of 690 bp in length
            *     8109     8208: gap of 100 bp
            *     8209     8915: contig of 707 bp in length
            *     8916     9015: gap of 100 bp
            *     9016     9743: contig of 728 bp in length
            *     9744     9843: gap of 100 bp
            *     9844    10568: contig of 725 bp in length
            *    10569    10668: gap of 100 bp
            *    10669    11383: contig of 715 bp in length
            *    11384    11483: gap of 100 bp
            *    11484    12241: contig of 758 bp in length
            *    12242    12341: gap of 100 bp
            *    12342    13072: contig of 731 bp in length
            *    13073    13172: gap of 100 bp
            *    13173    13893: contig of 721 bp in length
            *    13894    13993: gap of 100 bp
            *    13994    14707: contig of 714 bp in length
            *    14708    14807: gap of 100 bp
            *    14808    15532: contig of 725 bp in length
            *    15533    15632: gap of 100 bp
            *    15633    16363: contig of 731 bp in length
            *    16364    16463: gap of 100 bp
            *    16464    17183: contig of 720 bp in length
            *    17184    17283: gap of 100 bp
            *    17284    18015: contig of 732 bp in length
            *    18016    18115: gap of 100 bp
            *    18116    18816: contig of 701 bp in length
            *    18817    18916: gap of 100 bp
            *    18917    19648: contig of 732 bp in length
            *    19649    19748: gap of 100 bp
            *    19749    20461: contig of 713 bp in length
            *    20462    20561: gap of 100 bp
            *    20562    21299: contig of 738 bp in length
            *    21300    21399: gap of 100 bp
            *    21400    22137: contig of 738 bp in length
            *    22138    22237: gap of 100 bp
            *    22238    22971: contig of 734 bp in length
            *    22972    23071: gap of 100 bp
            *    23072    23789: contig of 718 bp in length
            *    23790    23889: gap of 100 bp
            *    23890    24613: contig of 724 bp in length
            *    24614    24713: gap of 100 bp
            *    24714    25415: contig of 702 bp in length
            *    25416    25515: gap of 100 bp
            *    25516    26248: contig of 733 bp in length
            *    26249    26348: gap of 100 bp
            *    26349    27084: contig of 736 bp in length
            *    27085    27184: gap of 100 bp
            *    27185    27911: contig of 727 bp in length
            *    27912    28011: gap of 100 bp
            *    28012    28715: contig of 704 bp in length
            *    28716    28815: gap of 100 bp
            *    28816    29500: contig of 685 bp in length
            *    29501    29600: gap of 100 bp
            *    29601    30339: contig of 739 bp in length
            *    30340    30439: gap of 100 bp
            *    30440    31163: contig of 724 bp in length
            *    31164    31263: gap of 100 bp
            *    31264    31998: contig of 735 bp in length
```

```
*    31999    32098: gap of 100 bp
*    32099    32814: contig of 716 bp in length
*    32815    32914: gap of 100 bp
*    32915    33635: contig of 721 bp in length
*    33636    33735: gap of 100 bp
*    33736    34457: contig of 722 bp in length
*    34458    34557: gap of 100 bp
*    34558    35274: contig of 717 bp in length
*    35275    35374: gap of 100 bp
*    35375    36103: contig of 729 bp in length
*    36104    36203: gap of 100 bp
*    36204    36934: contig of 731 bp in length
*    36935    37034: gap of 100 bp
*    37035    37761: contig of 727 bp in length
*    37762    37861: gap of 100 bp
*    37862    38569: contig of 708 bp in length
*    38570    38669: gap of 100 bp
*    38670    39404: contig of 735 bp in length
*    39405    39504: gap of 100 bp
*    39505    40242: contig of 738 bp in length
*    40243    40342: gap of 100 bp
*    40343    41072: contig of 730 bp in length
*    41073    41172: gap of 100 bp
*    41173    41905: contig of 733 bp in length
*    41906    42005: gap of 100 bp
*    42006    42733: contig of 728 bp in length
*    42734    42833: gap of 100 bp
*    42834    43554: contig of 721 bp in length
*    43555    43654: gap of 100 bp
*    43655    44381: contig of 727 bp in length
*    44382    44481: gap of 100 bp
*    44482    45205: contig of 724 bp in length
*    45206    45305: gap of 100 bp
*    45306    46044: contig of 739 bp in length
*    46045    46144: gap of 100 bp
*    46145    46865: contig of 721 bp in length
*    46866    46965: gap of 100 bp
*    46966    47688: contig of 723 bp in length
*    47689    47788: gap of 100 bp
*    47789    48520: contig of 732 bp in length
*    48521    48620: gap of 100 bp
*    48621    49363: contig of 743 bp in length
*    49364    49463: gap of 100 bp
*    49464    50198: contig of 735 bp in length
*    50199    50298: gap of 100 bp
*    50299    51017: contig of 719 bp in length
*    51018    51117: gap of 100 bp
*    51118    51840: contig of 723 bp in length
*    51841    51940: gap of 100 bp
*    51941    52658: contig of 718 bp in length
*    52659    52758: gap of 100 bp
*    52759    53485: contig of 727 bp in length
*    53486    53585: gap of 100 bp
*    53586    54310: contig of 725 bp in length
*    54311    54410: gap of 100 bp
*    54411    55128: contig of 718 bp in length
*    55129    55228: gap of 100 bp
*    55229    55957: contig of 729 bp in length
*    55958    56057: gap of 100 bp
*    56058    56763: contig of 706 bp in length

Alignment Scores:
Pred. No.:                1.45e+04        Length:          71665
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       14              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC101157 (1-71665)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
```

```
Db      55718 AGTTTGATTACCTGGAACATC 55738

RESULT 24
AC101157/c
LOCUS       AC101157               71665 bp    DNA     linear   HTG 23-NOV-2001
DEFINITION  Mus musculus clone RP23-172J12, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC101157
VERSION     AC101157.1  GI:17059931
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 71665)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP23-172J12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 71665)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L16045
                Center clone name: 172_J_12
            --------------.
            * NOTE: This record contains 87 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      709: contig of 709 bp in length
            *      710      809: gap of 100 bp
            *      810     1532: contig of 723 bp in length
            *     1533     1632: gap of 100 bp
            *     1633     2369: contig of 737 bp in length
            *     2370     2469: gap of 100 bp
```

```
*     2470     3180: contig of 711 bp in length
*     3181     3280: gap of 100 bp
*     3281     4009: contig of 729 bp in length
*     4010     4109: gap of 100 bp
*     4110     4846: contig of 737 bp in length
*     4847     4946: gap of 100 bp
*     4947     5674: contig of 728 bp in length
*     5675     5774: gap of 100 bp
*     5775     6501: contig of 727 bp in length
*     6502     6601: gap of 100 bp
*     6602     7318: contig of 717 bp in length
*     7319     7418: gap of 100 bp
*     7419     8108: contig of 690 bp in length
*     8109     8208: gap of 100 bp
*     8209     8915: contig of 707 bp in length
*     8916     9015: gap of 100 bp
*     9016     9743: contig of 728 bp in length
*     9744     9843: gap of 100 bp
*     9844    10568: contig of 725 bp in length
*    10569    10668: gap of 100 bp
*    10669    11383: contig of 715 bp in length
*    11384    11483: gap of 100 bp
*    11484    12241: contig of 758 bp in length
*    12242    12341: gap of 100 bp
*    12342    13072: contig of 731 bp in length
*    13073    13172: gap of 100 bp
*    13173    13893: contig of 721 bp in length
*    13894    13993: gap of 100 bp
*    13994    14707: contig of 714 bp in length
*    14708    14807: gap of 100 bp
*    14808    15532: contig of 725 bp in length
*    15533    15632: gap of 100 bp
*    15633    16363: contig of 731 bp in length
*    16364    16463: gap of 100 bp
*    16464    17183: contig of 720 bp in length
*    17184    17283: gap of 100 bp
*    17284    18015: contig of 732 bp in length
*    18016    18115: gap of 100 bp
*    18116    18816: contig of 701 bp in length
*    18817    18916: gap of 100 bp
*    18917    19648: contig of 732 bp in length
*    19649    19748: gap of 100 bp
*    19749    20461: contig of 713 bp in length
*    20462    20561: gap of 100 bp
*    20562    21299: contig of 738 bp in length
*    21300    21399: gap of 100 bp
*    21400    22137: contig of 738 bp in length
*    22138    22237: gap of 100 bp
*    22238    22971: contig of 734 bp in length
*    22972    23071: gap of 100 bp
*    23072    23789: contig of 718 bp in length
*    23790    23889: gap of 100 bp
*    23890    24613: contig of 724 bp in length
*    24614    24713: gap of 100 bp
*    24714    25415: contig of 702 bp in length
*    25416    25515: gap of 100 bp
*    25516    26248: contig of 733 bp in length
*    26249    26348: gap of 100 bp
*    26349    27084: contig of 736 bp in length
*    27085    27184: gap of 100 bp
*    27185    27911: contig of 727 bp in length
*    27912    28011: gap of 100 bp
*    28012    28715: contig of 704 bp in length
*    28716    28815: gap of 100 bp
*    28816    29500: contig of 685 bp in length
*    29501    29600: gap of 100 bp
*    29601    30339: contig of 739 bp in length
*    30340    30439: gap of 100 bp
*    30440    31163: contig of 724 bp in length
*    31164    31263: gap of 100 bp
*    31264    31998: contig of 735 bp in length
*    31999    32098: gap of 100 bp
*    32099    32814: contig of 716 bp in length
*    32815    32914: gap of 100 bp
*    32915    33635: contig of 721 bp in length
*    33636    33735: gap of 100 bp
*    33736    34457: contig of 722 bp in length
*    34458    34557: gap of 100 bp
*    34558    35274: contig of 717 bp in length
*    35275    35374: gap of 100 bp
*    35375    36103: contig of 729 bp in length
*    36104    36203: gap of 100 bp
*    36204    36934: contig of 731 bp in length
*    36935    37034: gap of 100 bp
*    37035    37761: contig of 727 bp in length
*    37762    37861: gap of 100 bp
*    37862    38569: contig of 708 bp in length
*    38570    38669: gap of 100 bp
*    38670    39404: contig of 735 bp in length
*    39405    39504: gap of 100 bp
*    39505    40242: contig of 738 bp in length
*    40243    40342: gap of 100 bp
*    40343    41072: contig of 730 bp in length
*    41073    41172: gap of 100 bp
*    41173    41905: contig of 733 bp in length
*    41906    42005: gap of 100 bp
*    42006    42733: contig of 728 bp in length
*    42734    42833: gap of 100 bp
*    42834    43554: contig of 721 bp in length
*    43555    43654: gap of 100 bp
*    43655    44381: contig of 727 bp in length
*    44382    44481: gap of 100 bp
*    44482    45205: contig of 724 bp in length
*    45206    45305: gap of 100 bp
*    45306    46044: contig of 739 bp in length
*    46045    46144: gap of 100 bp
*    46145    46865: contig of 721 bp in length
*    46866    46965: gap of 100 bp
*    46966    47688: contig of 723 bp in length
*    47689    47788: gap of 100 bp
*    47789    48520: contig of 732 bp in length
*    48521    48620: gap of 100 bp
*    48621    49363: contig of 743 bp in length
*    49364    49463: gap of 100 bp
*    49464    50198: contig of 735 bp in length
*    50199    50298: gap of 100 bp
*    50299    51017: contig of 719 bp in length
*    51018    51117: gap of 100 bp
*    51118    51840: contig of 723 bp in length
*    51841    51940: gap of 100 bp
*    51941    52658: contig of 718 bp in length
*    52659    52758: gap of 100 bp
*    52759    53485: contig of 727 bp in length
*    53486    53585: gap of 100 bp
*    53586    54310: contig of 725 bp in length
*    54311    54410: gap of 100 bp
*    54411    55128: contig of 718 bp in length
*    55129    55228: gap of 100 bp
*    55229    55957: contig of 729 bp in length
*    55958    56057: gap of 100 bp
*    56058    56763: contig of 706 bp in length

Alignment Scores:
Pred. No.:                1.45e+04        Length:          71665
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       14              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC101157 (1-71665)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      70506 AGTTTGATTACCTGGAACATC 70486
```

```
RESULT 25
AC005693/c
LOCUS       AC005693               80891 bp    DNA     linear   PLN 27-FEB-2002
DEFINITION  Arabidopsis thaliana chromosome 2 clone T25N22 map mi310, complete
            sequence.
ACCESSION   AC005693
VERSION     AC005693.3  GI:20197425
KEYWORDS    HTG.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 80891)
  AUTHORS   Rounsley,S.D., Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Mason,T.M.,
            Shen,M., Ronning,C.M., Fraser,C.M., Somerville,C.R. and Venter,J.C.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 80891)
  AUTHORS   Lin,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
REFERENCE   3  (bases 1 to 80891)
  AUTHORS   Town,C.D. and Kaul,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
COMMENT     On Apr 18, 2002 this sequence version replaced gi:6598480.
FEATURES             Location/Qualifiers
     source          1. .80891
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /map="mi310"
                     /clone="T25N22"
                     /ecotype="Columbia"
     gene            complement(5320. .7406)
                     /gene="At2g07070"
                     /note="synonym: T25N22.2"
     mRNA            complement(join(<5320. .5773,5785. .6269,6562. .6689,
                     6766. .7094,7297. .>7406))
                     /gene="At2g07070"
     CDS             complement(join(5320. .5773,5785. .6269,6562. .6689,
                     6766. .7094,7297. .7406))
                     /gene="At2g07070"
                     /codon_start=1
                     /product="putative zinc finger protein"
                     /protein_id="AAM15071.1"
                     /db_xref="GI:20197426"
                     /translation="MGRLTNPAAQRLWSLFPFLSNRWNLKGKVVGSGLGKGEELGEIL
                     DMEISSSSVKLKVLINGLQPLTKEIVVEFPDGSEALVLLDYKNLKNHCRHCLRLSHDI
                     KDCPGLISSSEKSKPPTSTHSRGSGGVSRNYYTPADNFVASKSHELTSPLEKGLEVSF
                     RKEVFHYYPQGIEHGLALKLKFITLTTLMSAARKQRVIQGEARNLMSNTTDQIIEAAA
                     ASISSADPIGLALPASPEPRVSVREDPSLPTSSHVVQVKRGRGRPPLAKTNSKNPANS
                     SSKLIGAKSQKRNLVRNSPKKGTGSNLPSSFAALQGGNPSSKAACPKSKKPANPKIFL
                     IPAAQKNKVDFHNPLAPIVSWNCCGLGNPMTVQRLKKIRKSIFPDILFLMETKNTSEN
                     VLQTLQWLDNNSNHLVAPLSPAAGGLALFWRSSLQVEIISSYIHFMDTKITAKGKSFY
                     ATFLYGEPERSERLQVWSLLQEVYSTRSGPWLITGDLPAIFFSLVLSTYIL"
     gene            9331. .11306
                     /gene="At2g07080"
                     /note="synonym: T25N22.4"
     mRNA            join(<9331. .9436,9493. .10104,10141. .>11306)
                     /gene="At2g07080"
     CDS             join(9331. .9436,9493. .10104,10141. .11306)
                     /gene="At2g07080"
                     /codon_start=1
                     /product="putative gag-protease polyprotein"
                     /protein_id="AAC69114.1"
                     /db_xref="GI:3805758"
                     /translation="MEPTDVSTGTGKVLLLDTKRYGYWKVCMTQIIRGQEDGFKITKP
                     KANWTAEEKLQSKFNARAMKAIFNGVDEDEFKLIQGCKSAKQAWDTLQKSHEGTSSVK
                     RTRLDHIATQFEYLKMEPDEKIVKFSSKISALANEAEVMGKTYKDQKLVKKLLRCLPP
                     KFAAHKAVMRVAGNTDKISFVDLVGMLKLEEMKADQDKVKPSKNIAFNADQGSEQFQE
                     IKDGMALLARNFGKALKRVEIDGERSRGRFSRSENDDLRKKKEIQCYECGGFGHIKPE
                     CPITKRKEMKCLKCKGVGHTKFECPNKSKLKEKSLISFSDSESDDEGEELLNFVAFMA
                     SSDSSKFMSDTDSDCDEELNPKDKYRVLYDSWVQLSKDKLKLVKEKLTLEAKLANVST
                     EDKQKLSGITVDGNSQDYYQKKLDCLQKECHRERDRTKLLERELNDKYKQIRMLNKGL
                     ESLDKILAMGRTDSQQRGLGYQGYTGKIKKEEGRVINFVSGGSTSETVVRQSYIEPKK
                     QVKSHVEIKRESVVRTRMVGVICCDHCGKRFHMREQCYKFKEKVRTLWNLAKCFIEPS
                     RFCTVWIKKKDLYGDMEAERYNHKLERMYEEEEMHYASCCNQVNKEMQDEARKKDDSF
                     GST"
     repeat_region   complement(11960. .11981)
                     /rpt_family="AT_rich"
     repeat_region   complement(12523. .12554)
                     /rpt_family="AT_rich"
     gene            14525. .15076
                     /gene="At2g07090"
                     /note="synonym: T25N22.5; predicted by genscan"
     mRNA            join(<14525. .14716,14789. .>15076)
                     /gene="At2g07090"
     CDS             join(14525. .14716,14789. .15076)
                     /gene="At2g07090"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAC69115.1"
                     /db_xref="GI:3805759"
                     /translation="MSSSSVTSRVARERGVAGERGVPSKCLCGSEVTIFVSKTQENQG
                     RPFFRCVRNRDVNTWTTKRDGHLFKWVEDPVFEEVEDALPRLAIVANEISKVKAETNE
                     MLNQIQELKEEAIARKKELVKCKEWLYVCFLLLCVITIVVAYVMFGQGNQKKLSIGY"
     gene            complement(15325. .18172)
                     /gene="At2g07100"
                     /note="synonym: T25N22.6; pseudogene, Mutator-like
                     transposase"
                     /pseudo
     repeat_region   complement(18114. .18170)
                     /rpt_family="(GAA)n"
     repeat_region   19556. .19650
                     /rpt_family="AT_rich"
     repeat_region   20562. .20585
                     /rpt_family="AT_rich"
     repeat_region   complement(20751. .20775)
                     /rpt_family="AT_rich"
     repeat_region   complement(20993. .21013)
                     /rpt_family="AT_rich"
     repeat_region   complement(21986. .22007)
                     /rpt_family="AT_rich"
     repeat_region   complement(22418. .22507)
                     /rpt_family="AT_rich"
     repeat_region   23054. .23094
                     /rpt_family="AT_rich"
     repeat_region   complement(23268. .23332)
                     /rpt_family="AT_rich"
     repeat_region   complement(23689. .23724)
                     /rpt_family="AT_rich"
     repeat_region   complement(24520. .24563)
                     /rpt_family="(TA)n"
     repeat_region   24960. .25050
                     /rpt_family="AT_rich"
     repeat_region   complement(25202. .25293)
                     /rpt_family="AT_rich"
     gene            complement(25702. .26300)
                     /gene="At2g07110"
                     /note="synonym: T25N22.7"
     mRNA            complement(join(<25702. .25821,25905. .26005,26069. .>26300))
                     /gene="At2g07110"
     CDS             complement(join(25702. .25821,25905. .26005,26069. .26300))
                     /gene="At2g07110"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="AAC69116.1"
                     /db_xref="GI:3805760"
                     /translation="MGTKLFLSFLNGIHGTLSFIFSEDTKVLSAQYGPKARTKKNENA
                     EKACLKSYGSLKPDRLVIVNAFSFSERGLFCRLRKVEKEYETILKRTLQSICVLRVVP
                     NTTTSVIIQVVHDGGSVSFLCSFHLGKHLLMLSETHIIRYKQCLPEYL"
     gene            complement(26697. .27814)
```

```
                     /gene="At2g07120"
                     /note="synonym: T25N22.8"
     mRNA            complement(join(<26697. .27683,27740. .>27814))
                     /gene="At2g07120"
     CDS             complement(join(26697. .27683,27740. .27814))
                     /gene="At2g07120"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="AAC69117.1"
                     /db_xref="GI:3805761"
                     /translation="MASGKLPWELEEEILCRLPPGSLVRKSLGCARPQFIILPESKIY
                     SIGTIGLDDFDPKIEVRELACQFHFKAEKWAFTRITACDGLLFRDFWNQGVTIWNPWL
                     RQVGWIEYKEDRDFRFCGVGYDACKPEKDYKIFGYFSRFYDTTLKIGHRFAIFECASQ
                     AFKFIDSPEWPMLAGRGEYVSLNGNLYWTAYNEDTREHFLGSFDFSTEISMHFCLLPC
                     GKHVSGLRDKLVLTVFKGDQFALLKQSRISSNTEIWVTNDMINSSNNVVWLNLMTLSI
                     PDFPSLFHQLSDISYFIHDMTLILCCDDNQTGVGCIFIARGDFCKKIQIDSVSLGFSQ
                     CVHLPSLTSVPLEFRSLQV"
     repeat_region   28079. .28121
                     /rpt_family="AT_rich"
     repeat_region   complement(28640. .28707)
                     /rpt_family="AT_rich"
     gene            29108. .29972
                     /gene="At2g07130"
                     /note="synonym: T25N22.9; predicted by genscan and
                     genefinder"
     mRNA            join(<29108. .29366,29602. .>29972)
                     /gene="At2g07130"
     CDS             join(29108. .29366,29602. .29972)
                     /gene="At2g07130"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAC69118.1"
                     /db_xref="GI:3805762"
                     /translation="MDLADPSAEVKGELTLPDPYFKNSADEFDIDRVFHCDGLLLCTS
                     QLERRMVVWNPLTGETKWIQTHEEDKNISCKKSYKIMGFYRSVRMVLIVFPLLEERRD
                     KRSKAEIWVTNKIDDTTEGAVSWTKVLVLDLSRDLQALFTSNFLVDEEKKVFICCVSW
                     KEDEDENKSKKVYIVGEDNRVKEIDSGEDATSGCEPTILSLYVPSLVYM"
     gene            30653. .31807
                     /gene="At2g07140"
                     /note="synonym: T25N22.10; predicted by genscan and
                     genefinder"
     mRNA            <30653. .>31807
                     /gene="At2g07140"
     CDS             30653. .31807
                     /gene="At2g07140"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAC69119.1"
                     /db_xref="GI:3805763"
                     /translation="MTLPELPKDLVEEILSFVPATSLKRLRSTCKGWNRLFKDDKRFT
                     RIHTEKAAKQFQPLTLTKNYRICPINVNLHGTTPSLEVKNEVSLLDPHSKNSAAQFNI
                     DRVFHCDGLLLCTSQKDSRFVVWNPLTGVTKWIELGDRYNEGMAFILGYDNKSCNKSY
                     KAMSFNYLDKDSEIYEFSSDSWRVIDDIIKPPHYMDYFRECFSLKGNTYWLGIDRRRR
                     PPDLRITLIKFDFGTEKFGYVSLPPPCQVHGFEASNLSVVGDEKLSVLVQAGSTSKTE
                     VWVTSKIGEANVVSWSKVLSLYPKPDVGFWHGLSFLLDEEKKVFLCCKSKGWMEEEDE
                     DNVYIVGEDNKFILLNFGVETIGGESPIITTYVPSLVQIELAGSKRKTDY"

Alignment Scores:
Pred. No.:              1.61e+04        Length:         80891
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     15              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC005693 (1-80891)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      33048 AGTTTGATTACTTGGAACATA 33028

RESULT 26
AB026658
```

```
LOCUS       AB026658               88989 bp    DNA     linear   PLN 14-FEB-2004
DEFINITION  Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MYF24.
ACCESSION   AB026658 BA000014
VERSION     AB026658.1  GI:4757414
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Sato,S., Nakamura,Y., Kaneko,T., Katoh,T., Asamizu,E. and Tabata,S.
  TITLE     Structural analysis of Arabidopsis thaliana chromosome 3. I.
            Sequence features of the regions of 4,504,864 bp covered by sixty
            P1 and TAC clones
  JOURNAL   DNA Res. 7 (2), 131-135 (2000)
   PUBMED   10819329
REFERENCE   2  (bases 1 to 88989)
  AUTHORS   Sato,S., Nakamura,Y., Kaneko,T., Kato,T., Asamizu,E. and Tabata,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-1999) Yasukazu Nakamura, Kazusa DNA Research
            Institute, Department of Plant Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
            Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT     Address for correspondence: kaos@kazusa.or.jp
            For the latest information on annotation of this clone, please see
            http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MYF24
            Genes with similarity to proteins in the databases are described in
            'product' or 'note' qualifiers.  Genes that have no significant
            protein similarity are described as 'unknown protein'.
            The software programs used to predict genes include: Grail
            (Informatics Group, Oak Ridge National Laboratory,
            http://compbio.ornl.gov/Grail-1.3/),
            GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
            NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
            Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
            SplicePredictor (Volker Brendel, Stanford University,
            http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).
            Genes encoding tRNAs are predicted by tRNAscan-SE
            (Sean Eddy, Washington University School of Medicine, St. Louis,
            http://genome.wustl.edu/eddy/tRNAscan-SE/).
            This sequence may not be the entire insert of this clone. It may be
            shorter because we remove overlaps between neighboring submissions.
            The 5' clone is MIE15 and the 3' clone is K24M9.
FEATURES             Location/Qualifiers
     source          1. .88989
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /chromosome="3"
                     /clone="MYF24"
                     /clone_lib="Mitsui P1"
                     /ecotype="Columbia"
     CDS             1511. .2335
                     /note="unnamed protein product; gb|AAF26481.1
                     gene_id:MYF24.2
                     similar to unknown protein"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="BAB01096.1"
                     /db_xref="GI:11994093"
                     /translation="MICTESNHQRFSFAGDLGQSDKGTPMEQQPSGPVRRDTTLLDSS
                     NSDFEFHISSNFDPGDSSPADEIFADGMILPVLPFQVTATSTMPKRLYKYELPPIVSA
                     PTLSSYLPPLPLPLPEHSRKYSVKETRGSLNGRGSGANSDSEAEKSSKSFWSFKRSSS
                     LNCDIKKSLICSFPRLTRSNSTGSVAISKREMLRDINKHSSQRHGVPRPGVNPSSHMR
                     PPSSFCCSSYQFRPQKHAGKNGGGRGGSFWIAPVIGGPSPFGLGSILRLTKEKKKK"
     CDS             complement(2638. .5259)
                     /note="unnamed protein product; gene_id:MYF24.3
                     unknown protein"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="BAB01097.1"
                     /db_xref="GI:11994094"
```

/translation="MITRRRSPSKNPMEMITDEWGRKLCITSSPEESIGPFFSNPSDS
QLLFSSPSLSPPILSITPHLTPARFLSVSGVPPSDSSAINSSFKISNPHDDTVRVLSY
NRLQFLPFPSKNSVLVFFPTGTNLDQIGFLLLSYGDSGGLQVTGSDEGDVFVATERLF
SRILKILVQPVSDFGAYKCSSSSGELGYVLVYSLYSIHWYCVKYDESQGKPVLRNLGC
KQFKRFVIVSASWSPHVTGECLLLLDNGEVFVFDLSQRHCRVRGCKLKVSWESQGKSV
NKSWLGCEFGWRVGVYIVARSDALFVIVKSTEDCSVRCLLEVESLNTAGAEVFVGFAK
AGSDGFRFVLASQSYVFLCDARSGVPLLKWQHDVEKPCFMDVYSLSELGVRTFESNTS
CLIIGSFWNAQSQMFCFGPSPSVGKDPSSLYVWELPHNLLLPVGKCLCGDCLFREVMI
KESLPEWIDWQKKSVLVLGFGVLNKYLPLGSSDQSSGFTLIRLTSSGKLEAVKFRASR
LKHLEVVAHKGSACKSDEVNLLYLPDDEEYKFPRRFNYLELEYLSAHRKGMLAGFLDS
KMRTESSDFKKSESFSLICHEELCKKLKICGFGKGRSASSITAVFENINSPTSVFDIA
LRETWSSLPKEILMLAFSNYSEFADVLVDKKKQSLEFLVVPEFPQLPPFLLRNPSSRS
SKWSKKEQPGVEVVGPVVPLPVLITLHEFHNGCLNSEQEFSPEAEFYNRCNQISKATR
QIANSGRHETTISLDEDRADEMWLNSDSQEEKKTFIAYRPITKTAESDRLQQEVTTFV
SRIRGCKEGDDNAVGRRGLELFDELSPVEMFFENREVNFDKFDMKAMLTDKTFHSQWQ
DRSSSYQEFLSQYHLKK"

CDS complement(join(7342. .7991,8069. .8099,8244. .8498))
/note="unnamed protein product; gb|AAC69119.1
gene_id:MYF24.4
strong similarity to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01098.1"
/db_xref="GI:11994095"
/translation="MTLPELPKDLVEEILCFVPATSLKRLRSTCKGWNRLFKDDKRFA
RKHIEKAAKQFQPLTLTKNYRICPINVNLHGTTPSLEVKNEWIELGDRYNEDKDSEIY
EFSSDSWRVIDDIIKPPHYMEYFRKCLSLKGNTYWLGIDRRRRPPDLRITLIKFDFGT
EKFGYVPLPPPCQVHGFEASRLSAVGDEKLSLLLVGDSTSNTELWVTSKIGEANVVSW
SKVLSLYPKPNVGFWHGLSFLLDEEKKVLLCCKSKGWMKEEDEDNVYIVGEDTKFILL
NFGVQTIGGYSPIIVNYVPSLGQIELAGSKRKRDY"

CDS complement(9242. .10372)
/note="unnamed protein product; gb|AAC69119.1
gene_id:MYF24.5
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01099.1"
/db_xref="GI:11994096"
/translation="MPMPNLPKELVEEILSFVPATYLKRLSATCKPWNRLIHNDKRFA
RKHYDNAAKEFLVFMMRKNFRIIRRGVNLHGADPSAEVKGELTLPDPYFKNSADEFHI
DRVFHCDGLLLCTSKLERRMVVWNPLTGETKWIQTHEEGDNFFLGYSQEDKNISCKKS
YKIMGFYRSGSKVWEYDFNSDSWRVLNGILPNWYFDKSYKCVSLKGNTYMLAGAVTDM
GFDLSLQSYDFSTEKFAPVSLPVPSQARSLNGANRLSVVRGEKLALLYRRDKRSKAEI
WVTNKIDDTTEGAVSWTKVLELDLSRELHALFTSNFLVDEEKKVFICCVSWKEDEDEN
KSNKVYIVGEDNIVKEIDSGEDATSGCEPTILSLYVPSLVYM"

CDS join(12133. .12216,12398. .13249)
/note="unnamed protein product; gb|AAC69117.1
gene_id:MYF24.6
strong similarity to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01100.1"
/db_xref="GI:11994097"
/translation="MASGKLPWELEEEILCRLPPGSLVRLRSAEKWAFTRITACNGLL
FRDFWNQGVTIWNPWLRQVGWIEYKEDKDFRFCGIGYDAGKPEKGYKIFGYFNRFYDT
KLKIGHRFAIFECASQAFKFIDSPEWPTLAGRGEYVSLNGNLYWTAYNEDTREHFLGS
FDFSTEISMHFCLLPCAKHVSGLQDKLVLTVFKGDRFALLKQSRISSNTEIWVTKDKI
NSSNNVVWLNLMTLSIPDFPSLFHQLSDISYFIHDMTLILCCDDNQTGVGCIFIARGD
LCKKIQINYVSLGFSQCVYLPSLTSVPLEFRSLQV"

CDS join(15639. .15823,16070. .16304,16383. .16552,16639. .16778,
16849. .17026,17139. .17476,17622. .17866,17947. .18303,
18385. .18615)
/note="unnamed protein product; gb|AAD20392.1
gene_id:MYF24.7
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01101.1"
/db_xref="GI:11994098"
/translation="MSSKTRNMLEGLVRDTSFKWLLGKQSSFDEEIEEMGRSPSAGTN
WIPELSPIANVVVRRCSKILGVSANELRDSFKQEAFESLKQPSLFPRNFLEYCCFRAL
SLSVGVTGHLADKKFRRLTFDMMVVWEVPAVASQALLSVEEDATVSLEAFSRIAPAVP
IIADVIICDNLFQMLTSSTGGRLQFSVYDKYLHGLERAIKKMRTQSESSLLSGVRSKR

EKILEIDGTVTTQPVLEHVGISTWPGRLILTDHSLYFEALKVVSYDTPKRYHLSEDLK
QIIKPELTGPWGTRLFDKAVSYQSISLSEPVVMEFPELKGHTRRDYWLTIIQEVLYVH
RYINKYKITGLARDEALSKAVLGVMRVQALQELNLTNAMRYENLLPFNLCDQLPGGDL
ILETLAEMSTSRELHRSNKSKDTGTLHSSASDMVSQLGSVFGGSSPRSRRETSSLVVG
EVVVGDVNPLERAVKESRKKYEKVVLAQETINGVKMGGIDTNLAVMKELMLPIMETWN
LILSVVYWDDPTKSSVFCLLTTFIIWRGWLVYVFALASLFSAIPMVLTRCFSREKLMI
ELKVTAPPPMNTMEQLLAVQNGISELEQNIQDANIVLLKFRALLFSLFPPQASQKFAIA
IVVAATMMAFVPGRYLLSVVFVELFTRYSPPRRASTERLIRRLREWWFSIPAAPVVLL
HDKNNKKKKN"

CDS complement(18942. .19799)
/note="unnamed protein product; gb|AAF16598.1
gene_id:MYF24.8
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01102.1"
/db_xref="GI:11994099"
/translation="MSSTYKDNHPYHHHPHHHHHHPKREPNENNSGILYTPSPSPSPP
TLKVNKDSHVIKKPPSPSSFSSAAKPRHPVIIYTHTPRIIHTNPKDFMALVQKLTGMT
HSDEDLGGGNAMTDPGVVKSINRTVSEPTVDRKNNRNRCGGGNNYLRNYSGAGNGKGI
FFNNTMISEDSESSSVITTEENIGEHGQVNSSLPYSAVAIPPQPPPHPPPPPPPPSMY
DAAGINYGAYLPTFPIFPPANPADNFLCGNQPFANFDDPLFFAPNMRSSFSSSSSSGF
DGLTEFRDF"

CDS join(24477. .24671,25127. .25201,25295. .25402,25514. .25546,
25816. .25941,26022. .26165,26254. .26664,26744. .26878,
26982. .27200,27291. .27551,27631. .27759,27850. .28318,
28416. .28471)
/note="unnamed protein product; gene_id:MYF24.9
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01103.1"
/db_xref="GI:11994100"
/translation="MGRRIKRKGLINTEAAREFINHLVAERHSLLLLVPLVLAFWAIE
RWVFAFSNWVPLVVAVWASLQYGSYQRALLAEDLTKKWRQTVFNASTITPLEHCQWLN
KLLSEIWLNYMNKKLSLRFSSMVEKRLRQRRSRLIKIMRLDFNWDTTDLSILLQAKLS
MPFNRTARIVVNSLCIKGDILIRPILEGRALLYSFVSNPEVRIGVAFGGGGGQSLPAT
ELPGVSSWLVKILTETLNKKMVEPRRGCFSLPATDLHKTAIGGIIYVTVVSGNNLNRR
ILRGSPSKSSEIGEGSSGNSSSKPVQTFVEVELEQLSRRTEMKSGPNPAYQSTFNMIL
HDNTGTLKFNLYENNPGSVRYDSLASCEVKMKYVGDDSTMFWAVGSDNGVIAKHAEFC
GQEIEMVVPPEGVSSGELTVRLLLKEWHFSDGSHSLNSVNSSSLHSLDSSSALLSKTG
RKIIVTVLAGKNLVSKDKSGKCDASVKLQYGKIIQKTKIVNAAECVWNQKFEFEELAG
EEYLKVKCYREEMLGTDNIGTATLSLQGINNSEMHIWVPLEDVNSGEIELLIEALDPE
YSEADSSKGLIELVLVEARDLVAADIRGTSDPYVRVQYGEKKQRTKVIYKTLQPKWNQ
TMEFPDDGSSLELHVKDYNTLLPTSSIGNCVVEYQGLKPNETADKWIILQGVKHGEVH
VRVTRKVTEIQRRASAGPGTPFNKALLLSNQMKQVMIKFQNLIDDGDLEGLAEALEEL
ESLEDEQEQYLLQLQTEQSLLINKIKDLGSGSGTGSSYNRRLPAPM"

CDS complement(join(29117. .29383,29522. .29641,29726. .29869,
29953. .30004,30121. .30196,30362. .30540,30667. .30708,
30805. .30896))
/note="unnamed protein product; gb|AAD39678.1
gene_id:MYF24.10
similar to unknown protein"
/codon_start=1

Alignment Scores:

Pred. No.:	1.76e+04	Length:	88989
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	15	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AB026658 (1-88989)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db       6166 AGTTTGATTACTTGGAACATA 6186
```

RESULT 27
HS30M3
LOCUS HS30M3 102200 bp DNA linear PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone RP1-30M3 on chromosome 6p22.1-22.3

```
            Contains the 5' end of gene KIAA0319, the gene for TRAF and TNF
            receptor-associated protein (TTRAP)(EAP2, AD022, MGC9099), gene
            FLJ20501, the 3' end of gene FLJ12619 (DKFZP564G182) and two CpG
            islands, complete sequence.
ACCESSION   AL031775
VERSION     AL031775.1  GI:4071041
KEYWORDS    HTG; AD022; CpG island; DKFZP564G182; FLJ12619; FLJ20501; KIAA0319;
            TNF; TRAF; TTRAP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 102200)
  AUTHORS   Phillips,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 29, 1998 this sequence version replaced gi:4007165.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP1-30M3 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES                Location/Qualifiers
     source             1. .102200
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="RZPD:RPCIP704M0330"
                        /db_xref="taxon:9606"
                        /chromosome="6"
                        /map="p22.1-22.3"
                        /clone="RP1-30M3"
                        /clone_lib="RPCI-1"
     gene               join(complement(AL133264.10:4611. .6907),
                        complement(AL133264.10:2219. .2395),
                        complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                        /gene="C6orf62"
                        /locus_tag="RP1-30M3.4-001"
     mRNA               join(complement(AL133264.10:4611. .6907),
                        complement(AL133264.10:2219. .2395),
                        complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                        /gene="C6orf62"
                        /locus_tag="RP1-30M3.4-001"
                        /product="chromosome 6 open reading frame 62"
                        /note="match: ESTs: AA084220 AA093151 AA168907 AA197971
                        AA200765 AA237701 AA292313 AA295166 AA402388 AA478637
                        AA591228 AA612170 AA638148 AA647423 AA647435 AA647742
                        AA672357 AA798123 AA798467 AA815944 AA833812 AA863774
                        AI053561 AI054330 AI122327 AI156738 AI191279 AI194277
                        AI284633 AI316639 AI390324 AI458673 AI688647 AI694146
                        AI694199 AI788240 AI802319 AI807413 AI831412 AI932363
                        AI935446 AI982078 AJ396992 AJ398211 AL529211 AL563388.1
                        AL601897.1 AU051650 AU125496 AU135776 AU137791 AU138464
                        AV656408 AV686324 AV698092 AV703346 AV708342 AV708873
                        AV709145 AV709526 AV739602 AV742411 AV743612 AV744587
                        AV745018 AW050065 AW063931 AW270998 AW300107 AW300121
                        AW301984 AW302318 AW302718 AW302841 AW302861 AW303029
                        AW355281 AW822881 AW859606 AW959243 BB067326 BB123906
                        BB609972 BB624305 BB628344 BB654156 BB823708 BE007675
                        BE049029 BE049182 BE120619 BE178692 BE178693 BE244058
                        BE246922 BE287246 BE289710 BE311050 BE369398 BE377181
                        BE386264 BE465032 BE484573 BE554360 BE571717 BE625094
                        BE766876 BE766886 BE770856 BE785982.1 BE813002 BE817780
                        BE829341 BE829344 BE829355 BE829359 BE893564 BE895461
                        BE896141 BE908143 BE913030 BE945891 BE991271 BE992722
                        BF015474 BF021751 BF022025 BF029248 BF037603 BF060909
                        BF064161 BF109576 BF111215 BF122334 BF137583 BF139381
                        BF140373 BF145038 BF148549 BF150174 BF165619 BF207643
                        BF326037 BF326041 BF326042 BF326045 BF334390 BF341149
                        BF376614 BF378829 BF380539 BF380550 BF465841 BF535383
                        BF578793 BF582307 BF748822 BF748938 BF748963 BF765346
                        BF766403 BF784199 BF789335 BF789696 BF885392 BF896088
                        BF940327 BF984391 BG030354 BG035327 BG088552 BG116557
                        BG148023 BG167018 BG168325 BG247693 BG252842 BG295272
                        BG295445 BG342500 BG610013 BG655426 BG667600 BG677384.1
                        BG678705 BG692520 BG717389 BG801183 BG802707 BG803992
                        BG863302 BG914583 BG917993 BG940215 BG943615 BG975946
                        BI080618 BI092275 BI106992 BI107793 BI108344 BI222784
                        BI258149 BI326830 BI651746 BI659175 BI666111 BI688353
                        BI689042 BI694119 BI696541 BI696575 BI696638 BI733610
                        BI753721 BI754669 BI851772 BI856689.1 BI872006 BI872428
                        BI903238 BI903250 BI960823 BI986171 BM016286 BM069622
                        BM069876 BM147734 BM237580 BM238719 BM238848 BM423047
                        BM423107 F11344 F20959 H17162 H98057 N56057 R25494 T11219
                        T51727 W47233
                        match: cDNAs: AB046018.1 AJ420475.1 AK022681.1 AL136632.1
                        BC002217.1 BC005537.1"
     gene               join(complement(AL133264.10:101. .211),5119. .5253,
                        7768. .8194)
                        /gene="C6orf62"
                        /locus_tag="RP1-30M3.4-002"
     mRNA               join(complement(AL133264.10:101. .211),5119. .5253,
                        7768. .8194)
                        /gene="C6orf62"
                        /locus_tag="RP1-30M3.4-002"
                        /product="chromosome 6 open reading frame 62"
                        /note="match: ESTs: BG772628 BG772628.1"
     gene               join(complement(AL133264.10:5931. .6069),
                        complement(AL133264.10:2219. .2395),
                        complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
                        /gene="C6orf62"
                        /locus_tag="RP1-30M3.4-003"
     mRNA               join(complement(AL133264.10:5931. .6069),
                        complement(AL133264.10:2219. .2395),
                        complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
                        /gene="C6orf62"
                        /locus_tag="RP1-30M3.4-003"
                        /product="chromosome 6 open reading frame 62"
                        /note="match: ESTs: BG342500.1 BG692520.1 BG803992.1"
     CDS                join(complement(AL133264.10:4611. .4739),
                        complement(AL133264.10:2219. .2395),
                        complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
                        /gene="C6orf62"
                        /locus_tag="RP1-30M3.4-003"
                        /standard_name="OTTHUMP00000016093"
                        /note="match: proteins: O95550 Q961D0 Q99K03 Q99LU8 Q9GZU0
                        Q9H1Z2 Q9N0B2 Q9QEJ3 Q9W384 Q9XWI7"
                        /codon_start=1
                        /product="chromosome 6 open reading frame 62"
                        /protein_id="CAI20135.1"
                        /db_xref="GI:56202469"
                        /translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
                        KKSALFEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
```

```
                     DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEPHHGDYEKQFLHV
                     LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
                     IEDHLRPYMPE"
     CDS             join(complement(AL133264.10:5931. .5972),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-003"
                     /standard_name="OTTHUMP00000016094"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 62"
                     /protein_id="CAI20136.1"
                     /db_xref="GI:56202470"
                     /db_xref="UniProt/TrEMBL:Q5JVZ2"
                     /translation="MSENLSDPVSPVVRKKKSALFEVSEVIPVMTNNYEENILKGVRD
                     SSYSLESSLELLQKDVVQLHAPRYQSMRRDVIGCTQEMDFILWPRNDIEKIVCLLFSR
                     WKESDEPFRPVQAKFEPHHGDYEKQFLHVLSRKDKTGIVVNNPNQSVFLFIDRQHLQT
                     PKNKATIFKLCSICLYLPQEQLTHWAVGTIEDHLRPYMPE"
     polyA_signal    9124. .9129
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-001"
     gene            complement(join(12316. .12571,15963. .16147,46686. .46959))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-001"
     mRNA            complement(join(12316. .12571,15963. .16147,46686. .46959))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-001"
                     /product="thioesterase superfamily member 2"
                     /note="match: ESTs: AA614605 AA928745 AI017897 AI052396
                     BF407527 BF475709 BF594161
                     match: cDNAs: AF155649.1 AF220186.1 AF274952.1 AK000508.1
                     AK002261.1 AK008624.1 BC000894.1"
     gene            complement(join(12316. .12571,15963. .16147,26337. .26628))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-002"
     mRNA            complement(join(12316. .12571,15963. .16147,26337. .26628))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-002"
                     /product="thioesterase superfamily member 2"
                     /note="match: ESTs: BG219890 BG256069"
     polyA_site      12318
     polyA_signal    12336. .12341
     CDS             complement(join(12415. .12571,15963. .16147,46686. .46766))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-001"
                     /standard_name="OTTHUMP00000016090"
                     /note="match: proteins: AAA28024 AAH18165 AAL48890 O95549
                     P34419 P87304 P93828 P95914 Q18187 Q9ASQ1 Q9NPJ3 Q9VZZ5
                     Q9VZZ6"
                     /codon_start=1
                     /product="thioesterase superfamily member 2"
                     /protein_id="CAD92509.1"

Alignment Scores:
Pred. No.:                1.99e+04        Length:         102200
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x HS30M3 (1-102200)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      55387 TCTCTCATTACCTGGAATATT 55407

RESULT 28
AC093026/c
LOCUS       AC093026              114468 bp    DNA     linear   PRI 30-JUL-2002
DEFINITION  Homo sapiens 12 BAC RP11-307E8 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC093026
VERSION     AC093026.3  GI:20334486
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 114468)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
            Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
            Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
            Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P.,
            Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M.,
            Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S.,
            Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
            Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
            King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
            Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O.,
            Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X.,
            Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P.,
            Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
            Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S.,
            Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
            Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
            Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 114468)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-AUG-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 114468)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 114468)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   5  (bases 1 to 114468)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
```

```
  JOURNAL    Submitted (30-JUL-2002) Human Genome Sequencing Center, Department
             of Molecular and Human Genetics, Baylor College of Medicine, One
             Baylor Plaza, Houston, TX 77030, USA
COMMENT      On Apr 27, 2002 this sequence version replaced gi:18853586.
             INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
             gc-help@bcm.tmc.edu

             CLONE LENGTH: This sequence does not necessarily represent the
             entire insert of this clone. Overlapping regions of clones are only
             sequenced and submitted once, so the sequence for the remainder of
             the insert may be found in the record for the adjacent clones.
             Overlapping clones are noted at the beginning and end of the
             Features listing.

             ANNOTATION OF FEATURES:
                STSs are identified using ePCR (Genome Res. 7:541-550) searches
             of a local database that includes entries from dbSTS, GDB, and
             local mapping efforts.
                Repeats are identified using RepeatMasker (A. Smit and P. Green,
             unpublished.) for Human and Mouse sequences.
                Genes and Region of sequence similarity are identified by BLAST
             (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
             EST and cDNA sequences. Genes demonstrate at least two exons
             flanked by consensus splice sites that maintained sequence
             continuity across the splice junctions. Sequences that are not
             identical matches are annotated as similar.

             SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
             standard of double strand coverage with a minimum of 2 clones and 2
             reads with no ambiguities or 2 chemistries with a minimum of 2
             clones and 3 reads with no ambiguities. If the sequence quality for
             a region does not meet this standard, it will be indicated in the
             annotation as Low Coverage.

             QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
             standards - estimated error rate less than 1 per 10,000 bases.
             Reports of lowest quality individual bases and measures of base
             quality are listed below. Description of the metrics can be found
             at URL:
             http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

             QUALSTAT-REPORT.
FEATURES             Location/Qualifiers
     source          1. .114468
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="RP11-307E8"
     misc_feature    1. .2000
                     /note="overlaps bases 9182. .11181 of clone AC128679"
                     /function="clone overlap"
     repeat_region   67. .89
                     /rpt_family="(TAAA)n"
     repeat_region   130. .151
                     /rpt_family="AT_rich"
     repeat_region   164. .186
                     /rpt_family="AT_rich"
     repeat_region   1084. .1380
                     /rpt_family="AluJb"
     repeat_region   1381. .1547
                     /rpt_family="L1MA1"
     repeat_region   1551. .1591
                     /rpt_family="AT_rich"
     repeat_region   1605. .1751
                     /rpt_family="FRAM"
     repeat_region   1752. .1812
                     /rpt_family="GA-rich"
     repeat_region   complement(1896. .2042)
                     /rpt_family="MER3"
     repeat_region   complement(2793. .3112)
                     /rpt_family="AluJb"
     repeat_region   3647. .3682
                     /rpt_family="(TG)n"
     repeat_region   complement(3798. .4109)
                     /rpt_family="AluSx"
     repeat_region   4322. .4427
                     /rpt_family="L2"
     repeat_region   complement(4508. .4734)
                     /rpt_family="MIR"
     repeat_region   5203. .5308
                     /rpt_family="MER102"
     repeat_region   complement(6610. .6919)
                     /rpt_family="AluJo"
     repeat_region   6928. .6971
                     /rpt_family="AT_rich"
     repeat_region   7458. .7477
                     /rpt_family="(TTTTC)n"
     repeat_region   complement(7484. .7616)
                     /rpt_family="AluY"
     repeat_region   complement(7675. .8479)
                     /rpt_family="LTR1B"
     repeat_region   8655. .8859
                     /rpt_family="L2"
     repeat_region   8966. .8995
                     /rpt_family="(TGAA)n"
     repeat_region   complement(10332. .10650)
                     /rpt_family="MER1B"
     repeat_region   complement(11031. .11242)
                     /rpt_family="MIR"
     repeat_region   complement(12260. .12381)
                     /rpt_family="MIR"
     repeat_region   12397. .12424
                     /rpt_family="(CAAAAA)n"
     repeat_region   13194. .13637
                     /rpt_family="LTR40b"
     repeat_region   complement(13667. .13800)
                     /rpt_family="MIR"
     repeat_region   complement(14211. .14471)
                     /rpt_family="L2"
     repeat_region   complement(14487. .14983)
                     /rpt_family="L2"
     repeat_region   complement(15016. .15117)
                     /rpt_family="MIR"
     repeat_region   15118. .15529
                     /rpt_family="MSTC"
     repeat_region   complement(15530. .15586)
                     /rpt_family="MIR"
     repeat_region   complement(16271. .16560)
                     /rpt_family="AluSx"
     repeat_region   16651. .16791
                     /rpt_family="L2"
     repeat_region   16988. .17112

Alignment Scores:
Pred. No.:                2.2e+04        Length:          114468
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       8              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC093026 (1-114468)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      77352 TCCCTTATCACATGGAATATA 77332

RESULT 29
AC135637/c
LOCUS       AC135637              130323 bp    DNA     linear   ROD 27-FEB-2004
DEFINITION  Mus musculus BAC clone RP23-298I23 from chromosome 14, complete
            sequence.
ACCESSION   AC135637
VERSION     AC135637.5  GI:42544095
KEYWORDS    HTG.
```

```
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 130323)
  AUTHORS   Shahid,S., Bielicki,L. and Meyer,R.
  TITLE     The sequence of Mus musculus BAC clone RP23-298I23
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 130323)
  AUTHORS   Wilson,R.
  TITLE     Sequencing of Mus musculus
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 130323)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-OCT-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 130323)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-FEB-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   5  (bases 1 to 130323)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   6  (bases 1 to 130323)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-2004) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Feb 13, 2004 this sequence version replaced gi:27923715.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: M_BA0298I23
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone, fosmid clone or direct clone walk sequence.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to obtain the consensus sequence; and
            the assembly was confirmed by  restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
            and Minako Tateno in the laboratory of Pieter de Jong
            (http://www.chori.org) from female C57BL/6J mouse kidney and/or
            brain genomic DNA.  The clone and detailed information can be
            obtained from Research Genetics, Inc. (http://www.resgen.com) or
            Pieter de Jong and coworkers at http://www.chori.org

            NEIGHBORING SEQUENCE INFORMATION:
            This sequence is the entire insert of the clone. This clone is
            overlapped by AC121567.
FEATURES              Location/Qualifiers
     source           1. .130323
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:10090"
                      /chromosome="14"
                      /map="14"
                      /clone="RP23-298I23"
                      /clone_lib="RPCI-23"
     misc_feature     13033. .13089
                      /note="Sequence derived from PCR product of project DNA."
     tRNA             73241. .73312
                      /product="tRNA-Ser"
                      /note="Likely pseudogene (HMM Sc=33.19 / Sec struct
                      Sc=-10.52)"
     tRNA             complement(93675. .93746)
                      /product="tRNA-Ser"
                      /note="Likely pseudogene (HMM Sc=33.94 / Sec struct
                      Sc=-11.27)"
     unsure           102229. .102770
                      /note="Unresolved simple sequence repeat."
ORIGIN

Alignment Scores:
Pred. No.:              2.48e+04        Length:         130323
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC135637 (1-130323)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     110459 TCACTTATAACCTGGAATATA 110439

RESULT 30
AC150890/c
LOCUS       AC150890              131005 bp    DNA     linear   HTG 03-JUN-2005
DEFINITION  Medicago truncatula chromosome 2 clone mth2-74h3, *** SEQUENCING IN
            PROGRESS ***, 12 unordered pieces.
ACCESSION   AC150890
VERSION     AC150890.3  GI:66912344
KEYWORDS    HTG; HTGS_PHASE1; HTGS_ACTIVEFIN.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 131005)
  AUTHORS   Town,C.D., Tallon,L.J., Arbogast,T., Althoff,R., Hine,E.,
            Monaghan,E., Smith,S.A., Utterback,T., Peldblyum,T., Koo,H. and
            Cheung,F.
  TITLE     Medicago truncatula BAC genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 131005)
  AUTHORS   Town,C.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-AUG-2004) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   3  (bases 1 to 131005)
  AUTHORS   Town,C.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-2005) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT     On Jun 3, 2005 this sequence version replaced gi:61967003.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
```

```
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    17642: contig of 17642 bp in length
            *    17643    17742: gap of unknown length
            *    17743    20485: contig of 2743 bp in length
            *    20486    20585: gap of unknown length
            *    20586    23045: contig of 2460 bp in length
            *    23046    23145: gap of unknown length
            *    23146    49182: contig of 26037 bp in length
            *    49183    49282: gap of unknown length
            *    49283    52132: contig of 2850 bp in length
            *    52133    52232: gap of unknown length
            *    52233    77714: contig of 25482 bp in length
            *    77715    77814: gap of unknown length
            *    77815    80269: contig of 2455 bp in length
            *    80270    80369: gap of unknown length
            *    80370    87965: contig of 7596 bp in length
            *    87966    88065: gap of unknown length
            *    88066    90240: contig of 2175 bp in length
            *    90241    90340: gap of unknown length
            *    90341   105239: contig of 14899 bp in length
            *   105240   105339: gap of unknown length
            *   105340   119636: contig of 14297 bp in length
            *   119637   119736: gap of unknown length
            *   119737   131005: contig of 11269 bp in length.
FEATURES             Location/Qualifiers
     source          1. .131005
                     /organism="Medicago truncatula"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3880"
                     /chromosome="2"
                     /clone="mth2-74h3"
     gap             17643. .17742
                     /estimated_length=unknown
     gap             20486. .20585
                     /estimated_length=unknown
     gap             23046. .23145
                     /estimated_length=unknown
     gap             49183. .49282
                     /estimated_length=unknown
     gap             52133. .52232
                     /estimated_length=unknown
     gap             77715. .77814
                     /estimated_length=unknown
     gap             80270. .80369
                     /estimated_length=unknown
     gap             87966. .88065
                     /estimated_length=unknown
     gap             90241. .90340
                     /estimated_length=unknown
     gap             105240. .105339
                     /estimated_length=unknown
     gap             119637. .119736
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:              2.49e+04        Length:         131005
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC150890 (1-131005)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     101655 TCCTTGATAACGTGGAATATT 101635
```

```
RESULT 31
HSAC000383
LOCUS       HSAC000383            135269 bp    DNA     linear   HTG 13-JUN-2002
DEFINITION  Homo sapiens chromosome 11 clone pDJ227b23, *** SEQUENCING IN
            PROGRESS ***, 19 unordered pieces.
ACCESSION   AC000383
VERSION     AC000383.2  GI:21406271
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 135269)
  AUTHORS   Evans,G.A., Bradbury,P., Brignac,S., Bumeister,R., Burbee,D.,
            Davie,J., Davies,C.J., Davis,C., English,C., Fondon,T.,
            Franklin,T.L., Garner,H.R., Gordon,M., Gotway,G., Grant,O.,
            Hahner,L., Harris,J., Hinson,S., Megarity,C., Narayanaswamy,U.,
            Newton,J., O'Brien,K., Oliver,T., Patel,P., Probst,S., Rayner,S.,
            Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
            Ward,T. and Wilson,R.
  TITLE     Template
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 135269)
  AUTHORS   Evans,G.A., Bradbury,P., Brignac,S., Bumeister,R., Burbee,D.,
            Davie,J., Davies,C.J., Davis,C., English,C., Fondon,T.,
            Franklin,T.L., Garner,H.R., Gordon,M., Gotway,G., Grant,O.,
            Hahner,L., Harris,J., Hinson,S., Megarity,C., Narayanaswamy,U.,
            Newton,J., O'Brien,K., Oliver,T., Patel,P., Probst,S., Rayner,S.,
            Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
            Ward,T. and Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-1997) Genome Science and Technology Center,
            University of Texas Southwestern Medical Center at Dallas, 5323
            Harry Hines Blvd, Dallas, TX 75235-8591, USA
COMMENT     On Jun 13, 2002 this sequence version replaced gi:1927149.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 19 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3136: contig of 3136 bp in length
            *     3137     3236: gap of 100 bp
            *     3237     5627: contig of 2391 bp in length
            *     5628     5727: gap of 100 bp
            *     5728     8251: contig of 2524 bp in length
            *     8252     8351: gap of 100 bp
            *     8352    12456: contig of 4105 bp in length
            *    12457    12556: gap of 100 bp
            *    12557    15314: contig of 2758 bp in length
            *    15315    15414: gap of 100 bp
            *    15415    18745: contig of 3331 bp in length
            *    18746    18845: gap of 100 bp
            *    18846    22715: contig of 3870 bp in length
            *    22716    22815: gap of 100 bp
            *    22816    26735: contig of 3920 bp in length
            *    26736    26835: gap of 100 bp
            *    26836    29292: contig of 2457 bp in length
            *    29293    29392: gap of 100 bp
            *    29393    33341: contig of 3949 bp in length
            *    33342    33441: gap of 100 bp
            *    33442    37782: contig of 4341 bp in length
            *    37783    37882: gap of 100 bp
            *    37883    45713: contig of 7831 bp in length
            *    45714    45813: gap of 100 bp
            *    45814    54572: contig of 8759 bp in length
            *    54573    54672: gap of 100 bp
            *    54673    62773: contig of 8101 bp in length
            *    62774    62873: gap of 100 bp
            *    62874    72447: contig of 9574 bp in length
```

```
            *    72448    72547: gap of 100 bp
            *    72548    85629: contig of 13082 bp in length
            *    85630    85729: gap of 100 bp
            *    85730   100997: contig of 15268 bp in length
            *   100998   101097: gap of 100 bp
            *   101098   116202: contig of 15105 bp in length
            *   116203   116302: gap of 100 bp
            *   116303   135269: contig of 18967 bp in length.
FEATURES             Location/Qualifiers
     source          1. .135269
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /clone="pDJ227b23"
     gap             3137. .3236
                     /estimated_length=100
     gap             5628. .5727
                     /estimated_length=100
     gap             8252. .8351
                     /estimated_length=100
     gap             12457. .12556
                     /estimated_length=100
     gap             15315. .15414
                     /estimated_length=100
     gap             18746. .18845
                     /estimated_length=100
     gap             22716. .22815
                     /estimated_length=100
     gap             26736. .26835
                     /estimated_length=100
     gap             29293. .29392
                     /estimated_length=100
     gap             33342. .33441
                     /estimated_length=100
     gap             37783. .37882
                     /estimated_length=100
     gap             45714. .45813
                     /estimated_length=100
     gap             54573. .54672
                     /estimated_length=100
     gap             62774. .62873
                     /estimated_length=100
     gap             72448. .72547
                     /estimated_length=100
     gap             85630. .85729
                     /estimated_length=100
     gap             100998. .101097
                     /estimated_length=100
     gap             116203. .116302
                     /estimated_length=100
ORIGIN

Alignment Scores:
Pred. No.:                2.56e+04        Length:         135269
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x HSAC000383 (1-135269)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      61295 TCATTGATAACTTGGAATATA 61315

RESULT 32
AC124017/c
LOCUS       AC124017               138271 bp    DNA     linear   PRI 27-JAN-2003
DEFINITION  Homo sapiens BAC clone RP11-317M11 from 4, complete sequence.
ACCESSION   AC124017 AC016488
VERSION     AC124017.2  GI:27819589
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 138271)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   PUBMED   9847074
REFERENCE   2  (bases 1 to 138271)
  AUTHORS   Cordum,H., Haglund,K., Elliott,G. and Delaney,K.
  TITLE     The sequence of Homo sapiens BAC clone RP11-317M11
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 138271)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 138271)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO, 63108, USA
REFERENCE   5  (bases 1 to 138271)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JAN-2003) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6  (bases 1 to 138271)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Jan 22, 2003 this sequence version replaced gi:21327652.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0317M11
                Drafting Center: WIBR
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO. For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
```

```
          libraries. Genomics 51:1-8. The clone may be obtained either from
          Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
          and coworkers at http://www.chori.org
          VECTOR:  pBACe3.6

          NEIGHBORING SEQUENCE INFORMATION:
          This sequence is not the entire insert of the clone. This clone is
          overlapped by AC095040 and AC105384.

          The sequence of AC016488 has been incorporated into AC124017.
FEATURES             Location/Qualifiers
     source          1. .138271
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /map="4"
                     /clone="RP11-317M11"
                     /clone_lib="RPCI-11"
     repeat_region   4. .107
                     /rpt_family="L1"
     repeat_region   108. .372
                     /rpt_family="Alu"
     repeat_region   1065. .1275
                     /rpt_family="MER1_type"
     repeat_region   1279. .1575
                     /rpt_family="Alu"
     repeat_region   1744. .1789
                     /rpt_family="CT-rich"
     repeat_region   1792. .2020
                     /rpt_family="MIR"
     repeat_region   2242. .2556
                     /rpt_family="Alu"
     repeat_region   4355. .4445
                     /rpt_family="MER1_type"
     repeat_region   4612. .4695
                     /rpt_family="L2"
     repeat_region   4698. .4829
                     /rpt_family="MER1_type"
     repeat_region   5502. .5534
                     /rpt_family="CT-rich"
     repeat_region   5721. .6737
                     /rpt_family="L1"
     repeat_region   7392. .7624
                     /rpt_family="MIR"
     repeat_region   7665. .8068
                     /rpt_family="L2"
     repeat_region   8546. .8732
                     /rpt_family="L1"
     repeat_region   9040. .9283
                     /rpt_family="MIR"
     repeat_region   10148. .10201
                     /rpt_family="(GA)n"
     repeat_region   11192. .11359
                     /rpt_family="MIR"
     repeat_region   13375. .14172
                     /rpt_family="L2"
     repeat_region   14328. .14409
                     /rpt_family="L2"
     repeat_region   15459. .15759
                     /rpt_family="Alu"
     repeat_region   16373. .16491
                     /rpt_family="MIR"
     repeat_region   17735. .17762
                     /rpt_family="AT_rich"
     repeat_region   17962. .18501
                     /rpt_family="MaLR"
     repeat_region   18502. .18594
                     /rpt_family="MaLR"
     repeat_region   18596. .18972
                     /rpt_family="L1"
     repeat_region   19177. .19330
                     /rpt_family="MaLR"
     repeat_region   19331. .19679
                     /rpt_family="L1"
     repeat_region   20264. .20338
                     /rpt_family="MIR"
     repeat_region   20454. .20854
                     /rpt_family="MaLR"
     repeat_region   21574. .21857
                     /rpt_family="ERVL"
     repeat_region   21878. .21919
                     /rpt_family="Mariner"
     repeat_region   28404. .28459
                     /rpt_family="Mariner"
     repeat_region   28497. .28953
                     /rpt_family="ERVL"
     repeat_region   28955. .29098
                     /rpt_family="CT-rich"
     repeat_region   29124. .29159
                     /rpt_family="ERVL"
     repeat_region   29160. .29502
                     /rpt_family="ERVL"
     repeat_region   30665. .30733
                     /rpt_family="MER1_type"
     repeat_region   30682. .30742
                     /rpt_family="MER1_type"
     repeat_region   31091. .31182
                     /rpt_family="CR1"
     repeat_region   31713. .31867
                     /rpt_family="L1"
     repeat_region   32103. .32350
                     /rpt_family="Alu"
     repeat_region   32717. .32822
                     /rpt_family="MIR"
     repeat_region   33667. .33816
                     /rpt_family="MIR"
     repeat_region   33835. .34135
                     /rpt_family="L1"
     repeat_region   34358. .34477
                     /rpt_family="MIR"
     repeat_region   34470. .34537
                     /rpt_family="L2"
     repeat_region   34918. .35007
                     /rpt_family="L2"
     repeat_region   35287. .35681
                     /rpt_family="MaLR"
     repeat_region   35696. .35806
                     /rpt_family="MIR"
     repeat_region   35825. .35927
                     /rpt_family="L1"
     repeat_region   35936. .36125
                     /rpt_family="L1"

Alignment Scores:
Pred. No.:              2.61e+04        Length:         138271
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC124017 (1-138271)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      72693 TCACTTATCACCTGGAACATC 72673

RESULT 33
AC021985/c
LOCUS       AC021985               149416 bp    DNA     linear   HTG 05-SEP-2000
DEFINITION  Homo sapiens clone RP11-3H20, WORKING DRAFT SEQUENCE, 8 unordered
            pieces.
ACCESSION   AC021985
VERSION     AC021985.3  GI:9966964
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 149416)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-3H20
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 149416)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 5, 2000 this sequence version replaced gi:9280768.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L2743
                Center clone name: 3_H_20
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 145817 bases at least Q40
                Consensus quality: 147697 bases at least Q30
                Consensus quality: 148336 bases at least Q20
                Insert size: 153000; agarose-fp
                Insert size: 148716; sum-of-contigs
                Quality coverage: 5.3 in Q20 bases; agarose-fp
                Quality coverage: 5.5 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      7330: contig of 7330 bp in length
            *     7331      7430: gap of 100 bp
            *     7431     44177: contig of 36747 bp in length
            *    44178     44277: gap of 100 bp
            *    44278     51152: contig of 6875 bp in length
            *    51153     51252: gap of 100 bp
            *    51253     68757: contig of 17505 bp in length
            *    68758     68857: gap of 100 bp
            *    68858     91289: contig of 22432 bp in length
            *    91290     91389: gap of 100 bp
            *    91390    120026: contig of 28637 bp in length
            *   120027    120126: gap of 100 bp
            *   120127    147727: contig of 27601 bp in length
            *   147728    147827: gap of 100 bp
            *   147828    149416: contig of 1589 bp in length.
FEATURES             Location/Qualifiers
     source          1. .149416
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-3H20"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .7330
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             7331. .7430
                     /estimated_length=100
     misc_feature    7431. .44177
                     /note="assembly_fragment"
     gap             44178. .44277
                     /estimated_length=100
     misc_feature    44278. .51152
                     /note="assembly_fragment"
     gap             51153. .51252
                     /estimated_length=100
     misc_feature    51253. .68757
                     /note="assembly_fragment"
     gap             68758. .68857
                     /estimated_length=100
     misc_feature    68858. .91289
                     /note="assembly_fragment"
     gap             91290. .91389
                     /estimated_length=100
     misc_feature    91390. .120026
                     /note="assembly_fragment"
     gap             120027. .120126
                     /estimated_length=100
     misc_feature    120127. .147727
                     /note="assembly_fragment"
     gap             147728. .147827
                     /estimated_length=100
     misc_feature    147828. .149416
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:                2.8e+04        Length:          149416
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       14             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC021985 (1-149416)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      25369 TCACTTATCACCTGGAACATC 25349

RESULT 34
AC152027/c
LOCUS       AC152027               150344 bp    DNA     linear   HTG 11-NOV-2004
DEFINITION  Dasypus novemcinctus clone VMRC5-12D21, WORKING DRAFT SEQUENCE, 5
            ordered pieces.
ACCESSION   AC152027
VERSION     AC152027.2  GI:55700150
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Dasypus novemcinctus (nine-banded armadillo)
  ORGANISM  Dasypus novemcinctus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Xenarthra; Dasypodidae; Dasypus.
```

```
REFERENCE   1  (bases 1 to 150344)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
            Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
            Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Laric,P.,
            Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L.,
            Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A.,
            Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,
            Shah,K., Sison,C., Stantripop,S., Stephen,E., Thomas,J.W.,
            Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
            Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625
            Fishers Lane, Rockville, MD 20852, USA
COMMENT     On Nov 11, 2004 this sequence version replaced gi:54312159.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: ifi
                 Center clone name: 012D21

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            -------------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 149480 bases at least Q40
                 Consensus quality: 149754 bases at least Q30
                 Consensus quality: 149909 bases at least Q20
                 Insert size: 155000; agarose-fp
                 Insert size: 149944; sum-of-contigs
                 Quality coverage: 9.52x in Q20 bases; agarose-fp
                 Quality coverage: 9.84x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    36273: contig of 36273 bp in length
            *    36274    36373: gap of unknown length
            *    36374    55049: contig of 18676 bp in length
            *    55050    55149: gap of unknown length
            *    55150   101939: contig of 46790 bp in length
            *   101940   102039: gap of unknown length
            *   102040   114806: contig of 12767 bp in length
            *   114807   114906: gap of unknown length
            *   114907   150344: contig of 35438 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150344
                     /organism="Dasypus novemcinctus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9361"
                     /clone="VMRC5-12D21"
                     /clone_lib="VMRC5"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .36273
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    1. .23436
                     /note="clone overlaps with GenBank Accession Number
                     AC152029 clone VMRC5-168B11 (center project name ifj)"
     gap             36274. .36373
                     /estimated_length=unknown
     misc_feature    36374. .55049
                     /note="assembly_fragment"
     gap             55050. .55149
                     /estimated_length=unknown
     misc_feature    55150. .101939
                     /note="assembly_fragment"
     gap             101940. .102039
                     /estimated_length=unknown
     misc_feature    102040. .114806
                     /note="assembly_fragment"
     gap             114807. .114906
                     /estimated_length=unknown
     misc_feature    114907. .150344
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:                2.82e+04        Length:          150344
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       14              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC152027 (1-150344)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      53255 TCTTTAATTACCTGGAATATT 53235

RESULT 35
AC007157/c
LOCUS       AC007157               152937 bp    DNA     linear   PRI 27-APR-1999
DEFINITION  Homo sapiens, clone hRPK.78_A_1, complete sequence.
ACCESSION   AC007157
VERSION     AC007157.6  GI:4646258
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 152937)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone hRPK.78_A_1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152937)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
            Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Depayre,E., Devon,K., Dewar,K.,
```

```
            Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
            Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
            Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
            Karatas,A., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P.,
            Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
            Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
            Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
            Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
            Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 152937)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
            Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Depayre,E., Devon,K., Dewar,K.,
            Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
            Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
            Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
            Karatas,A., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P.,
            Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
            Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
            Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
            Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
            Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 22, 1999 this sequence version replaced gi:4585945.
            All repeats were identified using RepeatMasker: Smit, A.F.A. &
            Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html.
FEATURES             Location/Qualifiers
     source          1. .152937
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="hRPK.78_A_1"
                     /clone_lib="RPCI-11 human BAC library"
     repeat_region   1. .1267
                     /rpt_family="L1M2"
     repeat_region   1268. .1564
                     /rpt_family="AluSp"
     repeat_region   1565. .2581
                     /rpt_family="L1M2"
     repeat_region   2582. .2612
                     /rpt_family="(CAAA)n"
     repeat_region   2613. .3942
                     /rpt_family="L1M2"
     repeat_region   3943. .4275
                     /rpt_family="L1PA7"
     repeat_region   4276. .4973
                     /rpt_family="L1M2"
     repeat_region   4974. .5234
                     /rpt_family="AluJo"
     repeat_region   5235. .5255
                     /rpt_family="(TAA)n"
     repeat_region   5256. .7139
                     /rpt_family="L1M2"
     repeat_region   7154. .7239
                     /rpt_family="GA-rich"
     repeat_region   7245. .7531
                     /rpt_family="AluJb"
     repeat_region   7544. .8081
                     /rpt_family="L1MA4A"
     repeat_region   8082. .8231
                     /rpt_family="L1MA4A"
     repeat_region   complement(9147. .9222)
                     /rpt_family="THE1B"
     repeat_region   9223. .9531
                     /rpt_family="AluSx"
     repeat_region   complement(9532. .9835)
                     /rpt_family="THE1B"
     repeat_region   11431. .11546
                     /rpt_family="L1PA4"
     repeat_region   11547. .11565
                     /rpt_family="AT_rich"
     repeat_region   complement(12122. .12180)
                     /rpt_family="L2"
     repeat_region   12199. .12507
                     /rpt_family="L1MC5"
     repeat_region   complement(12738. .13051)
                     /rpt_family="AluSx"
     repeat_region   13213. .14082
                     /rpt_family="L1PA10"
     repeat_region   14083. .14381
                     /rpt_family="AluSx"
     repeat_region   14382. .14488
                     /rpt_family="L1PA10"
     repeat_region   14500. .14552
                     /rpt_family="A-rich"
     repeat_region   16102. .16248
                     /rpt_family="MER63"
     repeat_region   16308. .16330
                     /rpt_family="(TC)n"
     repeat_region   16371. .16650
                     /rpt_family="L2"
     repeat_region   complement(16705. .16928)
                     /rpt_family="MIR"
     repeat_region   complement(17054. .17124)
                     /rpt_family="MER106"
     repeat_region   17314. .17622
                     /rpt_family="AluSx"
     repeat_region   complement(17827. .17905)
                     /rpt_family="L1MC/D"
     repeat_region   18738. .19043
                     /rpt_family="AluSx"
     repeat_region   complement(19130. .19310)
                     /rpt_family="MIR"
     repeat_region   20489. .20548
                     /rpt_family="MER44D"
     repeat_region   20549. .20860
                     /rpt_family="AluSx"
     repeat_region   20861. .20888
                     /rpt_family="MER44D"
     repeat_region   20907. .21125
                     /rpt_family="MER44D"
     repeat_region   complement(21164. .21266)
                     /rpt_family="MIR"
     repeat_region   21648. .21700
                     /rpt_family="AT_rich"
     repeat_region   21773. .22053
                     /rpt_family="AluSg"
     repeat_region   complement(23278. .23434)
                     /rpt_family="MIR"
     repeat_region   24095. .24381
                     /rpt_family="AluSg1"
     unsure          24989. .24994
                     /note="Single-stranded terminator coverage."
     repeat_region   25530. .25654
                     /rpt_family="L1PA8"
     repeat_region   25664. .25814
                     /rpt_family="(GGAA)n"
     repeat_region   26739. .27110
                     /rpt_family="L1MA8"
     repeat_region   27126. .27623
                     /rpt_family="HAL1"
     repeat_region   27620. .27760
                     /rpt_family="HAL1"
     repeat_region   27819. .28020
                     /rpt_family="HAL1"
```

```
     repeat_region   28209. .28338
                     /rpt_family="HAL1"
     repeat_region   28339. .28810
                     /rpt_family="L1MA5"
     repeat_region   complement(28902. .29104)
                     /rpt_family="L1MC5"
     repeat_region   29134. .29362
                     /rpt_family="MIR"
     repeat_region   29427. .29489
                     /rpt_family="AT_rich"
     repeat_region   29634. .30513
                     /rpt_family="L1PA7"
     repeat_region   30514. .30548
                     /rpt_family="AT_rich"
     repeat_region   30646. .30739
                     /rpt_family="L1MB2"
     repeat_region   30740. .30781
                     /rpt_family="(CA)n"
     repeat_region   30782. .31557
                     /rpt_family="L1MB2"
     repeat_region   31838. .31852
                     /rpt_family="AT_rich"
     repeat_region   complement(31853. .32504)
                     /rpt_family="L1MA3"
     repeat_region   32505. .32520
                     /rpt_family="AT_rich"
     repeat_region   complement(33807. .33935)
                     /rpt_family="L2"
     repeat_region   34307. .34437
                     /rpt_family="FLAM_C"
     repeat_region   35162. .35275
                     /rpt_family="(TA)n"
     repeat_region   35276. .35526
                     /rpt_family="LTR8"
     repeat_region   35527. .35548

Alignment Scores:
Pred. No.:              2.86e+04        Length:         152937
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC007157 (1-152937)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db      23465 TCTTTGATCACATGGAATATT 23445

RESULT 36
AC104664/c
LOCUS       AC104664               158145 bp    DNA     linear   PRI 21-FEB-2002
DEFINITION  Homo sapiens BAC clone RP11-724H6 from 4, complete sequence.
ACCESSION   AC104664 AC027502
VERSION     AC104664.2  GI:18042500
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 158145)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
   PUBMED   9847074
REFERENCE   2  (bases 1 to 158145)
  AUTHORS   Waligorski,J. and Haakenson,W.
  TITLE     The sequence of Homo sapiens BAC clone RP11-724H6
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 158145)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-DEC-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 158145)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 158145)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Jan 3, 2002 this sequence version replaced gi:17921279.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0724H06
                Drafting Center: WIBR
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO. For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            Actual start of this clone is at base position 1 of RP11-724H6;
            actual end is at base position 158145 of RP11-724H6.

            The sequence of AC027502 has been incorporated into AC104664.
FEATURES                Location/Qualifiers
     source             1. .158145
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="4"
                        /map="4"
                        /clone="RP11-724H6"
                        /clone_lib="RPCI-11"
     repeat_region      14. .68
                        /rpt_family="L2"
     repeat_region      288. .333
                        /rpt_family="A-rich"
```

November 1998

```
repeat_region   684. .734
                /rpt_family="(TTTA)n"
repeat_region   708. .1015
                /rpt_family="Alu"
repeat_region   1909. .1933
                /rpt_family="(CA)n"
repeat_region   2082. .2230
                /rpt_family="MIR"
repeat_region   2862. .3044
                /rpt_family="ERV1"
repeat_region   3013. .3058
                /rpt_family="(TTTG)n"
repeat_region   3045. .3342
                /rpt_family="Alu"
repeat_region   3343. .3602
                /rpt_family="ERV1"
repeat_region   3571. .3707
                /rpt_family="ERV1"
repeat_region   4255. .4435
                /rpt_family="MER53"
repeat_region   4548. .4578
                /rpt_family="AT_rich"
repeat_region   4963. .4985
                /rpt_family="AT_rich"
repeat_region   5433. .5795
                /rpt_family="L1"
repeat_region   5750. .5818
                /rpt_family="T-rich"
repeat_region   5796. .6165
                /rpt_family="L1"
repeat_region   6166. .6461
                /rpt_family="L1"
repeat_region   7517. .7545
                /rpt_family="AT_rich"
repeat_region   7761. .7844
                /rpt_family="MaLR"
repeat_region   7882. .8457
                /rpt_family="ERV1"
repeat_region   8768. .8890
                /rpt_family="Alu"
repeat_region   9555. .9631
                /rpt_family="(TA)n"
repeat_region   9669. .9704
                /rpt_family="AT_rich"
repeat_region   9760. .9885
                /rpt_family="MIR"
repeat_region   11686. .11840
                /rpt_family="MIR"
repeat_region   12301. .12523
                /rpt_family="L2"
repeat_region   12809. .13120
                /rpt_family="Alu"
repeat_region   13350. .13385
                /rpt_family="AT_rich"
repeat_region   13549. .13843
                /rpt_family="Alu"
repeat_region   13844. .14017
                /rpt_family="L2"
repeat_region   14016. .14043
                /rpt_family="(TTG)n"
repeat_region   14018. .14333
                /rpt_family="Alu"
repeat_region   14334. .14579
                /rpt_family="L2"
repeat_region   14667. .14830
                /rpt_family="MER1_type"
repeat_region   14689. .14723
                /rpt_family="AT_rich"
repeat_region   14974. .15397
                /rpt_family="MER1_type"
repeat_region   15393. .15418
                /rpt_family="AT_rich"
repeat_region   15434. .15619
                /rpt_family="MER1_type"
repeat_region   16183. .16376
                /rpt_family="L1"
repeat_region   16358. .16378
                /rpt_family="AT_rich"
repeat_region   16449. .16522
                /rpt_family="CR1"
repeat_region   18075. .18247
                /rpt_family="MER1_type"
repeat_region   19371. .19870
                /rpt_family="L1"
repeat_region   19871. .19967
                /rpt_family="Alu"
repeat_region   19971. .20877
                /rpt_family="L1"
repeat_region   20963. .21313
                /rpt_family="L1"
repeat_region   21324. .21405
                /rpt_family="MaLR"
repeat_region   21403. .21438
                /rpt_family="AT_rich"
repeat_region   21467. .21533
                /rpt_family="AT_rich"
repeat_region   21524. .21749
                /rpt_family="L1"
repeat_region   21752. .22021
                /rpt_family="L1"
repeat_region   22023. .22135
                /rpt_family="L1"
repeat_region   22136. .22253
                /rpt_family="L1"
repeat_region   22250. .22273
                /rpt_family="AT_rich"

Alignment Scores:
Pred. No.:              2.95e+04        Length:         158145
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC104664 (1-158145)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db      82050 TCATTAATAACTTGGAACATT 82030

RESULT 37
AL445242
LOCUS       AL445242              160177 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-321E11 on chromosome 13,
            complete sequence.
ACCESSION   AL445242
VERSION     AL445242.3  GI:11137901
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 160177)
  AUTHORS   Holt,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 10, 2000 this sequence version replaced gi:11062912.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
```

```
            was generated from part of bacterial clone contigs of human
            chromosome 13, constructed by the Sanger Centre Chromosome 13
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr13
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            -------------
            RP11-321E11 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .160177
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="RP11-321E11"
                     /clone_lib="RPCI-11.2"
     misc_feature    1
                     /note="Clone_left_end: RP11-321E11"
     misc_feature    160177
                     /note="Clone_right_end: RP11-321E11"
ORIGIN

Alignment Scores:
Pred. No.:              2.98e+04        Length:         160177
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AL445242 (1-160177)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      68428 TCATTGATAACTTGGAATATA 68448

RESULT 38
CR361565/c
LOCUS       CR361565               164475 bp    DNA     linear   VRT 07-SEP-2004
DEFINITION  Zebrafish DNA sequence from clone CH211-242L1 in linkage group 22,
            complete sequence.
ACCESSION   CR361565
VERSION     CR361565.8  GI:51950524
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 164475)
  AUTHORS   Barker,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 9, 2004 this sequence version replaced gi:51870436.
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-242L1
            is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .164475
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-242L1"
                     /clone_lib="CHORI-211"
ORIGIN

Alignment Scores:
Pred. No.:              3.06e+04        Length:         164475
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x CR361565 (1-164475)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      97645 AGTTTAATAACCTGGAACATA 97625

RESULT 39
AC016800/c
LOCUS       AC016800               164528 bp    DNA     linear   HTG 14-MAR-2000
DEFINITION  Homo sapiens clone RP11-2L21, *** SEQUENCING IN PROGRESS ***, 67
            unordered pieces.
ACCESSION   AC016800
VERSION     AC016800.2  GI:7229990
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 164528)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-2L21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 164528)
```

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (07-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Mar 12, 2000 this sequence version replaced gi:6532115.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-------------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
Center project name: L2661
Center clone name: 2_L_21
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 67 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	1	1095:	contig of 1095 bp in length
*	1096	1195:	gap of 100 bp
*	1196	2386:	contig of 1191 bp in length
*	2387	2486:	gap of 100 bp
*	2487	3802:	contig of 1316 bp in length
*	3803	3902:	gap of 100 bp
*	3903	4913:	contig of 1011 bp in length
*	4914	5013:	gap of 100 bp
*	5014	6071:	contig of 1058 bp in length
*	6072	6171:	gap of 100 bp
*	6172	7188:	contig of 1017 bp in length
*	7189	7288:	gap of 100 bp
*	7289	8489:	contig of 1201 bp in length
*	8490	8589:	gap of 100 bp
*	8590	9666:	contig of 1077 bp in length
*	9667	9766:	gap of 100 bp
*	9767	11233:	contig of 1467 bp in length
*	11234	11333:	gap of 100 bp
*	11334	12674:	contig of 1341 bp in length
*	12675	12774:	gap of 100 bp
*	12775	14216:	contig of 1442 bp in length
*	14217	14316:	gap of 100 bp
*	14317	15463:	contig of 1147 bp in length
*	15464	15563:	gap of 100 bp
*	15564	16945:	contig of 1382 bp in length
*	16946	17045:	gap of 100 bp
*	17046	18571:	contig of 1526 bp in length
*	18572	18671:	gap of 100 bp
*	18672	20246:	contig of 1575 bp in length
*	20247	20346:	gap of 100 bp
*	20347	22236:	contig of 1890 bp in length
*	22237	22336:	gap of 100 bp
*	22337	23343:	contig of 1007 bp in length
*	23344	23443:	gap of 100 bp
*	23444	24783:	contig of 1340 bp in length
*	24784	24883:	gap of 100 bp
*	24884	25955:	contig of 1072 bp in length
*	25956	26055:	gap of 100 bp
*	26056	27353:	contig of 1298 bp in length
*	27354	27453:	gap of 100 bp
*	27454	28779:	contig of 1326 bp in length
*	28780	28879:	gap of 100 bp
*	28880	30525:	contig of 1646 bp in length
*	30526	30625:	gap of 100 bp
*	30626	31806:	contig of 1181 bp in length
*	31807	31906:	gap of 100 bp
*	31907	33591:	contig of 1685 bp in length
*	33592	33691:	gap of 100 bp
*	33692	35096:	contig of 1405 bp in length
*	35097	35196:	gap of 100 bp
*	35197	37249:	contig of 2053 bp in length
*	37250	37349:	gap of 100 bp
*	37350	38428:	contig of 1079 bp in length
*	38429	38528:	gap of 100 bp
*	38529	40301:	contig of 1773 bp in length
*	40302	40401:	gap of 100 bp
*	40402	42088:	contig of 1687 bp in length
*	42089	42188:	gap of 100 bp
*	42189	43933:	contig of 1745 bp in length
*	43934	44033:	gap of 100 bp
*	44034	45578:	contig of 1545 bp in length
*	45579	45678:	gap of 100 bp
*	45679	47001:	contig of 1323 bp in length
*	47002	47101:	gap of 100 bp
*	47102	48554:	contig of 1453 bp in length
*	48555	48654:	gap of 100 bp
*	48655	50560:	contig of 1906 bp in length
*	50561	50660:	gap of 100 bp
*	50661	52356:	contig of 1696 bp in length
*	52357	52456:	gap of 100 bp
*	52457	54079:	contig of 1623 bp in length
*	54080	54179:	gap of 100 bp
*	54180	55927:	contig of 1748 bp in length
*	55928	56027:	gap of 100 bp
*	56028	58921:	contig of 2894 bp in length
*	58922	59021:	gap of 100 bp
*	59022	61172:	contig of 2151 bp in length
*	61173	61272:	gap of 100 bp
*	61273	63203:	contig of 1931 bp in length
*	63204	63303:	gap of 100 bp
*	63304	64868:	contig of 1565 bp in length
*	64869	64968:	gap of 100 bp
*	64969	67516:	contig of 2548 bp in length
*	67517	67616:	gap of 100 bp
*	67617	69749:	contig of 2133 bp in length
*	69750	69849:	gap of 100 bp
*	69850	72841:	contig of 2992 bp in length
*	72842	72941:	gap of 100 bp
*	72942	74456:	contig of 1515 bp in length
*	74457	74556:	gap of 100 bp
*	74557	78257:	contig of 3701 bp in length
*	78258	78357:	gap of 100 bp
*	78358	81266:	contig of 2909 bp in length
*	81267	81366:	gap of 100 bp
*	81367	84282:	contig of 2916 bp in length
*	84283	84382:	gap of 100 bp
*	84383	86181:	contig of 1799 bp in length
*	86182	86281:	gap of 100 bp
*	86282	89197:	contig of 2916 bp in length
*	89198	89297:	gap of 100 bp
*	89298	91641:	contig of 2344 bp in length
*	91642	91741:	gap of 100 bp
*	91742	94732:	contig of 2991 bp in length
*	94733	94832:	gap of 100 bp
*	94833	99130:	contig of 4298 bp in length
*	99131	99230:	gap of 100 bp
*	99231	102224:	contig of 2994 bp in length
*	102225	102324:	gap of 100 bp

```
            *   102325   105640: contig of 3316 bp in length
            *   105641   105740: gap of 100 bp
            *   105741   109057: contig of 3317 bp in length
            *   109058   109157: gap of 100 bp
            *   109158   113699: contig of 4542 bp in length
            *   113700   113799: gap of 100 bp
            *   113800   117438: contig of 3639 bp in length
            *   117439   117538: gap of 100 bp
            *   117539   120283: contig of 2745 bp in length
            *   120284   120383: gap of 100 bp
            *   120384   124529: contig of 4146 bp in length
            *   124530   124629: gap of 100 bp
            *   124630   128684: contig of 4055 bp in length
            *   128685   128784: gap of 100 bp
            *   128785   132342: contig of 3558 bp in length
            *   132343   132442: gap of 100 bp
            *   132443   138252: contig of 5810 bp in length
            *   138253   138352: gap of 100 bp
            *   138353   144132: contig of 5780 bp in length
            *   144133   144232: gap of 100 bp
            *   144233   150567: contig of 6335 bp in length
            *   150568   150667: gap of 100 bp
            *   150668   156867: contig of 6200 bp in length
            *   156868   156967: gap of 100 bp
            *   156968   164528: contig of 7561 bp in length.
FEATURES             Location/Qualifiers
     source          1. .164528
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-2L21"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1095
                     /note="assembly_fragment"
     gap             1096. .1195
                     /estimated_length=100
     misc_feature    1196. .2386
                     /note="assembly_fragment"

Alignment Scores:
Pred. No.:              3.06e+04        Length:         164528
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC016800 (1-164528)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db      57637 TCCCTTATCACATGGAATATA 57617

RESULT 40
AC121567/c
LOCUS       AC121567               165586 bp    DNA     linear   ROD 08-NOV-2003
DEFINITION  Mus musculus BAC clone RP23-225D14 from 14, complete sequence.
ACCESSION   AC121567
VERSION     AC121567.3  GI:29294384
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 165586)
  AUTHORS   Wang,C., Elliott,G., Spalding,L., Mangiapanello,L. and Kozlowicz,A.
  TITLE     The sequence of Mus musculus BAC clone RP23-225D14
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 165586)
  AUTHORS   Wilson,R.
  TITLE     Sequencing of Mus musculus
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 165586)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 165586)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   5  (bases 1 to 165586)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   6  (bases 1 to 165586)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Mar 27, 2003 this sequence version replaced gi:22475483.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: M_BA0225D14
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
            and Minako Tateno in the laboratory of Pieter de Jong
            (http://www.chori.org) from female C57BL/6J mouse kidney and/or
            brain genomic DNA.  The clone and detailed information can be
            obtained from Research Genetics, Inc. (http://www.resgen.com) or
            Pieter de Jong and coworkers at http://www.chori.org

            NEIGHBORING SEQUENCE INFORMATION:
            This sequence is the entire insert of the clone.
FEATURES             Location/Qualifiers
     source          1. .165586
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="14"
                     /map="14"
                     /clone="RP23-225D14"
                     /clone_lib="RPCI-23"
     unsure          296. .379
                     /note="Sequence derived from one plasmid subclone."
     repeat_region   1327. .1509
                     /rpt_family="B2"
     tRNA            complement(1433. .1504)
```

```
                     /product="tRNA-Ser"
                     /note="Likely pseudogene (HMM Sc=33.94 / Sec struct
                     Sc=-11.27)"
     repeat_region   2949. .3008
                     /rpt_family="ERV1"
     repeat_region   3054. .3245
                     /rpt_family="MaLR"
     repeat_region   3261. .4254
                     /rpt_family="ERVK"
     repeat_region   4255. .4365
                     /rpt_family="L1"
     repeat_region   4367. .5308
                     /rpt_family="L1"
     misc_feature    5612. .5667
                     /note="Sequence derived from PCR product of project DNA."
     repeat_region   5806. .5895
                     /rpt_family="ERVK"
     unsure          5814. .5895
                     /note="Unresolved simple sequence repeat."
     unsure          5896. .5938
                     /note="Unresolved simple sequence repeat."
     repeat_region   6042. .6320
                     /rpt_family="B4"
     repeat_region   6434. .6544
                     /rpt_family="L1"
     repeat_region   6637. .6983
                     /rpt_family="MaLR"
     repeat_region   7128. .7793
                     /rpt_family="L1"
     repeat_region   8124. .8477
                     /rpt_family="L1"
     repeat_region   8465. .8600
                     /rpt_family="L1"
     repeat_region   8764. .8937
                     /rpt_family="MaLR"
     repeat_region   9154. .9302
                     /rpt_family="Alu"
     unsure          9989. .10519
                     /note="Unresolved simple sequence repeat."
     repeat_region   10580. .11083
                     /rpt_family="L1"
     repeat_region   11150. .11332
                     /rpt_family="B4"
     repeat_region   11466. .11720
                     /rpt_family="L1"
     repeat_region   11849. .11985
                     /rpt_family="L1"
     repeat_region   13685. .13824
                     /rpt_family="Alu"
     repeat_region   14208. .14582
                     /rpt_family="MaLR"
     repeat_region   15212. .15293
                     /rpt_family="MER1_type"
     repeat_region   15729. .15798
                     /rpt_family="L1"
     repeat_region   15810. .16213
                     /rpt_family="ERVK"
     repeat_region   16217. .16462
                     /rpt_family="L1"
     repeat_region   16659. .16720
                     /rpt_family="L1"
     repeat_region   17355. .17802
                     /rpt_family="L1"
     repeat_region   17891. .18326
                     /rpt_family="L1"
     repeat_region   18378. .18754
                     /rpt_family="MaLR"
     repeat_region   18786. .19395
                     /rpt_family="L1"
     repeat_region   19416. .19576
                     /rpt_family="L1"
     repeat_region   19649. .20154
                     /rpt_family="RMER19B"
     repeat_region   20160. .20198
                     /rpt_family="L1"
     repeat_region   20190. .20562
                     /rpt_family="L1"
     repeat_region   20567. .20704
                     /rpt_family="MaLR"
     repeat_region   20701. .20988
                     /rpt_family="MaLR"
     repeat_region   21033. .21163
                     /rpt_family="L1"
     repeat_region   21398. .21858
                     /rpt_family="MaLR"
     repeat_region   21644. .21924
                     /rpt_family="ERVL"
     repeat_region   22319. .22411
                     /rpt_family="L1"
     repeat_region   22392. .22453
                     /rpt_family="L1"
     repeat_region   22451. .23065
                     /rpt_family="L1"
     repeat_region   23220. .23798
                     /rpt_family="L1"
     repeat_region   23900. .24046
                     /rpt_family="Alu"
     repeat_region   24053. .24471
                     /rpt_family="L1"
     repeat_region   24473. .25041
                     /rpt_family="ERV1"
     repeat_region   28893. .29041
                     /rpt_family="Alu"
     repeat_region   29109. .29248
                     /rpt_family="Alu"
     repeat_region   29249. .30612
                     /rpt_family="L1"
     repeat_region   30611. .31161
                     /rpt_family="L1"

Alignment Scores:
Pred. No.:              3.07e+04        Length:         165586
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC121567 (1-165586)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      18243 TCACTTATAACCTGGAATATA 18223

RESULT 41
CR925830
LOCUS       CR925830               167847 bp    DNA     linear   PRI 08-JAN-2005
DEFINITION  Human DNA sequence from clone RP11-195J19 on chromosome 6, complete
            sequence.
ACCESSION   CR925830
VERSION     CR925830.3  GI:56798018
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 167847)
  AUTHORS   Palmer,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 24, 2004 this sequence version replaced gi:56542150.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
```

```
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-195J19 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .167847
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-195J19"
                     /clone_lib="RPCI-11.1"
ORIGIN

Alignment Scores:
Pred. No.:                3.11e+04      Length:          167847
Score:                    38.00         Matches:         7
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       8             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x CR925830 (1-167847)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      35630 TCTCTCATTACCTGGAATATT 35650

RESULT 42
AC151746/c
LOCUS       AC151746               168291 bp    DNA     linear   ROD 08-MAR-2005
DEFINITION  Mus musculus BAC clone RP23-19M11 from 9, complete sequence.
ACCESSION   AC151746
VERSION     AC151746.3  GI:58531643
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 168291)
  AUTHORS   Nguyen,C., Cotton,M. and Kozlowicz,A.
  TITLE     The sequence of Mus musculus BAC clone RP23-19M11
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 168291)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 168291)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 168291)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2005) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   5  (bases 1 to 168291)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-MAR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Feb 3, 2005 this sequence version replaced gi:58038189.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: M_BA0019M11
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
            and Minako Tateno in the laboratory of Pieter de Jong
            (http://www.chori.org) from female C57BL/6J mouse kidney and/or
            brain genomic DNA.  The clone and detailed information can be
            obtained from Research Genetics, Inc. (http://www.resgen.com) or
            Pieter de Jong and coworkers at http://www.chori.org

            NEIGHBORING SEQUENCE INFORMATION:
            This sequence is the entire insert of the clone.
FEATURES             Location/Qualifiers
     source          1. .168291
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /map="9"
                     /clone="RP23-19M11"
                     /clone_lib="RPCI-23"
     unsure          6140. .6216
                     /note="Unresolved simple sequence repeat."
     unsure          59914. .60150
                     /note="Sequence derived from one plasmid subclone."
     unsure          95077. .95247
                     /note="Unresolved simple sequence repeat."
     unsure          134490. .134498
                     /note="Sequence derived from one plasmid subclone."
ORIGIN
```

```
Alignment Scores:
Pred. No.:              3.12e+04        Length:         168291
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC151746 (1-168291)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     166758 AGTTTGATTACCTGGAACATC 166738

RESULT 43
AL442064/c
LOCUS       AL442064               169905 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-272D1 on chromosome 9 Contains
            the 3' end of the EDG2 gene for endothelial differentiation,
            lysophosphatidic acid G-protein-coupled receptor, 2, complete
            sequence.
ACCESSION   AL442064
VERSION     AL442064.10  GI:11544548
KEYWORDS    HTG; EDG2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 169905)
  AUTHORS   Garner,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 4, 2000 this sequence version replaced gi:11414642.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            RP11-272D1 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .169905
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /clone="RP11-272D1"
                     /clone_lib="RPCI-11.1"
     misc_feature    1
                     /note="Clone_left_end: RP11-272D1"
     polyA_site      complement(22599)
     polyA_signal    complement(22612. .22617)
     polyA_site      complement(23114)
     polyA_signal    complement(23125. .23130)
     gene            join(complement(148728. .148803),
                     complement(122818. .122895),complement(121410. .121557),
                     complement(90758. .91505),
                     complement(AL157881.14:106498. .108442))
                     /gene="EDG2"
                     /locus_tag="RP11-104M22.2-002"
     mRNA            join(complement(148728. .148803),
                     complement(122818. .122895),complement(121410. .121557),
                     complement(90758. .91505),
                     complement(AL157881.14:106498. .108442))
                     /gene="EDG2"
                     /locus_tag="RP11-104M22.2-002"
                     /product="endothelial differentiation, lysophosphatidic
                     acid G-protein-coupled receptor, 2"
                     /note="match: ESTs: AI687023.1 AW150071.1 AW269335.1
                     AW452359.1 BG701440.1
                     match: cDNAs: BC030615.1"
     gene            join(complement(AC007157.6:85725. .85927),
                     complement(122818. .122895),complement(121410. .121557),
                     complement(90758. .91505),
                     complement(AL157881.14:105983. .108442))
                     /gene="EDG2"
                     /locus_tag="RP11-104M22.2-001"
     mRNA            join(complement(AC007157.6:85725. .85927),
                     complement(122818. .122895),complement(121410. .121557),
                     complement(90758. .91505),
                     complement(AL157881.14:105983. .108442))
                     /gene="EDG2"
                     /locus_tag="RP11-104M22.2-001"
                     /product="endothelial differentiation, lysophosphatidic
                     acid G-protein-coupled receptor, 2"
                     /note="match: ESTs: AI150725.1 AU120057.1 BE882064.1
                     BF055366.1 BF214735.1 BF241764.1 BF966982.1 BG119082.1
                     BG611648.1 BI461641.1 BI493987.1 BI820204.1 BM807438.1
                     BM972828.1 BM988491.1 BM988931.1 BM993722.1 BQ005319.1
                     BQ435615.1 BU171826.1 BU521528.1 BU838473.1 CA397755.1
                     match: cDNAs: AF014418.1 AF090347.1 AJ249843.1 AJ249844.1
                     BC025425.1 BC036034.1 U18405.1 U48235.1 U48236.1 U70622.1
                     U78192.1 U80811.1 Y09479.1"
     gene            join(complement(AC007157.6:85725. .85969),
                     complement(121410. .121557),complement(90758. .91505),
                     complement(AL157881.14:106897. .108442))
                     /gene="EDG2"
                     /locus_tag="RP11-104M22.2-003"
     mRNA            join(complement(AC007157.6:85725. .85969),
                     complement(121410. .121557),complement(90758. .91505),
                     complement(AL157881.14:106897. .108442))
                     /gene="EDG2"
                     /locus_tag="RP11-104M22.2-003"
                     /product="endothelial differentiation, lysophosphatidic
                     acid G-protein-coupled receptor, 2"
                     /note="match: cDNAs: AK022808.1"
     CDS             join(complement(121410. .121454),complement(90758. .91505),
                     complement(AL157881.14:108141. .108442))
                     /gene="EDG2"
                     /locus_tag="RP11-104M22.2-002"
                     /standard_name="OTTHUMP00000021904"
                     /codon_start=1
                     /product="endothelial differentiation, lysophosphatidic
                     acid G-protein-coupled receptor, 2"
                     /protein_id="CAI13010.1"
                     /db_xref="GI:55959843"
                     /db_xref="InterPro:IPR000276"
                     /db_xref="InterPro:IPR002277"
                     /db_xref="InterPro:IPR004065"
                     /translation="MAAISTSIPVISQPQFTAMNEPQCFYNESIAFFYNRSGKHLATE
                     WNTVSKLVMGLGITVCIFIMLANLLVMVAIYVNRRFHFPIYYLMANLAAADFFAGLAY
                     FYLMFNTGPNTRRLTVSTWLLRQGLIDTSLTASVANLLAIAIERHITVFRMQLHTRMS
                     NRRVVVVIVVIWTMAIVMGAIPSVGWNCICDIENCSNMAPLYSDSYLVFWAIFNLVTF
```

```
                     VVMVVLYAHIFGYVRQRTMRMSRHSSGPRRNRDTMMSLLKTVVIVLGAFIICWTPGLV
                     LLLLDVCCPQCDVLAYEKFFLLLAEFNSAMNPIIYSYRDKEMSATFRQILCCQRSENP
                     TGPTEGSDRSASSLNHTILAGVHSNDHSVV"
     gene            join(complement(AC007157.6:86051. .86342),
                     complement(122818. .122895),complement(121410. .121557),
                     complement(<91005. .91505))
                     /gene="EDG2"
                     /locus_tag="RP11-104M22.2-004"
     mRNA            join(complement(AC007157.6:86051. .86342),
                     complement(122818. .122895),complement(121410. .121557),
                     complement(91005. .91505))
                     /gene="EDG2"
                     /locus_tag="RP11-104M22.2-004"
                     /product="endothelial differentiation, lysophosphatidic
                     acid G-protein-coupled receptor, 2"
                     /note="match: ESTs: BG699220.1 BI834404.1 BM805943.1"
     CDS             complement(join(<91005. .91505,121410. .121454))
                     /gene="EDG2"
                     /locus_tag="RP11-104M22.2-003"
                     /standard_name="OTTHUMP00000021906"
                     /codon_start=1
                     /product="endothelial differentiation, lysophosphatidic
                     acid G-protein-coupled receptor, 2"
                     /protein_id="CAI13009.1"
                     /db_xref="GI:55959842"
                     /db_xref="GOA:Q5T261"
                     /db_xref="InterPro:IPR000276"
                     /db_xref="InterPro:IPR002277"
                     /db_xref="InterPro:IPR004065"
                     /db_xref="UniProt/TrEMBL:Q5T261"
                     /translation="MAAISTSIPVISQPQFTAMNEPQCFYNESIAFFYNRSGKHLATE
                     WNTVSKLVMGLGITVCIFIMLANLLVMVAIYVNRRFHFPIYYLMANLAAADFFAGLAY
                     FYLMFNTGPNTRRLTVSTWLLRQGLIDTSLTASVANLLAIAIERHITVFRMQLHTRMS
                     NRRVVVVIVVIWTMAIVMGAIP"
     misc_feature    113875. .113911
                     /note="Single clone region. Assembly confirmed by
                     restriction digest data."
     misc_feature    169905
                     /note="Clone_right_end: RP11-272D1"
ORIGIN

Alignment Scores:
Pred. No.:              3.15e+04        Length:         169905
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AL442064 (1-169905)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     124918 TCTTTGATCACATGGAATATT 124898

RESULT 44
AC140199/c
LOCUS       AC140199             170896 bp    DNA     linear   ROD 15-MAY-2004
DEFINITION  Mus musculus BAC clone RP23-480D7 from chromosome 9, complete
            sequence.
ACCESSION   AC140199
VERSION     AC140199.2  GI:46276004
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 170896)
  AUTHORS   Hodges,J., Bielicki,L. and Haakenson,W.
  TITLE     The sequence of Mus musculus BAC clone RP23-480D7
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 170896)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 170896)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-APR-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 170896)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On Apr 8, 2004 this sequence version replaced gi:28475434.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: M_BA0480D07
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone, fosmid clone or direct clone walk sequence.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to obtain the consensus sequence; and
            the assembly was confirmed by  restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
            and Minako Tateno in the laboratory of Pieter de Jong
            (http://www.chori.org) from female C57BL/6J mouse kidney and/or
            brain genomic DNA.  The clone and detailed information can be
            obtained from Research Genetics, Inc. (http://www.resgen.com) or
            Pieter de Jong and coworkers at http://www.chori.org

            NEIGHBORING SEQUENCE INFORMATION:
            This sequence is the entire insert of the clone.
FEATURES             Location/Qualifiers
     source          1. .170896
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /map="9"
                     /clone="RP23-480D7"
                     /clone_lib="RPCI-23"
     repeat_region   870. .1117
                     /rpt_family="L1"
     repeat_region   1316. .1817
                     /rpt_family="L1"
     repeat_region   1817. .2205
                     /rpt_family="L1"
     repeat_region   4149. .4349
                     /rpt_family="B2"
     repeat_region   6631. .7353
                     /rpt_family="L1"
     repeat_region   7352. .7531
                     /rpt_family="L1"
```

```
     repeat_region   7533. .8289
                     /rpt_family="L1"
     repeat_region   8640. .8691
                     /rpt_family="L1"
     repeat_region   8793. .9419
                     /rpt_family="L1"
     repeat_region   9429. .9575
                     /rpt_family="B2"
     repeat_region   9768. .10617
                     /rpt_family="L1"
     repeat_region   10640. .10856
                     /rpt_family="B2"
     repeat_region   11071. .11141
                     /rpt_family="ID"
     repeat_region   11189. .11863
                     /rpt_family="L1"
     repeat_region   12044. .12323
                     /rpt_family="L1"
     repeat_region   12406. .12470
                     /rpt_family="B4"
     repeat_region   12677. .13004
                     /rpt_family="L1"
     repeat_region   13251. .13778
                     /rpt_family="L1"
     repeat_region   13989. .14034
                     /rpt_family="B4"
     repeat_region   14534. .14604
                     /rpt_family="ID"
     repeat_region   17169. .17382
                     /rpt_family="B2"
     repeat_region   17669. .18032
                     /rpt_family="RMER1B"
     repeat_region   18148. .18310
                     /rpt_family="RMER1B"
     repeat_region   18328. .18488
                     /rpt_family="RMER1B"
     repeat_region   19242. .19431
                     /rpt_family="L1"
     repeat_region   20764. .20943
                     /rpt_family="L1"
     repeat_region   20932. .21278
                     /rpt_family="L1"
     repeat_region   21519. .21697
                     /rpt_family="L1"
     repeat_region   22058. .23079
                     /rpt_family="L1"
     repeat_region   23080. .23302
                     /rpt_family="MER1_type"
     repeat_region   23303. .23690
                     /rpt_family="L1"
     repeat_region   24031. .25221
                     /rpt_family="L1"
     repeat_region   25402. .25708
                     /rpt_family="L1"
     repeat_region   25806. .26004
                     /rpt_family="L1"
     repeat_region   26515. .26948
                     /rpt_family="L1"
     repeat_region   27425. .27632
                     /rpt_family="L1"
     repeat_region   28292. .28491
                     /rpt_family="MaLR"
     repeat_region   28486. .28621
                     /rpt_family="Alu"
     repeat_region   28622. .28702
                     /rpt_family="MaLR"
     repeat_region   29638. .30400
                     /rpt_family="ERVK"
     misc_feature    29887. .29888
                     /note="Sequence derived from PCR product of genomic DNA."
     repeat_region   30668. .31049
                     /rpt_family="L1"
     repeat_region   31107. .31312
                     /rpt_family="L1"
     repeat_region   31356. .31519
                     /rpt_family="ERVK"
     repeat_region   31530. .31602
                     /rpt_family="ERVK"
     repeat_region   31897. .32071
                     /rpt_family="B2"
     repeat_region   32394. .32537
                     /rpt_family="ERVK"
     repeat_region   32998. .33876
                     /rpt_family="ERVK"
     repeat_region   34183. .34239
                     /rpt_family="ERVK"
     repeat_region   34278. .34505
                     /rpt_family="ERVK"
     repeat_region   34584. .34748
                     /rpt_family="ERVK"
     repeat_region   34759. .34831
                     /rpt_family="ERVK"
     repeat_region   34834. .35021
                     /rpt_family="L1"
     repeat_region   35961. .36136
                     /rpt_family="L1"
     repeat_region   38375. .38526
                     /rpt_family="L1"
     repeat_region   38536. .38983
                     /rpt_family="L1"
     unsure          41067. .41211
                     /note="Sequence derived from one plasmid subclone."
     repeat_region   43352. .43928
                     /rpt_family="ERV1"
     repeat_region   43929. .45060
                     /rpt_family="ERV1"
     repeat_region   45465. .45621
                     /rpt_family="ERV1"
     repeat_region   51539. .52115
                     /rpt_family="ERV1"
     repeat_region   53981. .54397

Alignment Scores:
Pred. No.:              3.16e+04        Length:         170896
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC140199 (1-170896)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      18735 AGTTTGATTACCTGGAACATC 18715

RESULT 45
CR942205
LOCUS       CR942205              176819 bp    DNA     linear   PRI 29-APR-2005
DEFINITION  Human DNA sequence from clone RP11-948M1 on chromosome 6, complete
            sequence.
ACCESSION   CR942205
VERSION     CR942205.4  GI:62989869
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 176819)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:62551221.
```

---------- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6
RP11-948M1 is from the library RPCI-11.4 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES Location/Qualifiers
source 1. .176819
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-948M1"
/clone_lib="RPCI-11.4"

ORIGIN

Alignment Scores:

Pred. No.:	3.26e+04	Length:	176819
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x CR942205 (1-176819)

Qy 1 SerLeuIleThrTrpAsnIle 7
||||||||||||||||||||
Db 29609 TCTCTCATTACCTGGAATATT 29629

RESULT 46
AC164269/c
LOCUS AC164269 178252 bp DNA linear HTG 01-JUL-2005
DEFINITION Bos taurus clone CH240-146A5, WORKING DRAFT SEQUENCE, 11 unordered pieces.
ACCESSION AC164269
VERSION AC164269.2 GI:68301121
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 178252)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 178252)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 178252)
AUTHORS Cow Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jun 29, 2005 this sequence version replaced gi:67972703.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

```
table.
-------------- Genome Center
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
    Center project name: FIIJ
    Center clone name: CH240-146A5
-------------- Summary Statistics
    Assembly program: Atlas 3.0;
    Consensus quality: 174486 bases at least Q40
    Consensus quality: 175765 bases at least Q30
    Consensus quality: 176748 bases at least Q20
    Estimated insert size: 178361; sum-of-contigs estimation
    Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
--------------
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1    36455: contig of 36455 bp in length
*    36456    36505: gap of 50 bp
*    36506    42142: contig of 5637 bp in length
*    42143    42192: gap of 50 bp
*    42193    47072: contig of 4880 bp in length
*    47073    47122: gap of 50 bp
*    47123    87942: contig of 40820 bp in length
*    87943    87992: gap of 50 bp
*    87993    94381: contig of 6389 bp in length
*    94382    94431: gap of 50 bp
*    94432   123017: contig of 28586 bp in length
*   123018   123067: gap of 50 bp
*   123068   146637: contig of 23570 bp in length
*   146638   146687: gap of 50 bp
*   146688   172253: contig of 25566 bp in length
*   172254   172303: gap of 50 bp
*   172304   175780: contig of 3477 bp in length
*   175781   175880: gap of unknown length
*   175881   176918: contig of 1038 bp in length
*   176919   177018: gap of unknown length
*   177019   178252: contig of 1234 bp in length.
FEATURES             Location/Qualifiers
     source          1. .178252
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9913"
                     /clone="CH240-146A5"
     gap             36456. .36505
                     /estimated_length=50
     gap             42143. .42192
                     /estimated_length=50
     gap             47073. .47122
                     /estimated_length=50
     gap             87943. .87992
                     /estimated_length=50
     gap             94382. .94431
                     /estimated_length=50
     gap             123018. .123067
                     /estimated_length=50
     gap             146638. .146687
                     /estimated_length=50
     gap             172254. .172303
                     /estimated_length=50
     gap             175781. .175880
                     /estimated_length=unknown
     gap             176919. .177018
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:                3.29e+04        Length:          178252
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       14              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC164269 (1-178252)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      89657 TCACTCATTACATGGAATATA 89637

RESULT 47
BS000021
LOCUS       BS000021              183533 bp    DNA     linear   PRI 12-JUN-2004
DEFINITION  Pan troglodytes chromosome 22 clone:PTB-141K15, map 22, complete
            sequences.
ACCESSION   BS000021 BA000046
VERSION     BS000021.1  GI:37537288
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   The International Chimpanzee Chromosome 22 Consortium.
  TITLE     DNA sequence and comparative analysis of chimpanzee chromosome 22
  JOURNAL   Nature 429, 382-388 (2004)
REFERENCE   2  (bases 1 to 183533)
  AUTHORS   Hattori,M., Toyoda,A., Watanabe,H., Taylor,T.D., Kuroki,Y.,
            Fujiyama,A. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
            *Chinese National Human Genome Center at Shanghai, Shanghai, China;
            *GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
            of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research
            Center, Daejeon, Korea;
            *Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
            *National Institute of Genetics, Mishima, Japan;
            *National Yang Ming University Genome Research Center, Taipei,
            Taiwan;
            *RIKEN Genomic Sciences Center, Yokohama, Japan.
            -------------- Genome Center
            Center: RIKEN Genomic Sciences Center
            Center code: RIKEN
            Web site: http://hgp.gsc.riken.go.jp/
            Contact: hattori@gsc.riken.go.jp
            -------------- Project Information
            Center project name:The Chimpanzee Chromosome 22 Sequencing Project
            Center clone name: PTB-141K15
            -------------- Summary Statistics
            Sequencing vector: pUC18,pUC13,pTZ19R; 100% of reads Chemistry:
            Dye-terminator Big Dye and ET; 100% of reads Assembly program:
            Phrap; version 0.990329
            Consensus quality:   183,519 bases at least Q40
            Consensus quality:         2 bases at least Q30
            Consensus quality:         0 bases at least Q20
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30);
```

an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at one plasmid
subclone or more than one M13 subclone;
and the assembly was confirmed by restriction digest.

Source information:
The RPCI-43 chimpanzee BAC library was prepared from DNA isolated from the blood of a single male chimpanzee using published protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the chimpanzee ('Clint') was obtained from the Yerkes Primate Center in Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu, Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library characteristics are described at
http://www.chori.org/bacpac/mchimp43.htm.
The clone may be obtained from Pieter J. de Jong and coworkers (http://www.chori.org/bacpac).
VECTOR: pBACe3.6
The CHORI-251 chimpanzee BAC library was prepared from DNA isolated from the blood of a single male chimpanzee using published protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the chimpanzee ('Clint') was obtained from the Yerkes Primate Center in Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu, Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library characteristics are described at
http://www.chori.org/bacpac/chimpanzee251.htm.
The clone may be obtained from Pieter J. de Jong and coworkers (http://www.chori.org/bacpac).
VECTOR: pTARBAC2.1
The PTB1 chimpanzee BAC library was prepared from DNA isolated from cultured cells established from the blood of a single male chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers (http://www.gsc.riken.go.jp).
VECTOR: pKS145
The PTF22 chimpanzee Fosmid library was prepared from DNA isolated from cultured cells established from the blood of a single male chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers (http://www.gsc.riken.go.jp).
VECTOR: pKS143

Sequence Quality Assessment:
This entry has been annotated with sequence estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Neighboring clones: PTB-084P08(left) and PTB-131J13(right).

FEATURES Location/Qualifiers
source 1. .183533
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="PTB-141K15"
/clone_lib="PTB1 chimpanzee BAC"

ORIGIN

Alignment Scores:

Pred. No.:	3.37e+04	Length:	183533
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BS000021 (1-183533)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     126704 AGTTTGATAACATGGAACATT 126724
```

RESULT 48
AC112957/c
LOCUS AC112957 186051 bp DNA linear ROD 10-MAR-2005
DEFINITION Mus musculus chromosome 15, clone RP23-418C6, complete sequence.
ACCESSION AC112957
VERSION AC112957.14 GI:60678471
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 186051)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus chromosome 15, clone RP23-418C6
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 186051)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 186051)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 186051)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,

```
          DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
          Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
          Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
          Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
          Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
          Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
          MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
          McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
          Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
          O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
          Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
          Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
          Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
          Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
          Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
          Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
          Zimmer,A. and Zody,M.
  TITLE   Direct Submission
  JOURNAL Submitted (10-MAR-2005) Broad Institute of MIT and Harvard, 320
          Charles Street, Cambridge, MA 02141, USA
COMMENT   On Mar 10, 2005 this sequence version replaced gi:59958519.
          All repeats were identified using RepeatMasker:
          Smit, A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          -------------- Genome Center
              Center: Broad Institute of MIT and Harvard
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@broad.mit.edu
          -------------- Project Information
              Center project name: L20693
              Center clone name: 418_C_6
          --------------.
FEATURES             Location/Qualifiers
     source          1. .186051
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="15"
                     /map="15"
                     /clone="RP23-418C6"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     repeat_region   1194. .1406
                     /rpt_family="ORR1C"
     repeat_region   3463. .3483
                     /rpt_family="AT_rich"
     repeat_region   3595. .3775
                     /rpt_family="MIR3"
     repeat_region   4148. .4454
                     /rpt_family="MER33"
     repeat_region   4732. .4784
                     /rpt_family="(TTTTA)n"
     repeat_region   complement(4900. .5037)
                     /rpt_family="B1F"
     repeat_region   7252. .7455
                     /rpt_family="URR1A"
     repeat_region   8561. .8598
                     /rpt_family="C-rich"
     repeat_region   8673. .8748
                     /rpt_family="PB1D9"
     repeat_region   8876. .8979
                     /rpt_family="B1_MM"
     repeat_region   9019. .9196
                     /rpt_family="B3"
     repeat_region   9809. .10168
                     /rpt_family="MTE"
     repeat_region   complement(10609. .10801)
                     /rpt_family="B3"
     repeat_region   10991. .11056
                     /rpt_family="MIR"
     repeat_region   11113. .11149
                     /rpt_family="AT_rich"
     repeat_region   12773. .12820
                     /rpt_family="(TG)n"
     repeat_region   complement(13362. .13379)
                     /rpt_family="B4"
     repeat_region   complement(13380. .13449)
                     /rpt_family="ID5"
     repeat_region   15370. .15428
                     /rpt_family="MLT1D"
     repeat_region   15459. .15506
                     /rpt_family="T-rich"
     repeat_region   16035. .16359
                     /rpt_family="MTE"
     repeat_region   17209. .17259
                     /rpt_family="(TA)n"
     repeat_region   17406. .17485
                     /rpt_family="MIR"
     repeat_region   complement(17835. .17974)
                     /rpt_family="MER20"
     repeat_region   18005. .18138
                     /rpt_family="BC1_MM"
     repeat_region   18140. .18169
                     /rpt_family="(CA)n"
     repeat_region   19412. .19499
                     /rpt_family="(GA)n"
     repeat_region   19637. .19669
                     /rpt_family="AT_rich"
     repeat_region   20536. .20637
                     /rpt_family="MTC"
     repeat_region   20861. .21172
                     /rpt_family="MTC"
     repeat_region   complement(21192. .21254)
                     /rpt_family="MIR"
     repeat_region   complement(21885. .22213)
                     /rpt_family="ORR1A2"
     repeat_region   complement(23391. .23562)
                     /rpt_family="L1_MM"
     repeat_region   23595. .23614
                     /rpt_family="(TA)n"
     repeat_region   24336. .24413
                     /rpt_family="PB1D7"
     repeat_region   24414. .24540
                     /rpt_family="(CAGAGA)n"
     repeat_region   24541. .24584
                     /rpt_family="PB1D7"
     repeat_region   24610. .24651
                     /rpt_family="(TG)n"
     repeat_region   complement(24700. .24846)
                     /rpt_family="MLT2D"
     repeat_region   25242. .25266
                     /rpt_family="AT_rich"
     repeat_region   26331. .26385
                     /rpt_family="(T)n"
     repeat_region   26555. .26585
                     /rpt_family="AT_rich"
     repeat_region   complement(26704. .26805)
                     /rpt_family="RMER15"
     repeat_region   complement(26921. .27179)
                     /rpt_family="Lx8"

Alignment Scores:
Pred. No.:                3.41e+04        Length:          186051
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC112957 (1-186051)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      49372 AGTTTAATCACGTGGAACATT 49352

RESULT 49
```

```
AC011844
LOCUS       AC011844               191676 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 12 clone RP11-307E8 map 12, WORKING DRAFT
            SEQUENCE, 22 unordered pieces.
ACCESSION   AC011844
VERSION     AC011844.3  GI:7107846
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 191676)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 12, clone RP11-307E8
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 191676)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 191676)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
            Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K.,
            Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
            Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 28, 2000 this sequence version replaced gi:6587908.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L2611
                Center clone name: 307_E_8
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 148426 bases at least Q40
                Consensus quality: 171214 bases at least Q30
                Consensus quality: 181482 bases at least Q20
                Insert size: 133000; agarose-fp
                Insert size: 189576; sum-of-contigs
                Quality coverage: 4.8 in Q20 bases; agarose-fp
                Quality coverage: 3.4 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 22 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1138: contig of 1138 bp in length
            *     1139     1238: gap of 100 bp
            *     1239     2510: contig of 1272 bp in length
            *     2511     2610: gap of 100 bp
            *     2611     4716: contig of 2106 bp in length
            *     4717     4816: gap of 100 bp
            *     4817     6902: contig of 2086 bp in length
            *     6903     7002: gap of 100 bp
            *     7003    10042: contig of 3040 bp in length
            *    10043    10142: gap of 100 bp
            *    10143    13270: contig of 3128 bp in length
            *    13271    13370: gap of 100 bp
            *    13371    16754: contig of 3384 bp in length
            *    16755    16854: gap of 100 bp
            *    16855    18739: contig of 1885 bp in length
            *    18740    18839: gap of 100 bp
            *    18840    21926: contig of 3087 bp in length
            *    21927    22026: gap of 100 bp
            *    22027    26211: contig of 4185 bp in length
            *    26212    26311: gap of 100 bp
            *    26312    29775: contig of 3464 bp in length
            *    29776    29875: gap of 100 bp
            *    29876    35862: contig of 5987 bp in length
            *    35863    35962: gap of 100 bp
            *    35963    42223: contig of 6261 bp in length
            *    42224    42323: gap of 100 bp
            *    42324    52123: contig of 9800 bp in length
            *    52124    52223: gap of 100 bp
            *    52224    62392: contig of 10169 bp in length
            *    62393    62492: gap of 100 bp
            *    62493    72478: contig of 9986 bp in length
            *    72479    72578: gap of 100 bp
            *    72579    85551: contig of 12973 bp in length
            *    85552    85651: gap of 100 bp
            *    85652   102897: contig of 17246 bp in length
            *   102898   102997: gap of 100 bp
            *   102998   118677: contig of 15680 bp in length
            *   118678   118777: gap of 100 bp
            *   118778   136020: contig of 17243 bp in length
            *   136021   136120: gap of 100 bp
            *   136121   159054: contig of 22934 bp in length
            *   159055   159154: gap of 100 bp
            *   159155   191676: contig of 32522 bp in length.
FEATURES             Location/Qualifiers
     source          1. .191676
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12"
                     /clone="RP11-307E8"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1138
                     /note="assembly_fragment"
     gap             1139. .1238
```

```
                /estimated_length=100
     misc_feature    1239. .2510
                /note="assembly_fragment"
     gap             2511. .2610
                /estimated_length=100
     misc_feature    2611. .4716
                /note="assembly_fragment"
     gap             4717. .4816
                /estimated_length=100
     misc_feature    4817. .6902
                /note="assembly_fragment"
     gap             6903. .7002
                /estimated_length=100
     misc_feature    7003. .10042
                /note="assembly_fragment"
     gap             10043. .10142
                /estimated_length=100
     misc_feature    10143. .13270
                /note="assembly_fragment"
     gap             13271. .13370
                /estimated_length=100
     misc_feature    13371. .16754
                /note="assembly_fragment"
     gap             16755. .16854
                /estimated_length=100
     misc_feature    16855. .18739
                /note="assembly_fragment
                clone_end:SP6
                vector_side:left"
     gap             18740. .18839
                /estimated_length=100
     misc_feature    18840. .21926
                /note="assembly_fragment"
     gap             21927. .22026
                /estimated_length=100
     misc_feature    22027. .26211
                /note="assembly_fragment"
     gap             26212. .26311
                /estimated_length=100
     misc_feature    26312. .29775
                /note="assembly_fragment"
     gap             29776. .29875
                /estimated_length=100
     misc_feature    29876. .35862
                /note="assembly_fragment"
     gap             35863. .35962
                /estimated_length=100
     misc_feature    35963. .42223
                /note="assembly_fragment"
     gap             42224. .42323
                /estimated_length=100
     misc_feature    42324. .52123
                /note="assembly_fragment"
     gap             52124. .52223
                /estimated_length=100

Alignment Scores:
Pred. No.:              3.51e+04        Length:         191676
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC011844 (1-191676)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db     182120 TCCCTTATCACATGGAATATA 182140

RESULT 50
AC148903
LOCUS       AC148903               193842 bp    DNA     linear   HTG 21-APR-2004
DEFINITION  Otolemur garnettii clone CH256-44P16, WORKING DRAFT SEQUENCE, 3
            unordered pieces.
ACCESSION   AC148903
VERSION     AC148903.1  GI:46430798
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Otolemur garnettii (small-eared galago)
  ORGANISM  Otolemur garnettii
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Strepsirhini;
            Galagonidae; Otolemur.
REFERENCE   1  (bases 1 to 193842)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
            Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
            Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
            Idol,J.R., Jones,C., Karlins,E., Kim,H., Kwong,P., Laric,P.,
            Larson,S., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
            Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Mullikin,J.C., Paguirigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
            Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K., Sison,C.,
            Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L.,
            Wetherby,K.D., Young,A. and Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 193842)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     -------------- Genome Center
                Center: NIH Intramural Sequencing Center
                Center code: NISC
                Web site: http://www.nisc.nih.gov
                Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                Center project name: gqq
                Center clone name: 044P16
            -------------- Summary Statistics
                Sequencing vector: plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.990319
                Consensus quality: 192449 bases at least Q40
                Consensus quality: 192817 bases at least Q30
                Consensus quality: 193009 bases at least Q20
                Insert size: 204000; agarose-fp
                Insert size: 193642; sum-of-contigs
                Quality coverage: 10.93x in Q20 bases; agarose-fp
                Quality coverage: 11.51x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    56067: contig of 56067 bp in length
            *    56068    56167: gap of unknown length
            *    56168   113946: contig of 57779 bp in length
            *   113947   114046: gap of unknown length
            *   114047   193842: contig of 79796 bp in length.
FEATURES             Location/Qualifiers
     source          1. .193842
                     /organism="Otolemur garnettii"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:30611"
                     /clone="CH256-44P16"
                     /clone_lib="CH256"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .56067
                     /note="assembly_fragment"
     gap             56068. .56167
```

```
                     /estimated_length=unknown
     misc_feature    56168. .113946
                     /note="assembly_fragment"
     gap             113947. .114046
                     /estimated_length=unknown
     misc_feature    114047. .193842
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:                3.54e+04        Length:          193842
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       14              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC148903 (1-193842)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      88100 TCTCTTATCACCTGGAATATT 88120

RESULT 51
AC012404
LOCUS       AC012404               198857 bp    DNA     linear   PRI 31-MAR-2002
DEFINITION  Homo sapiens, clone RP11-39M21, complete sequence.
ACCESSION   AC012404
VERSION     AC012404.9  GI:19849373
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 198857)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-39M21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 198857)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 198857)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 31, 2002 this sequence version replaced gi:17386436.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L2261
                Center clone name: 39_M_21
            --------------.
FEATURES                Location/Qualifiers
     source             1. .198857
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /clone="RP11-39M21"
                        /clone_lib="RPCI-11 Human Male BAC"
     repeat_region      2302. .2322
                        /rpt_family="AT_rich"
     repeat_region      3608. .3637
                        /rpt_family="AT_rich"
     repeat_region      complement(4201. .4504)
                        /rpt_family="AluSx"
     repeat_region      complement(5590. .5661)
                        /rpt_family="FLAM_A"
     repeat_region      complement(7741. .7870)
                        /rpt_family="MIR3"
     repeat_region      complement(7943. .8244)
                        /rpt_family="AluY"
     repeat_region      8383. .8404
                        /rpt_family="AT_rich"
     repeat_region      8679. .8826
                        /rpt_family="L1ME4A"
     repeat_region      8827. .8862
                        /rpt_family="(TG)n"
     repeat_region      8863. .9029
                        /rpt_family="L1ME4A"
     repeat_region      9329. .9365
                        /rpt_family="(CA)n"
     repeat_region      complement(9817. .9933)
                        /rpt_family="L2"
     repeat_region      10650. .10740
                        /rpt_family="MIR"
     repeat_region      11839. .11903
                        /rpt_family="(CATATA)n"
     repeat_region      12355. .12673
                        /rpt_family="AluY"
     repeat_region      complement(13683. .13743)
                        /rpt_family="L2"
     repeat_region      14846. .15143
                        /rpt_family="AluSx"
     repeat_region      16234. .16278
                        /rpt_family="T-rich"
     repeat_region      complement(17849. .17937)
                        /rpt_family="MLT1A2"
     repeat_region      complement(17922. .18019)
                        /rpt_family="MLT1A2"
     repeat_region      18739. .18912
```

```
                     /rpt_family="MIR3"
     repeat_region   complement(18970. .19382)
                     /rpt_family="L1MB8"
     repeat_region   20364. .20453
                     /rpt_family="(TATATG)n"
     repeat_region   21037. .21149
                     /rpt_family="L2"
     repeat_region   21263. .21293
                     /rpt_family="polypyrimidine"
     repeat_region   22859. .23122
                     /rpt_family="L2"
     repeat_region   complement(23182. .23301)
                     /rpt_family="L2"
     repeat_region   23305. .23377
                     /rpt_family="L2"
     repeat_region   23503. .23541
                     /rpt_family="L2"
     repeat_region   complement(23664. .23915)
                     /rpt_family="HAL1"
     repeat_region   complement(24113. .24314)
                     /rpt_family="L1ME3"
     repeat_region   complement(24763. .24971)
                     /rpt_family="L1MEd"
     repeat_region   25132. .25207
                     /rpt_family="MER5A"
     repeat_region   25228. .25527
                     /rpt_family="AluSg"
     repeat_region   complement(29303. .29482)
                     /rpt_family="MER96B"
     repeat_region   complement(31428. .31468)
                     /rpt_family="MADE1"
     repeat_region   31469. .31525
                     /rpt_family="MIR"
     repeat_region   complement(32574. .32768)
                     /rpt_family="MIR"
     repeat_region   complement(32997. .33259)
                     /rpt_family="AluSg"
     repeat_region   33731. .33811
                     /rpt_family="MIR"
     repeat_region   complement(33887. .34030)
                     /rpt_family="MLT1K"
     repeat_region   complement(34590. .34777)
                     /rpt_family="MIR"
     repeat_region   complement(36228. .36375)
                     /rpt_family="MIR"
     repeat_region   36660. .36688
                     /rpt_family="AT_rich"
     repeat_region   complement(36762. .37048)
                     /rpt_family="AluJb"
     repeat_region   complement(37474. .37638)
                     /rpt_family="L1ME3"
     repeat_region   complement(37655. .37961)
                     /rpt_family="AluY"
     repeat_region   38868. .38919
                     /rpt_family="AT_rich"
     repeat_region   39123. .39342
                     /rpt_family="MER58A"
     repeat_region   complement(40870. .41178)
                     /rpt_family="AluSx"
     repeat_region   41184. .41243
                     /rpt_family="MER5A"
     repeat_region   complement(41557. .41643)
                     /rpt_family="L2"
     repeat_region   complement(42521. .42834)
                     /rpt_family="AluJo"
     repeat_region   complement(42856. .42979)
                     /rpt_family="L2"
     repeat_region   complement(43150. .43257)
                     /rpt_family="L2"
     repeat_region   complement(43979. .44283)
                     /rpt_family="AluSq"
     repeat_region   complement(44299. .44917)
                     /rpt_family="L1MC4a"
     repeat_region   45208. .45373
                     /rpt_family="MIR3"
     repeat_region   45835. .45856
                     /rpt_family="(T)n"
     repeat_region   complement(46413. .46720)
                     /rpt_family="AluSx"
     repeat_region   complement(46742. .47018)
                     /rpt_family="L2"

Alignment Scores:
Pred. No.:              3.63e+04        Length:         198857
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC012404 (1-198857)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      62678 TCTCTAATTACCTGGAACATT 62698

RESULT 52
AC154641
LOCUS       AC154641               201511 bp    DNA     linear   HTG 30-DEC-2004
DEFINITION  Mus musculus chromosome 16 clone RP23-276C17, WORKING DRAFT
            SEQUENCE, 11 unordered pieces.
ACCESSION   AC154641 AC116756
VERSION     AC154641.1  GI:56900267
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 201511)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 201511)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On Dec 30, 2004 this sequence version replaced gi:50355833.

            ---------------- Genome Center ----------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information -------------------
            Center project name: M_BA0276C17
            Drafting center: WIBR
            ------------------------------------------------------
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 194862 bases at least Q40
            Consensus quality: 196442 bases at least Q30
            Consensus quality: 197550 bases at least Q20
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
```

```
            * be preserved.
            *        1     1351: contig of 1351 bp in length
            *     1352     1451: gap of unknown length
            *     1452     2642: contig of 1191 bp in length
            *     2643     2742: gap of unknown length
            *     2743     3981: contig of 1239 bp in length
            *     3982     4081: gap of unknown length
            *     4082     5217: contig of 1136 bp in length
            *     5218     5317: gap of unknown length
            *     5318     6609: contig of 1292 bp in length
            *     6610     6709: gap of unknown length
            *     6710     9555: contig of 2846 bp in length
            *     9556     9655: gap of unknown length
            *     9656    13588: contig of 3933 bp in length
            *    13589    13688: gap of unknown length
            *    13689    24309: contig of 10621 bp in length
            *    24310    24409: gap of unknown length
            *    24410    41354: contig of 16945 bp in length
            *    41355    41454: gap of unknown length
            *    41455   200194: contig of 158740 bp in length
            *   200195   200294: gap of unknown length
            *   200295   201511: contig of 1217 bp in length.
FEATURES             Location/Qualifiers
     source          1. .201511
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="16"
                     /clone="RP23-276C17"
     misc_feature    1. .1351
                     /note="assembly_name:Contig14"
     gap             1352. .1451
                     /estimated_length=unknown
     misc_feature    1452. .2642
                     /note="assembly_name:Contig15"
     gap             2643. .2742
                     /estimated_length=unknown
     misc_feature    2743. .3981
                     /note="assembly_name:Contig19"
     gap             3982. .4081
                     /estimated_length=unknown
     misc_feature    4082. .5217
                     /note="assembly_name:Contig20"
     gap             5218. .5317
                     /estimated_length=unknown
     misc_feature    5318. .6609
                     /note="assembly_name:Contig21"
     gap             6610. .6709
                     /estimated_length=unknown
     misc_feature    6710. .9555
                     /note="assembly_name:Contig26"
     gap             9556. .9655
                     /estimated_length=unknown
     misc_feature    9656. .13588
                     /note="assembly_name:Contig27"
     gap             13589. .13688
                     /estimated_length=unknown
     misc_feature    13689. .24309
                     /note="assembly_name:Contig28
                     clone_end:T7
                     vector_side:right"
     gap             24310. .24409
                     /estimated_length=unknown
     misc_feature    24410. .41354
                     /note="assembly_name:Contig29"
     gap             41355. .41454
                     /estimated_length=unknown
     misc_feature    41455. .200194
                     /note="assembly_name:Contig30"
     gap             200195. .200294
                     /estimated_length=unknown
     misc_feature    200295. .201511
                     /note="assembly_name:Contig9"

ORIGIN

Alignment Scores:
Pred. No.:              3.67e+04        Length:         201511
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC154641 (1-201511)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     121528 TCATTAATAACCTGGAATATC 121548

RESULT 53
AL928789/c
LOCUS       AL928789              202887 bp    DNA     linear   ROD 07-NOV-2002
DEFINITION  Mouse DNA sequence from clone RP23-435F21 on chromosome 2, complete
            sequence.
ACCESSION   AL928789
VERSION     AL928789.4  GI:24474491
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 202887)
  AUTHORS   Howden,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 1, 2002 this sequence version replaced gi:23896231.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-435F21 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .202887
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
                     /clone="RP23-435F21"
                     /clone_lib="RPCI-23"

ORIGIN
```

```
Alignment Scores:
Pred. No.:                3.69e+04        Length:          202887
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AL928789 (1-202887)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      91065 TCTCTTATTACCTGGAATATT 91045

RESULT 54
AC120951/c
LOCUS       AC120951               205582 bp    DNA     linear   HTG 19-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-451P6, WORKING DRAFT SEQUENCE.
ACCESSION   AC120951
VERSION     AC120951.5  GI:24635440
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 205582)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 205582)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 205582)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 6, 2002 this sequence version replaced gi:23907972.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GWRT
                Center clone name: CH230-451P6
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 170359 bases at least Q40
                Consensus quality: 172159 bases at least Q30
                Consensus quality: 173141 bases at least Q20
                Estimated insert size: 172237; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *         1   205582: contig of 205582 bp in length.
FEATURES                Location/Qualifiers
     source             1. .205582
                        /organism="Rattus norvegicus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10116"
                        /clone="CH230-451P6"
     misc_feature       1. .1276
                        /note="wgs_end_extension
                        clone_end:Sp6"
     misc_feature       1461. .3565
                        /note="wgs_end_extension
                        clone_end:Sp6"
     misc_feature       6403. .7275
                        /note="clone_boundary
```

```
                clone_end:Sp6
                site:
                end_sequence:BZ125882"
     misc_feature    complement(169435. .170317)
                /note="clone_boundary
                clone_end:T7
                site:
                end_sequence:BZ125866"
     misc_feature    173399. .174794
                /note="wgs_end_extension
                clone_end:T7"
     misc_feature    204149. .205582
                /note="wgs_end_extension
                clone_end:T7"
ORIGIN

Alignment Scores:
Pred. No.:              3.74e+04        Length:         205582
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC120951 (1-205582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     123632 AGTCTTATTACTTGGAACATT 123612

RESULT 55
BS000022
LOCUS       BS000022              214189 bp    DNA     linear   PRI 12-JUN-2004
DEFINITION  Pan troglodytes chromosome 22 clone:PTB-131J13, map 22, complete
            sequences.
ACCESSION   BS000022 BA000046
VERSION     BS000022.1  GI:37537289
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   The International Chimpanzee Chromosome 22 Consortium.
  TITLE     DNA sequence and comparative analysis of chimpanzee chromosome 22
  JOURNAL   Nature 429, 382-388 (2004)
REFERENCE   2  (bases 1 to 214189)
  AUTHORS   Hattori,M., Toyoda,A., Watanabe,H., Taylor,T.D., Kuroki,Y.,
            Fujiyama,A. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
            *Chinese National Human Genome Center at Shanghai, Shanghai, China;
            *GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
            of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research
            Center, Daejeon, Korea;
            *Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
            *National Institute of Genetics, Mishima, Japan;
            *National Yang Ming University Genome Research Center, Taipei,
            Taiwan;
            *RIKEN Genomic Sciences Center, Yokohama, Japan.
            -------------- Genome Center
            Center: RIKEN Genomic Sciences Center
            Center code: RIKEN
            Web site: http://hgp.gsc.riken.go.jp/
            Contact: hattori@gsc.riken.go.jp
            -------------- Project Information
            Center project name:The Chimpanzee Chromosome 22 Sequencing Project
            Center clone name: PTB-131J13
            -------------- Summary Statistics
            Sequencing vector: pUC18,pUC13,pTZ19R; 100% of reads Chemistry:
            Dye-terminator Big Dye and ET; 100% of reads Assembly program:
            Phrap; version 0.990329
            Consensus quality:   211,905 bases at least Q40
            Consensus quality:     2,026 bases at least Q30
            Consensus quality:       258 bases at least Q20
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30);
            an attempt was made to resolve all sequencing problems, such as
            compressions and repeats; all regions were covered by at one
            plasmid
            subclone or more than one M13 subclone;
            and the assembly was confirmed by restriction digest.
            --------------
            Source information:
            The RPCI-43 chimpanzee BAC library was prepared from DNA isolated
            from the blood of a single male chimpanzee using published
            protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the
            chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
            Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
            Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
            characteristics are described at
            http://www.chori.org/bacpac/mchimp43.htm.
            The clone may be obtained from Pieter J. de Jong and coworkers
            (http://www.chori.org/bacpac).
            VECTOR:  pBACe3.6
            The CHORI-251 chimpanzee BAC library was prepared from DNA isolated
            from the blood of a single male chimpanzee using published
            protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the
            chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
            Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
            Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
            characteristics are described at
            http://www.chori.org/bacpac/chimpanzee251.htm.
            The clone may be obtained from Pieter J. de Jong and coworkers
            (http://www.chori.org/bacpac).
            VECTOR:  pTARBAC2.1
            The PTB1 chimpanzee BAC library was prepared from DNA isolated from
            cultured cells established from the blood of a single male
            chimpanzee.
            Clones may be obtained from Asao Fujiyama and co-workers
            (http://www.gsc.riken.go.jp).
            VECTOR:  pKS145
            The PTF22 chimpanzee Fosmid library was prepared from DNA isolated
            from cultured cells established from the blood of a single male
            chimpanzee.
            Clones may be obtained from Asao Fujiyama and co-workers
            (http://www.gsc.riken.go.jp).
            VECTOR:  pKS143
            --------------
            Sequence Quality Assessment:
            This entry has been annotated with sequence
            estimates computed by the Phrap assembly program.
            All manually edited bases have been reduced to quality zero.
            Quality levels above 40 are expected to have less than 1 error in
            10,000 bp.
            ----------------------
            Neighboring clones: PTB-141K15(left) and PTB-018G23(right).
FEATURES                Location/Qualifiers
     source             1. .214189
                        /organism="Pan troglodytes"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9598"
                        /chromosome="22"
                        /clone="PTB-131J13"
                        /clone_lib="PTB1 chimpanzee BAC"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                3.88e+04        Length:          214189
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BS000022 (1-214189)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      20858 AGTTTGATAACATGGAACATT 20878

RESULT 56
AC152083
LOCUS       AC152083               214287 bp    DNA     linear   HTG 01-JUL-2005
DEFINITION  Bos taurus clone CH240-2J13, *** SEQUENCING IN PROGRESS ***, 19
            unordered pieces.
ACCESSION   AC152083
VERSION     AC152083.4  GI:68227252
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 214287)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 214287)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-OCT-2004) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 214287)
  AUTHORS   Cow Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jun 26, 2005 this sequence version replaced gi:58038148.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: FABG
                Center clone name: CH240-2J13
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 206062 bases at least Q40
                Consensus quality: 207854 bases at least Q30
                Consensus quality: 209322 bases at least Q20
                Estimated insert size: 210645; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 19 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     6195: contig of 6195 bp in length
            *     6196     6245: gap of 50 bp
            *     6246    33572: contig of 27327 bp in length
            *    33573    33672: gap of unknown length
            *    33673    47814: contig of 14142 bp in length
            *    47815    47864: gap of 50 bp
            *    47865    84772: contig of 36908 bp in length
            *    84773    84822: gap of 50 bp
            *    84823    86731: contig of 1909 bp in length
            *    86732    86781: gap of 50 bp
            *    86782    97707: contig of 10926 bp in length
            *    97708    99475: gap of 1768 bp
            *    99476   100498: contig of 1023 bp in length
            *   100499   100548: gap of 50 bp
            *   100549   122202: contig of 21654 bp in length
            *   122203   122252: gap of 50 bp
```

```
          *   122253   139783: contig of 17531 bp in length
          *   139784   139833: gap of 50 bp
          *   139834   144186: contig of 4353 bp in length
          *   144187   144236: gap of 50 bp
          *   144237   155649: contig of 11413 bp in length
          *   155650   155699: gap of 50 bp
          *   155700   159612: contig of 3913 bp in length
          *   159613   159662: gap of 50 bp
          *   159663   166383: contig of 6721 bp in length
          *   166384   166433: gap of 50 bp
          *   166434   174142: contig of 7709 bp in length
          *   174143   174192: gap of 50 bp
          *   174193   192338: contig of 18146 bp in length
          *   192339   193403: gap of 1065 bp
          *   193404   209185: contig of 15782 bp in length
          *   209186   209285: gap of unknown length
          *   209286   210701: contig of 1416 bp in length
          *   210702   210801: gap of unknown length
          *   210802   212504: contig of 1703 bp in length
          *   212505   212604: gap of unknown length
          *   212605   214287: contig of 1683 bp in length.
FEATURES             Location/Qualifiers
     source          1. .214287
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9913"
                     /clone="CH240-2J13"
     gap             6196. .6245
                     /estimated_length=50
     gap             33573. .33672
                     /estimated_length=unknown
     gap             47815. .47864
                     /estimated_length=50
     gap             84773. .84822
                     /estimated_length=50
     gap             86732. .86781
                     /estimated_length=50
     gap             97708. .99475
                     /estimated_length=1768
     gap             100499. .100548
                     /estimated_length=50
     gap             122203. .122252
                     /estimated_length=50
     gap             139784. .139833
                     /estimated_length=50
     gap             144187. .144236
                     /estimated_length=50
     gap             155650. .155699
                     /estimated_length=50
     gap             159613. .159662
                     /estimated_length=50
     gap             166384. .166433
                     /estimated_length=50
     gap             174143. .174192
                     /estimated_length=50
     gap             192339. .193403
                     /estimated_length=1065
     gap             209186. .209285
                     /estimated_length=unknown
     gap             210702. .210801
                     /estimated_length=unknown
     gap             212505. .212604
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:                3.88e+04      Length:          214287
Score:                    38.00         Matches:         7
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       14            Gaps:            0
```

```
US-10-757-745-2_COPY_115_121 (1-7) x AC152083 (1-214287)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     146675 TCCTTGATAACATGGAACATT 146695

RESULT 57
AC115541
LOCUS       AC115541               218833 bp    DNA     linear   HTG 10-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-40C8, *** SEQUENCING IN PROGRESS ***,
            2 unordered pieces.
ACCESSION   AC115541
VERSION     AC115541.6  GI:30522302
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 218833)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 218833)
  AUTHORS   Worley,K.C.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 218833)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:24818369.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GESL
                Center clone name: CH230-40C8
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 212249 bases at least Q40
                Consensus quality: 214966 bases at least Q30
                Consensus quality: 216584 bases at least Q20
                Estimated insert size: 230246; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   217029: contig of 217029 bp in length
            *  217030   217129: gap of unknown length
            *  217130   218833: contig of 1704 bp in length.
FEATURES             Location/Qualifiers
     source          1. .218833
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-40C8"
     misc_feature    96300. .98297
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH315248"
     gap             217030. .217129
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:                3.95e+04        Length:          218833
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0

DB:                       14              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC115541 (1-218833)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db      43979 TCACTGATAACCTGGAATATT 43999

RESULT 58
AC094940/c
LOCUS       AC094940               229363 bp    DNA     linear   HTG 09-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-6E9, *** SEQUENCING IN PROGRESS ***,
            5 unordered pieces.
ACCESSION   AC094940
VERSION     AC094940.5  GI:30466994
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 229363)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
```

```
REFERENCE   2  (bases 1 to 229363)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 229363)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 9, 2003 this sequence version replaced gi:22771362.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
                Center project name: GBTU
                Center clone name: CH230-6E9
            ---------- Summary Statistics
                Assembly program: Atlas;
                Consensus quality: 208118 bases at least Q40
                Consensus quality: 212027 bases at least Q30
                Consensus quality: 214790 bases at least Q20
                Estimated insert size: 226740; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            ----------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   131104: contig of 131104 bp in length
            *  131105   131204: gap of unknown length
            *  131205   223436: contig of 92232 bp in length
            *  223437   223536: gap of unknown length
            *  223537   224657: contig of 1121 bp in length
            *  224658   224757: gap of unknown length
            *  224758   227566: contig of 2809 bp in length
            *  227567   227666: gap of unknown length
            *  227667   229363: contig of 1697 bp in length.
FEATURES             Location/Qualifiers
     source          1. .229363
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-6E9"
     gap             131105. .131204
                     /estimated_length=unknown
     gap             223437. .223536
                     /estimated_length=unknown
     gap             224658. .224757
                     /estimated_length=unknown
     gap             227567. .227666
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:              4.12e+04        Length:         229363
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC094940 (1-229363)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     206144 TCACTAATAACTTGGAATATC 206124

RESULT 59
AC113013/c
LOCUS       AC113013               235823 bp    DNA     linear   ROD 17-DEC-2004
DEFINITION  Mus musculus chromosome 7, clone RP23-174K7, complete sequence.
ACCESSION   AC113013
VERSION     AC113013.15  GI:56693440
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 235823)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 7, clone RP23-174K7
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 235823)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-FEB-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 235823)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
            Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
            Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
            DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
            Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
            Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
            Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
            MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
```

```
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
            Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
            Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 235823)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
            Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
            Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
            DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
            Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
            Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
            Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
            MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
            Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
            Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Dec 17, 2004 this sequence version replaced gi:54312251.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@broad.mit.edu
            -------------- Project Information
                Center project name: L23481
                Center clone name: 174_K_7
            --------------.
FEATURES             Location/Qualifiers
     source          1. .235823
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="7"
                     /map="7"
                     /clone="RP23-174K7"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     repeat_region   complement(2. .107)
                     /rpt_family="MSTD"
     repeat_region   631. .749
                     /rpt_family="B4A"
     repeat_region   755. .795
                     /rpt_family="(CA)n"
     repeat_region   801. .856
                     /rpt_family="(GA)n"
     repeat_region   964. .1001
                     /rpt_family="(CA)n"
     repeat_region   1022. .1076
                     /rpt_family="(GA)n"
     repeat_region   3132. .3196
                     /rpt_family="(GA)n"
     repeat_region   3196. .3426
                     /rpt_family="GA-rich"
     repeat_region   4496. .4598
                     /rpt_family="MIR"
     repeat_region   4666. .4784
                     /rpt_family="CT-rich"
     repeat_region   complement(5220. .5420)
                     /rpt_family="ID_B1"
     repeat_region   complement(5542. .5719)
                     /rpt_family="L1MB3"
     repeat_region   5880. .5955
                     /rpt_family="(TG)n"
     repeat_region   complement(6005. .6260)
                     /rpt_family="B4A"
     repeat_region   complement(6267. .6425)
                     /rpt_family="L1M4"
     repeat_region   6435. .6574
                     /rpt_family="RSINE1"
     repeat_region   complement(6823. .7217)
                     /rpt_family="MTC"
     repeat_region   complement(7497. .7969)
                     /rpt_family="L1M4"
     repeat_region   complement(7943. .8433)
                     /rpt_family="L1_MM"
     repeat_region   8429. .8760
                     /rpt_family="Lx"
     repeat_region   8761. .8791
                     /rpt_family="(CA)n"
     repeat_region   8792. .8862
                     /rpt_family="Lx"
     unsure          complement(8793. .8802)
                     /note="<30 Qual SNGL region"
     unsure          complement(8794. .8797)
                     /note="probably 4 C's, possibly 3"
     unsure          complement(8802. .8803)
                     /note="probably GG, possibly G"
     repeat_region   complement(8867. .8966)
                     /rpt_family="L1M4"
     repeat_region   9402. .9503
                     /rpt_family="B1_MM"
     repeat_region   9522. .9549
                     /rpt_family="(CAAA)n"
     repeat_region   9552. .9587
                     /rpt_family="(GGAA)n"
     repeat_region   9588. .9643
                     /rpt_family="(GGAA)n"
     repeat_region   complement(9705. .10004)
                     /rpt_family="MTD"
     repeat_region   complement(10070. .10122)
                     /rpt_family="L2"
     repeat_region   10317. .10381
                     /rpt_family="CT-rich"
     repeat_region   complement(11204. .11280)
                     /rpt_family="L2"
     repeat_region   11412. .11486
                     /rpt_family="(TG)n"
     repeat_region   11502. .11657
                     /rpt_family="(TG)n"
     repeat_region   11706. .11727
                     /rpt_family="(TG)n"
     repeat_region   complement(11730. .12010)
                     /rpt_family="B4"
     repeat_region   13341. .13460
                     /rpt_family="B1_MM"
     repeat_region   13462. .13492
                     /rpt_family="(CAAA)n"
     repeat_region   complement(13538. .13805)
                     /rpt_family="L1_MM"
     repeat_region   13796. .13878
                     /rpt_family="L1_MM"
     repeat_region   13882. .13932
                     /rpt_family="MIR"
     unsure          14334. .14490
                     /note="single clone coverage"
```

```
Alignment Scores:
Pred. No.:                4.23e+04        Length:          235823
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x AC113013 (1-235823)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db      99976 TCATTGATTACTTGGAACATT 99956

RESULT 60
AC131436
LOCUS       AC131436               236814 bp    DNA     linear   HTG 09-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-5P1, *** SEQUENCING IN PROGRESS ***,
            3 unordered pieces.
ACCESSION   AC131436
VERSION     AC131436.4  GI:30467795
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 236814)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 236814)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 236814)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 9, 2003 this sequence version replaced gi:24818187.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GCHQ
                Center clone name: CH230-5P1
            -------------- Summary Statistics
                Assembly program: Atlas;
                Consensus quality: 222197 bases at least Q40
                Consensus quality: 224952 bases at least Q30
                Consensus quality: 226289 bases at least Q20
                Estimated insert size: 238764; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *         1   233065: contig of 233065 bp in length
            *    233066   233165: gap of unknown length
            *    233166   234362: contig of 1197 bp in length
            *    234363   234462: gap of unknown length
            *    234463   236814: contig of 2352 bp in length.
FEATURES             Location/Qualifiers
     source          1. .236814
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-5P1"
     misc_feature    199902. .200779
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
```

```
                    end_sequence:BH363605"
     misc_feature   complement(230476. .231247)
                    /note="clone_boundary
                    clone_end:T7
                    site:EcoRI
                    end_sequence:BH363603"
     misc_feature   231330. .233065
                    /note="wgs_end_extension
                    clone_end:T7"
     gap            233066. .233165
                    /estimated_length=unknown
     gap            234363. .234462
                    /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:              4.24e+04        Length:         236814
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC131436 (1-236814)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     223344 AGTCTTATTACTTGGAACATT 223364

RESULT 61
AL590646/c
LOCUS       AL590646              252828 bp    DNA     linear   HTG 10-JUL-2001
DEFINITION  Homo sapiens chromosome 9 clone RP11-370F12, 3 unordered pieces.
ACCESSION   AL590646
VERSION     AL590646.6  GI:14625672
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Skuce,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 6, 2001 this sequence version replaced gi:13752059.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA370F12
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 149340 bases at least Q40
            Consensus quality: 149818 bases at least Q30
            Consensus quality: 150084 bases at least Q20
            Insert size: 252628; sum-of-contigs
            Insert size: 148931; 6.6% error; agarose-fp
            Quality coverage: 3.50x in Q20 bases; sum-of-contigs Quality
            coverage: 5.97x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   151923: contig of 151923 bp in length
            *   151924   152023: gap of 100 bp
            *   152024   176128: contig of 24105 bp in length
            *   176129   176228: gap of 100 bp
            *   176229   252828: contig of 76600 bp in length.
FEATURES            Location/Qualifiers
     source         1. .252828
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /chromosome="9"
                    /clone="RP11-370F12"
                    /clone_lib="RPCI-11.2"
     misc_feature   1. .151923
                    /note="assembly_fragment:00278
                    fragment_chain:1
                    clone_end:T7
                    vector_side:left"
     misc_feature   152024. .176128
                    /note="assembly_fragment:02847
                    fragment_chain:1"
     misc_feature   176229. .252828
                    /note="assembly_fragment:02588"
ORIGIN

Alignment Scores:
Pred. No.:              4.5e+04         Length:         252828
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AL590646 (1-252828)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     131232 TCTTTGATCACATGGAATATT 131212

RESULT 62
AC073798
LOCUS       AC073798              255090 bp    DNA     linear   HTG 29-JUN-2000
DEFINITION  Mus musculus clone RP23-435F21, WORKING DRAFT SEQUENCE, 43
            unordered pieces.
ACCESSION   AC073798
VERSION     AC073798.1  GI:8810415
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 255090)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Mouse
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 255090)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 1904053
            Center clone name: RPCI-23_435F21
            -------
            Summary Statistics
```

```
Consensus quality: 224231 bases at least Q40
Consensus quality: 242804 bases at least Q30
Consensus quality: 246055 bases at least Q20
Estimated insert size: 149000; agarose-fp estimation
Estimated insert size: 250890; sum-of-contigs estimation
Quality coverage:  9.31 in Q20 bases; agarose-fp estimation
Quality coverage:  5.53 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 43 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1031: contig of 1031 bp in length
*     1032     1131: gap of unknown length
*     1132     2549: contig of 1418 bp in length
*     2550     2649: gap of unknown length
*     2650     3848: contig of 1199 bp in length
*     3849     3948: gap of unknown length
*     3949     5338: contig of 1390 bp in length
*     5339     5438: gap of unknown length
*     5439     7250: contig of 1812 bp in length
*     7251     7350: gap of unknown length
*     7351     9466: contig of 2116 bp in length
*     9467     9566: gap of unknown length
*     9567    10931: contig of 1365 bp in length
*    10932    11031: gap of unknown length
*    11032    12376: contig of 1345 bp in length
*    12377    12476: gap of unknown length
*    12477    14339: contig of 1863 bp in length
*    14340    14439: gap of unknown length
*    14440    15871: contig of 1432 bp in length
*    15872    15971: gap of unknown length
*    15972    17705: contig of 1734 bp in length
*    17706    17805: gap of unknown length
*    17806    19369: contig of 1564 bp in length
*    19370    19469: gap of unknown length
*    19470    21454: contig of 1985 bp in length
*    21455    21554: gap of unknown length
*    21555    23903: contig of 2349 bp in length
*    23904    24003: gap of unknown length
*    24004    25898: contig of 1895 bp in length
*    25899    25998: gap of unknown length
*    25999    28132: contig of 2134 bp in length
*    28133    28232: gap of unknown length
*    28233    30806: contig of 2574 bp in length
*    30807    30906: gap of unknown length
*    30907    34436: contig of 3530 bp in length
*    34437    34536: gap of unknown length
*    34537    37102: contig of 2566 bp in length
*    37103    37202: gap of unknown length
*    37203    41173: contig of 3971 bp in length
*    41174    41273: gap of unknown length
*    41274    45203: contig of 3930 bp in length
*    45204    45303: gap of unknown length
*    45304    50577: contig of 5274 bp in length
*    50578    50677: gap of unknown length
*    50678    55144: contig of 4467 bp in length
*    55145    55244: gap of unknown length
*    55245    59858: contig of 4614 bp in length
*    59859    59958: gap of unknown length
*    59959    64822: contig of 4864 bp in length
*    64823    64922: gap of unknown length
*    64923    69300: contig of 4378 bp in length
*    69301    69400: gap of unknown length
*    69401    75949: contig of 6549 bp in length
*    75950    76049: gap of unknown length
*    76050    82872: contig of 6823 bp in length
*    82873    82972: gap of unknown length
*    82973    88042: contig of 5070 bp in length
*    88043    88142: gap of unknown length
*    88143    95523: contig of 7381 bp in length
*    95524    95623: gap of unknown length
*    95624   102434: contig of 6811 bp in length
*   102435   102534: gap of unknown length
*   102535   109964: contig of 7430 bp in length
*   109965   110064: gap of unknown length
*   110065   117440: contig of 7376 bp in length
*   117441   117540: gap of unknown length
*   117541   126621: contig of 9081 bp in length
*   126622   126721: gap of unknown length
*   126722   138396: contig of 11675 bp in length
*   138397   138496: gap of unknown length
*   138497   148713: contig of 10217 bp in length
*   148714   148813: gap of unknown length
*   148814   158610: contig of 9797 bp in length
*   158611   158710: gap of unknown length
*   158711   171006: contig of 12296 bp in length
*   171007   171106: gap of unknown length
*   171107   181592: contig of 10486 bp in length
*   181593   181692: gap of unknown length
*   181693   195347: contig of 13655 bp in length
*   195348   195447: gap of unknown length
*   195448   210917: contig of 15470 bp in length
*   210918   211017: gap of unknown length
*   211018   228602: contig of 17585 bp in length
*   228603   228702: gap of unknown length
*   228703   255090: contig of 26388 bp in length.
FEATURES             Location/Qualifiers
     source          1. .255090
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="RP23-435F21"
                     /clone_lib="RPCI mouse BAC library 23"
     gap             1032. .1131
                     /estimated_length=unknown
     gap             2550. .2649
                     /estimated_length=unknown
     gap             3849. .3948
                     /estimated_length=unknown
     gap             5339. .5438
                     /estimated_length=unknown
     gap             7251. .7350
                     /estimated_length=unknown
     gap             9467. .9566
                     /estimated_length=unknown
     gap             10932. .11031
                     /estimated_length=unknown
     gap             12377. .12476
                     /estimated_length=unknown
     gap             14340. .14439
                     /estimated_length=unknown
     gap             15872. .15971
                     /estimated_length=unknown
     gap             17706. .17805
                     /estimated_length=unknown
     gap             19370. .19469
                     /estimated_length=unknown
     gap             21455. .21554
                     /estimated_length=unknown
     gap             23904. .24003
                     /estimated_length=unknown
     gap             25899. .25998
                     /estimated_length=unknown
     gap             28133. .28232
                     /estimated_length=unknown
     gap             30807. .30906
                     /estimated_length=unknown
     gap             34437. .34536
                     /estimated_length=unknown
     gap             37103. .37202
                     /estimated_length=unknown
     gap             41174. .41273
```

```
                /estimated_length=unknown
     gap        45204. .45303
                /estimated_length=unknown
     gap        50578. .50677
                /estimated_length=unknown
     gap        55145. .55244
                /estimated_length=unknown
     gap        59859. .59958
                /estimated_length=unknown
     gap        64823. .64922
                /estimated_length=unknown
     gap        69301. .69400
                /estimated_length=unknown
     gap        75950. .76049
                /estimated_length=unknown
     gap        82873. .82972
                /estimated_length=unknown
     gap        88043. .88142
                /estimated_length=unknown
     gap        95524. .95623
                /estimated_length=unknown
     gap        102435. .102534
                /estimated_length=unknown
     gap        109965. .110064

Alignment Scores:
Pred. No.:              4.54e+04        Length:         255090
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC073798 (1-255090)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     133361 TCTCTTATTACCTGGAATATT 133381

RESULT 63
AC094186
LOCUS       AC094186               255706 bp    DNA     linear   HTG 09-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-2K14, *** SEQUENCING IN PROGRESS ***,
            2 unordered pieces.
ACCESSION   AC094186
VERSION     AC094186.6  GI:30467744
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 255706)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 255706)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 255706)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 9, 2003 this sequence version replaced gi:24942737.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GADZ
                Center clone name: CH230-2K14
            -------------- Summary Statistics
                Assembly program: Atlas;
                Consensus quality: 244860 bases at least Q40
                Consensus quality: 247754 bases at least Q30
                Consensus quality: 249340 bases at least Q20
                Estimated insert size: 268126; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
```

```
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
            * NOTE: This sequence may represent more than one clone.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   254326: contig of 254326 bp in length
            *  254327   254426: gap of unknown length
            *  254427   255706: contig of 1280 bp in length.
FEATURES             Location/Qualifiers
     source          1. .255706
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-2K14"
     misc_feature    38438. .43555
                     /note="wgs_contig"
     misc_feature    43606. .46302
                     /note="wgs_contig"
     gap             254327. .254426
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:              4.55e+04        Length:         255706
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC094186 (1-255706)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db     180107 TCACTAATAACTTGGAATATC 180127

RESULT 64
AL445583/c
LOCUS       AL445583              348479 bp    DNA     linear   HTG 10-JUL-2001
DEFINITION  Homo sapiens chromosome 1 clone RP11-130J18, 49 unordered pieces.
ACCESSION   AL445583
VERSION     AL445583.9  GI:11995190
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mclay,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 28, 2000 this sequence version replaced gi:11995115.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA130J18
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator ET-amersham; 40% of reads Chemistry:
            Dye-terminator Big Dye; 59% of reads
            Consensus quality: 319527 bases at least Q40
            Consensus quality: 331637 bases at least Q30
            Consensus quality: 338755 bases at least Q20
            Insert size: 343679; sum-of-contigs
            Insert size: 228267; 48.0% error; agarose-fp
            Quality coverage: 3.54x in Q20 bases; sum-of-contigs Quality
            coverage: 6.60x in Q20 bases; agarose-fp
            --------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 49 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     5244: contig of 5244 bp in length
            *    5245     5344: gap of 100 bp
            *    5345     8561: contig of 3217 bp in length
            *    8562     8661: gap of 100 bp
            *    8662    11176: contig of 2515 bp in length
            *   11177    11276: gap of 100 bp
            *   11277    17711: contig of 6435 bp in length
            *   17712    17811: gap of 100 bp
            *   17812    22357: contig of 4546 bp in length
            *   22358    22457: gap of 100 bp
            *   22458    24879: contig of 2422 bp in length
            *   24880    24979: gap of 100 bp
            *   24980    79557: contig of 54578 bp in length
            *   79558    79657: gap of 100 bp
            *   79658    90547: contig of 10890 bp in length
            *   90548    90647: gap of 100 bp
            *   90648   118615: contig of 27968 bp in length
            *  118616   118715: gap of 100 bp
            *  118716   132646: contig of 13931 bp in length
            *  132647   132746: gap of 100 bp
            *  132747   144970: contig of 12224 bp in length
            *  144971   145070: gap of 100 bp
            *  145071   160944: contig of 15874 bp in length
            *  160945   161044: gap of 100 bp
            *  161045   190206: contig of 29162 bp in length
            *  190207   190306: gap of 100 bp
            *  190307   198700: contig of 8394 bp in length
            *  198701   198800: gap of 100 bp
            *  198801   202791: contig of 3991 bp in length
            *  202792   202891: gap of 100 bp
            *  202892   215487: contig of 12596 bp in length
            *  215488   215587: gap of 100 bp
            *  215588   219219: contig of 3632 bp in length
            *  219220   219319: gap of 100 bp
            *  219320   223305: contig of 3986 bp in length
            *  223306   223405: gap of 100 bp
            *  223406   226370: contig of 2965 bp in length
            *  226371   226470: gap of 100 bp
            *  226471   230403: contig of 3933 bp in length
            *  230404   230503: gap of 100 bp
            *  230504   232529: contig of 2026 bp in length
            *  232530   232629: gap of 100 bp
            *  232630   235117: contig of 2488 bp in length
            *  235118   235217: gap of 100 bp
            *  235218   238440: contig of 3223 bp in length
            *  238441   238540: gap of 100 bp
            *  238541   242282: contig of 3742 bp in length
            *  242283   242382: gap of 100 bp
            *  242383   244740: contig of 2358 bp in length
            *  244741   244840: gap of 100 bp
            *  244841   253991: contig of 9151 bp in length
            *  253992   254091: gap of 100 bp
            *  254092   257021: contig of 2930 bp in length
            *  257022   257121: gap of 100 bp
            *  257122   259350: contig of 2229 bp in length
            *  259351   259450: gap of 100 bp
            *  259451   261686: contig of 2236 bp in length
            *  261687   261786: gap of 100 bp
```

```
            *   261787   270474: contig of 8688 bp in length
            *   270475   270574: gap of 100 bp
            *   270575   275968: contig of 5394 bp in length
            *   275969   276068: gap of 100 bp
            *   276069   279460: contig of 3392 bp in length
            *   279461   279560: gap of 100 bp
            *   279561   293186: contig of 13626 bp in length
            *   293187   293286: gap of 100 bp
            *   293287   295854: contig of 2568 bp in length
            *   295855   295954: gap of 100 bp
            *   295955   298918: contig of 2964 bp in length
            *   298919   299018: gap of 100 bp
            *   299019   305708: contig of 6690 bp in length
            *   305709   305808: gap of 100 bp
            *   305809   311158: contig of 5350 bp in length
            *   311159   311258: gap of 100 bp
            *   311259   314235: contig of 2977 bp in length
            *   314236   314335: gap of 100 bp
            *   314336   316682: contig of 2347 bp in length
            *   316683   316782: gap of 100 bp
            *   316783   325107: contig of 8325 bp in length
            *   325108   325207: gap of 100 bp
            *   325208   327743: contig of 2536 bp in length
            *   327744   327843: gap of 100 bp
            *   327844   330004: contig of 2161 bp in length
            *   330005   330104: gap of 100 bp
            *   330105   332640: contig of 2536 bp in length
            *   332641   332740: gap of 100 bp
            *   332741   335236: contig of 2496 bp in length
            *   335237   335336: gap of 100 bp
            *   335337   337385: contig of 2049 bp in length
            *   337386   337485: gap of 100 bp
            *   337486   340993: contig of 3508 bp in length
            *   340994   341093: gap of 100 bp
            *   341094   343260: contig of 2167 bp in length
            *   343261   343360: gap of 100 bp
            *   343361   346283: contig of 2923 bp in length
            *   346284   346383: gap of 100 bp
            *   346384   348479: contig of 2096 bp in length.
FEATURES             Location/Qualifiers
     source          1. .348479
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-130J18"
                     /clone_lib="RPCI-11.1"
     misc_feature    1. .5244
                     /note="assembly_fragment:00025
                     fragment_chain:1"
     misc_feature    5345. .8561
                     /note="assembly_fragment:00516
                     fragment_chain:1"
     misc_feature    8662. .11176
                     /note="assembly_fragment:03202
                     fragment_chain:1"
     misc_feature    11277. .17711
                     /note="assembly_fragment:00100
                     fragment_chain:1"
     misc_feature    17812. .22357
                     /note="assembly_fragment:00263
                     fragment_chain:1"
     misc_feature    22458. .24879
                     /note="assembly_fragment:01354
                     fragment_chain:1"
     misc_feature    24980. .79557
                     /note="assembly_fragment:04401
                     fragment_chain:2"
     misc_feature    79658. .90547
                     /note="assembly_fragment:04074
                     fragment_chain:2"
     misc_feature    90648. .118615
                     /note="assembly_fragment:00736
                     fragment_chain:2"
     misc_feature    118716. .132646
                     /note="assembly_fragment:00038
                     fragment_chain:2"
     misc_feature    132747. .144970
                     /note="assembly_fragment:00577
                     fragment_chain:2"
     misc_feature    145071. .160944
                     /note="assembly_fragment:03276
                     fragment_chain:2"
     misc_feature    161045. .190206
                     /note="assembly_fragment:00536
                     fragment_chain:3"
     misc_feature    190307. .198700
                     /note="assembly_fragment:01928
                     fragment_chain:3"
     misc_feature    198801. .202791
                     /note="assembly_fragment:01816
                     fragment_chain:4"
     misc_feature    202892. .215487
                     /note="assembly_fragment:03100
                     fragment_chain:4"
     misc_feature    215588. .219219
                     /note="assembly_fragment:03529

Alignment Scores:
Pred. No.:              6.01e+04        Length:         348479
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AL445583 (1-348479)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db     224370 TCATTGATAACTTGGAATATA 224350

RESULT 65
BV289706/c
LOCUS       BV289706                444 bp    DNA     linear   STS 22-JAN-2005
DEFINITION  S232P6107RH2.T0 Beagle Canis familiaris STS genomic, sequence
            tagged site.
ACCESSION   BV289706
VERSION     BV289706.1  GI:57451941
KEYWORDS    STS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Lindblad-Toh,K.
  TITLE     The genome sequence of Canis familiaris
  JOURNAL   Unpublished (2004)
COMMENT
            Contact: Kerstin Lindblad-Toh
            Whitehead Institute for Biomedical Research, Center for Genome
            Research
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172521477
            Fax: 6172580903
            Email: kersli@genome.wi.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 444
            Protocol:
            WGS-discovery (WGS):
            Paired-end low-coverage whole genome shotgun reads were generated
            from 9 breeds
            (German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador
            Retriever, English
```

Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese Water Dog -100,000 each)
and five other canids (Chinese, Alaskan, Indian and Spanish Gray Wolf as well as the
Californian Coyote).
The WGS reads were placed uniquely on the CanFam1.0 boxer assembly and SNP detection was
carried out by SSAHA-SNP. 863872 reads were annotated as STSs and 485941 SNPs were
annotated with alleles from the boxer and the breed or canid from which the particular
read came. The validation rate for these SNPs was estimated at approximately 98%.
WGA-discovery (WGA) of Boxer/Poodle SNPs:
A second set of SNPs was generated using a similar methodology except that the contigs
from the 1.5x poodle assembly (Kirkness 2003) were used instead of WGS reads. Since this
sequence lacked base quality scores, arbitrary quality scores of phred 40 were assigned
before the poddle seqeunce was placed uniquely on the CanFam1.0 boxer assembly and SNP
detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated with alleles from the
boxer and the poodle. The validation rate for these SNPs was estimated at approximately TBD%.
Internal-WGA-discovery (I-WGA):
A third set of SNPs were discovered by comparing reads in the WGA assembly. SNPs were
defined as mismatch positions that had a base quality of >= 30 on both reads in a region
that aligned without gaps, and with at most one additional mismatch in the ten flanking
bases. For each allele, at least one additional read had to confirm it. 731476 SNPs were
annotated with alleles between the two boxer alleles. The validation rate for these SNPs
was estimated at approximately TBD%.

```
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Beagle"
                     /db_xref="taxon:9615"
                     /map="+ 3 22-400 21983507-21983885"
                     /clone_lib="Beagle"
     STS             <1. .>444
ORIGIN
```

Alignment Scores:

Pred. No.:	243	Length:	444
Score:	37.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BV289706 (1-444)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        413 TCACTAATTACTTGGAACGTA 393
```

```
RESULT 66
BV197165
LOCUS       BV197165                601 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm191843 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV197165
VERSION     BV197165.1  GI:48041661
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 601)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 601.
FEATURES             Location/Qualifiers
     source          1. .601
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>601
ORIGIN
```

Alignment Scores:

Pred. No.:	320	Length:	601
Score:	37.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BV197165 (1-601)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        549 AGTCTAATCACATGGAATGTT 569
```

```
RESULT 67
BV197166
LOCUS       BV197166                601 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm191844 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV197166
VERSION     BV197166.1  GI:48041662
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 601)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 601.
```

```
FEATURES             Location/Qualifiers
     source          1. .601
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>601
ORIGIN

Alignment Scores:
Pred. No.:              320          Length:        601
Score:                  37.00        Matches:       6
Percent Similarity:     100.0%       Conservative:  1
Best Local Similarity:  85.7%        Mismatches:    0
Query Match:            97.4%        Indels:        0
DB:                     10           Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x BV197166 (1-601)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        499 AGTCTAATCACATGGAATGTT 519

RESULT 68
BC087162/c
LOCUS       BC087162               3185 bp    mRNA    linear   ROD 06-JUL-2005
DEFINITION  Rattus norvegicus checkpoint with forkhead and ring finger domains
            (predicted), mRNA (cDNA clone MGC:95316 IMAGE:7135077), complete
            cds.
ACCESSION   BC087162
VERSION     BC087162.1  GI:56268937
KEYWORDS    MGC.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 3185)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
    PUBMED  12477932
REFERENCE   2  (bases 1 to 3185)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Howard Jacobs
            cDNA Library Preparation: Express Genomics
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 185 Row: n Column: 18
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 57164148.
FEATURES             Location/Qualifiers
     source          1. .3185
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="MGC:95316 IMAGE:7135077"
                     /tissue_type="Testis, rat (Brown Norway)"
                     /clone_lib="NIH_MGC_238"
                     /lab_host="DH10B"
                     /note="Vector: pExpress1"
     gene            1. .3185
                     /gene="Chfr_predicted"
                     /note="synonym: Chfr"
                     /db_xref="GeneID:288734"
                     /db_xref="RGD:1306360"
     CDS             55. .2046
                     /gene="Chfr_predicted"
                     /codon_start=1
                     /product="checkpoint with forkhead and ring finger domains
                     (predicted)"
                     /protein_id="AAH87162.1"
                     /db_xref="GI:56268938"
                     /db_xref="GeneID:288734"
                     /db_xref="RGD:1306360"
                     /translation="MEPHGEDEQPPPREPWGRLLRLGAEEDEPQILLWKREWTIGRRR
                     GCDLSFPSNKLVSGDHCKLRVDEKSGEVTLEDTSTNGTVINKLRVVKKQTYPLQSGDI
                     IYLVYRKNEPEHNVAYLYESLNGKQSLTQESLEANKENMPHVTKESSGPGQGDDPQVP
                     LLSPMAQTCLEEPQPSTSTSDLLPTASTSSTEPELTSAGQKHSSSSGPGTISISPKEC
                     SSLVANGELSSLSSVFQDNEASFSLLKTEDTEELEPAKKKMKGDGELDLNLQLSVADQ
                     RGIAQTSSEDVKDASVKPDKMEETLTCIICQDLLHDCVSLQPCMHTFCAACYSGWMER
                     SSLCPTCRCPVERICKNHILNNLVEAYLLQHPDKSRSEEDVRSMDARNKITQDMLQPK
                     VRRSFSDEEGSSEDLLELSDVDSESSDISQPYIVCRQCPEYRRQAVQSLPCPVPESEL
                     GATQALGGEAPSTSASLTTAPGYTCPLQGSHAICTCCFQPMPDRRAEHEQDSRIAPQQ
                     CAVCLQPFCHLYWGCTRTGCFGCLAPFCELNLGDKCLDGVLNNNNYESDILKNYLTTR
                     GLTWKNMLTESLLALQRGVFILSDYRITGNTVLCYCCGLRSFRELTYQYRQNIPASEL
                     PVTVTSRPDCYWGRNCRTQVKAHHAMKFNHICEQTRFKN"
ORIGIN

Alignment Scores:
Pred. No.:              1.44e+03     Length:        3185
Score:                  37.00        Matches:       6
Percent Similarity:     100.0%       Conservative:  1
Best Local Similarity:  85.7%        Mismatches:    0
Query Match:            97.4%        Indels:        0
DB:                     9            Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x BC087162 (1-3185)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        494 TCTTTGGTCACATGGAACATA 474

RESULT 69
BC049792/c
LOCUS       BC049792               3194 bp    mRNA    linear   ROD 30-JUN-2004
DEFINITION  Mus musculus checkpoint with forkhead and ring finger domains, mRNA
            (cDNA clone MGC:59419 IMAGE:6510480), complete cds.
ACCESSION   BC049792
VERSION     BC049792.1  GI:29437110
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3194)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3194)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. David Rowe
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 109 Row: l Column: 17
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 27370049.
FEATURES            Location/Qualifiers
     source         1. .3194
                    /organism="Mus musculus"
                    /mol_type="mRNA"
                    /db_xref="taxon:10090"
                    /clone="MGC:59419 IMAGE:6510480"
                    /tissue_type="Limb, mouse, day 10.5 and 11.5, pool of
                    undifferentiated limb containing undifferentiated limb
                    mesenchyme and early condensing mesenchyme."
                    /clone_lib="NIH_MGC_134"
                    /lab_host="DH10B"
                    /note="Vector: pCMV-SPORT6.1"
     gene           1. .3194
                    /gene="Chfr"
                    /note="synonyms: RNF116, C230082M18"
                    /db_xref="GeneID:231600"
                    /db_xref="MGI:2444898"
     CDS            71. .2065
                    /gene="Chfr"
                    /codon_start=1
                    /product="checkpoint with forkhead and ring finger
                    domains"
                    /protein_id="AAH49792.1"
                    /db_xref="GI:29437111"
                    /db_xref="GeneID:231600"
                    /db_xref="MGI:2444898"
                    /translation="MELHGEEQPPPPQEPWGRLLRLGAEEDEPQILLWKREWTIGRRR
                    GCDLSFPSNKLVSGDHCKLTVDEISGEVTLEDTSTNGTVINKLQVVKKQTYPLQSGDI
                    IYLVYRKNEPEHNVAYLYESLSGKQSLTQDSLEANKENMFHVTKDCSGPGQGDDPQVP
                    LLSPMAQTCLEEPQPSTSTSDLLPTASTSSTEPELTSAGQKHSSSSGPGNTSISPKGR
                    SSLVANGELSSLSPVFQDKEASFSLLESKDHEELEPAKKKMKGDGELDTNLQLLVSGQ
                    RGNAQTSSEDVKDASVKPDKMEETLTCIICQDLLHDCVSLQPCMHTFCAACYSGWMER
                    SSLCPTCRCPVERICKNHILNNLVEAYLIQHPDKSRSEEDVRSMDARNKITQDMLQPK
                    VRRSFSDEEGSSEDLLELSDVDSESSDISQPYIVCRQCPEYRRQAVQSLPCPVPESEL
                    GATLALGGEAPSTSASLPTAAPDYMCPLQGSHAICTCCFQPMPDRRAEREQDPRVAPQ
                    QCAVCLQPFCHLYWGCTRTGCFGCLAPFCELNLGDKCLDGVLNNNNYESDILKNYLAT
                    RGLTWKSVLTESLLALQRGVFMLSDYRITGNTVLCYCCGLRSFRELTYQYRQNIPASE
                    LPVTVTSRPDCYWGRNCRTQVKAHHAMKFNHICEQTRFKN"
ORIGIN

Alignment Scores:
Pred. No.:                1.44e+03        Length:          3194
Score:                    37.00           Matches:         6
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    85.7%           Mismatches:      0
Query Match:              97.4%           Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BC049792 (1-3194)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||
Db        510 TCTTTGGTCACATGGAACATA 490

RESULT 70
AC135376/c
LOCUS       AC135376               53047 bp    DNA     linear   HTG 13-OCT-2002
DEFINITION  Rattus norvegicus clone CH230-7N18, *** SEQUENCING IN PROGRESS ***,
            33 unordered pieces.
ACCESSION   AC135376
VERSION     AC135376.1  GI:23928479
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 53047)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
```

```
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., Mcleod,M., Mcneill,T., Meenen,E., Milosavljevic,A.,
            Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
            Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
            Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G.,
            Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,
            Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A.,
            Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E.,
            Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
            Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A.,
            Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S.,
            Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
            Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,
            Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,
            Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
            Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D.,
            Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
            Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R.,
            Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
            Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
            Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
            Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 53047)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GCNK
                Center clone name: CH230-7N18
            -------------- Summary Statistics
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 18027 bases at least Q40
                Consensus quality: 19915 bases at least Q30
                Consensus quality: 21196 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 33 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1045: contig of 1045 bp in length
            *     1046     1145: gap of unknown length
            *     1146     2264: contig of 1119 bp in length
            *     2265     2364: gap of unknown length
            *     2365     3459: contig of 1095 bp in length
            *     3460     3559: gap of unknown length
            *     3560     4636: contig of 1077 bp in length
            *     4637     4736: gap of unknown length
            *     4737     5950: contig of 1214 bp in length
            *     5951     6050: gap of unknown length
            *     6051     7388: contig of 1338 bp in length
            *     7389     7488: gap of unknown length
            *     7489     8971: contig of 1483 bp in length
            *     8972     9071: gap of unknown length
            *     9072    10569: contig of 1498 bp in length
            *    10570    10669: gap of unknown length
            *    10670    11953: contig of 1284 bp in length
            *    11954    12053: gap of unknown length
            *    12054    13135: contig of 1082 bp in length
            *    13136    13235: gap of unknown length
            *    13236    14412: contig of 1177 bp in length
            *    14413    14512: gap of unknown length
            *    14513    15513: contig of 1001 bp in length
            *    15514    15613: gap of unknown length
            *    15614    16965: contig of 1352 bp in length
            *    16966    17065: gap of unknown length
            *    17066    18165: contig of 1100 bp in length
            *    18166    18265: gap of unknown length
            *    18266    19308: contig of 1043 bp in length
            *    19309    19408: gap of unknown length
            *    19409    20877: contig of 1469 bp in length
            *    20878    20977: gap of unknown length
            *    20978    22109: contig of 1132 bp in length
            *    22110    22209: gap of unknown length
            *    22210    23472: contig of 1263 bp in length
            *    23473    23572: gap of unknown length
            *    23573    24944: contig of 1372 bp in length
            *    24945    25044: gap of unknown length
            *    25045    26991: contig of 1947 bp in length
            *    26992    27091: gap of unknown length
            *    27092    28807: contig of 1716 bp in length
            *    28808    28907: gap of unknown length
            *    28908    31265: contig of 2358 bp in length
            *    31266    31365: gap of unknown length
            *    31366    33069: contig of 1704 bp in length
            *    33070    33169: gap of unknown length
            *    33170    35374: contig of 2205 bp in length
            *    35375    35474: gap of unknown length
            *    35475    36508: contig of 1034 bp in length
            *    36509    36608: gap of unknown length
            *    36609    38186: contig of 1578 bp in length
            *    38187    38286: gap of unknown length
            *    38287    40240: contig of 1954 bp in length
            *    40241    40340: gap of unknown length
            *    40341    41599: contig of 1259 bp in length
            *    41600    41699: gap of unknown length
            *    41700    43832: contig of 2133 bp in length
            *    43833    43932: gap of unknown length
            *    43933    46528: contig of 2596 bp in length
            *    46529    46628: gap of unknown length
            *    46629    48174: contig of 1546 bp in length
            *    48175    48274: gap of unknown length
            *    48275    50392: contig of 2118 bp in length
            *    50393    50492: gap of unknown length
            *    50493    53047: contig of 2555 bp in length.
FEATURES             Location/Qualifiers
     source          1. .53047
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-7N18"
     gap             1046. .1145
                     /estimated_length=unknown
     gap             2265. .2364
                     /estimated_length=unknown
     gap             3460. .3559
                     /estimated_length=unknown
     gap             4637. .4736
                     /estimated_length=unknown
     gap             5951. .6050
                     /estimated_length=unknown
     gap             7389. .7488
                     /estimated_length=unknown
     gap             8972. .9071
                     /estimated_length=unknown
     gap             10570. .10669
```

```
                     /estimated_length=unknown
     gap             11954. .12053
                     /estimated_length=unknown
     gap             13136. .13235
                     /estimated_length=unknown
     gap             14413. .14512
                     /estimated_length=unknown
     gap             15514. .15613
                     /estimated_length=unknown
     gap             16966. .17065
                     /estimated_length=unknown
     gap             18166. .18265
                     /estimated_length=unknown
     gap             19309. .19408
                     /estimated_length=unknown
     gap             20878. .20977
                     /estimated_length=unknown
     gap             22110. .22209
                     /estimated_length=unknown
     gap             23473. .23572
                     /estimated_length=unknown

Alignment Scores:
Pred. No.:                1.81e+04        Length:         53047
Score:                    37.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              97.4%           Indels:         0
DB:                       14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC135376 (1-53047)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db      11042 TCATTGATCACCTGGAATGTT 11022

RESULT 71
AL162613
LOCUS       AL162613               79949 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-182C2 on chromosome 10 Contains
            the 5' end of the ADD3 gene for adducin 3 (gamma) and a CpG island,
            complete sequence.
ACCESSION   AL162613
VERSION     AL162613.22  GI:15617245
KEYWORDS    HTG; ADD3.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 79949)
  AUTHORS   Whitehead,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 13, 2001 this sequence version replaced gi:15131971.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-182C2 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .79949
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-182C2"
                     /clone_lib="RPCI-11.1"
     misc_feature    2000
                     /note="Clone_right_end: RP11-451M19"
     gene            join(5699. .5777,AL590628.6:32007. .32230,
                     AL590628.6:44159. .44297,AL590628.6:47641. .47792,
                     AL590628.6:48724. .48804,AL590628.6:49969. .49990)
                     /gene="ADD3"
                     /locus_tag="RP11-252O7.3-007"
     mRNA            join(5699. .5777,AL590628.6:32007. .32230,
                     AL590628.6:44159. .44297,AL590628.6:47641. .47792,
                     AL590628.6:48724. .48804,AL590628.6:49969. .49990)
                     /gene="ADD3"
                     /locus_tag="RP11-252O7.3-007"
                     /product="adducin 3 (gamma)"
                     /note="match: ESTs: BG723361.1"
     gene            join(15135. .15482,AL590628.6:32007. .32230,
                     AL590628.6:44159. .44297,AL590628.6:47641. .47792,
                     AL590628.6:48724. .48804,AL590628.6:49969. .50118,
                     AL590628.6:50593. .50736,AL590628.6:50848. .50946,
                     AL590628.6:53492. .53674,AL590628.6:55399. .55656,
                     AL590628.6:57208. .57327,AL590628.6:57799. .57885,
                     AL590628.6:61745. .61868,AL590628.6:64708. .65712)
                     /gene="ADD3"
                     /locus_tag="RP11-252O7.3-003"
     mRNA            join(15135. .15482,AL590628.6:32007. .32230,
                     AL590628.6:44159. .44297,AL590628.6:47641. .47792,
                     AL590628.6:48724. .48804,AL590628.6:49969. .50118,
                     AL590628.6:50593. .50736,AL590628.6:50848. .50946,
                     AL590628.6:53492. .53674,AL590628.6:55399. .55656,
                     AL590628.6:57208. .57327,AL590628.6:57799. .57885,
                     AL590628.6:61745. .61868,AL590628.6:64708. .65712)
                     /gene="ADD3"
                     /locus_tag="RP11-252O7.3-003"
                     /product="adducin 3 (gamma)"
                     /note="match: ESTs: AL600519.1 AL600528.1 AL600737.1
                     AL702110.1 AL702111.1 AL709221.1 BI459529.1 BI459783.1
                     BI544596.1 BI599681.1
                     match: cDNAs: D67031.1"
     gene            join(17293. .17630,AL590628.6:32007. .32230,
                     AL590628.6:44159. .44297,AL590628.6:47641. .47792,
                     AL590628.6:48724. .48804,AL590628.6:49969. .50118,
                     AL590628.6:50593. .50736,AL590628.6:50848. .50946,
                     AL590628.6:53492. .53674,AL590628.6:55399. .55656,
                     AL590628.6:57208. .57327,AL590628.6:57799. .57885,
                     AL590628.6:61745. .61868,AL590628.6:63687. .63782,
                     AL590628.6:64708. .65593)
                     /gene="ADD3"
                     /locus_tag="RP11-252O7.3-001"
     mRNA            join(17293. .17630,AL590628.6:32007. .32230,
                     AL590628.6:44159. .44297,AL590628.6:47641. .47792,
                     AL590628.6:48724. .48804,AL590628.6:49969. .50118,
                     AL590628.6:50593. .50736,AL590628.6:50848. .50946,
                     AL590628.6:53492. .53674,AL590628.6:55399. .55656,
                     AL590628.6:57208. .57327,AL590628.6:57799. .57885,
                     AL590628.6:61745. .61868,AL590628.6:63687. .63782,
                     AL590628.6:64708. .65593)
```

```
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-001"
                /product="adducin 3 (gamma)"
                /note="match: ESTs: AA017171.1 AA054174.1 AI538560.1
                AI763123.1 AI818488.1 AU132010.1 AW103081.1 AW103506.1
                BF132006.1 BF665245.1 BG615692.1 BM453623.1 BQ431006.1
                BU855635.1
                match: cDNAs: AF100425.1 BC004746.1"
     gene       join(17295. .17630,AL590628.6:32007. .32230,
                AL590628.6:44159. .44297,AL590628.6:47641. .47792,
                AL590628.6:48724. .48804,AL590628.6:49969. .50118,
                AL590628.6:50593. .50736,AL590628.6:50848. .50946,
                AL590628.6:53492. .53674,AL590628.6:55399. .55656,
                AL590628.6:57208. .57327,AL590628.6:57799. .57885,
                AL590628.6:61745. .61868,AL590628.6:64708. .66947)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-002"
     mRNA       join(17295. .17630,AL590628.6:32007. .32230,
                AL590628.6:44159. .44297,AL590628.6:47641. .47792,
                AL590628.6:48724. .48804,AL590628.6:49969. .50118,
                AL590628.6:50593. .50736,AL590628.6:50848. .50946,
                AL590628.6:53492. .53674,AL590628.6:55399. .55656,
                AL590628.6:57208. .57327,AL590628.6:57799. .57885,
                AL590628.6:61745. .61868,AL590628.6:64708. .66947)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-002"
                /product="adducin 3 (gamma)"
                /note="match: ESTs: AA315146.1 AA331788.1 AI056886.1
                AI379520.1 AI924752.1 AL135243.1 AL598826.1 AL709391.1
                AL713383.1 AU116934.1 AU120029.1 AU120982.1 AU122189.1
                AU123534.1 AU134494.1 AU138596.1 AV696958.1 AV698453.1
                AV724496.1 AV724742.1 AW149374.1 AW500746.1 AW517131.1
                AW952357.1 AW956707.1 AW965276.1 BE545756.1 BE819429.1
                BF666309.1 BF791724.1 BF814551.1 BF961262.1 BG109015.1
                BG330754.1 BG429041.1 BG435678.1 BG545349.1 BG546181.1
                BG576985.1 BG618588.1 BG623163.1 BG701842.1 BG715663.1
                BG742508.1 BG773158.1 BG820357.1 BI497049.1 BI551827.1
                BI553636.1 BI559957.1 BI560091.1 BI598843.1 BI599567.1
                BI599650.1 BI767694.1 BI769647.1 BI859671.1 BM147962.1
                BM153313.1 BM171744.1 BM458464.1 BM560499.1 BM721846.1
                BM803391.1 BM875698.1 BM969461.1 BM995424.1 BQ006529.1
                BQ014141.1 BQ016034.1 BQ019413.1 BQ130250.1 BQ574778.1
                BQ575118.1 BQ638385.1 BQ719076.1 BQ723220.1 BQ892959.1
                BU166302.1 BU674831.1 BU682086.1 BU688678.1 BU688873.1
                BU732810.1 BU787485.1 BU859168.1 CA314013.1 CA390909.1
                R84729.1 Z99390.1 Z99428.1
                match: cDNAs: U37122.1"
     gene       join(17314. .17376,17602. .17630,AL590628.6:32007. .32081)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-004"
     mRNA       join(17314. .17376,17602. .17630,AL590628.6:32007. .32081)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-004"
                /product="adducin 3 (gamma)"
                /note="match: ESTs: BI562227.1"
     gene       join(17539. .17630,AL590628.6:44159. .44297,
                AL590628.6:47641. .47792,AL590628.6:48724. .48741)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-005"
     gene       join(17539. .17630,AL590628.6:32007. .32230,
                AL590628.6:37382. .37435,AL590628.6:43523. .43737,
                AL590628.6:44159. .44281)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-013"
     mRNA       join(17539. .17630,AL590628.6:32007. .32230,
                AL590628.6:37382. .37435,AL590628.6:43523. .43737,
                AL590628.6:44159. .44281)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-005"
                /product="adducin 3 (gamma)"
                /note="match: ESTs: AA020975.1 AA054129.1"
     mRNA       join(17539. .17630,AL590628.6:44159. .44297,
                AL590628.6:47641. .47792,AL590628.6:48724. .48741)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-005"
                /product="adducin 3 (gamma)"
                /note="match: ESTs: AI219408.1 AI221608.1"
     gene       join(17540. .17630,AL590628.6:10622. .10679,
                AL590628.6:32007. .32086)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-006"
     mRNA       join(17540. .17630,AL590628.6:10622. .10679,
                AL590628.6:32007. .32086)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-006"
                /product="adducin 3 (gamma)"
                /note="match: ESTs: AW135967.1"
     gene       join(17562. .17630,AL590628.6:7000. .7091,
                AL590628.6:32007. .32230,AL590628.6:37382. .37435,
                AL590628.6:44159. .44276)
                /gene="ADD3"
                /locus_tag="RP11-252O7.3-009"

Alignment Scores:
Pred. No.:              2.62e+04        Length:         79949
Score:                  37.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            97.4%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AL162613 (1-79949)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db      51911 AGTCTAATAACATGGAATGTG 51931

RESULT 72
BX908798_00
WPCOMMENT
  Sequence split into 25 fragments  LOCUS BX908798 Accession BX908798
     Fragment Name      Begin     End
     BX908798_00            1    110000
     BX908798_01       100001    210000
     BX908798_02       200001    310000
     BX908798_03       300001    410000
     BX908798_04       400001    510000
     BX908798_05       500001    610000
     BX908798_06       600001    710000
     BX908798_07       700001    810000
     BX908798_08       800001    910000
     BX908798_09       900001   1010000
     BX908798_10      1000001   1110000
     BX908798_11      1100001   1210000
     BX908798_12      1200001   1310000
     BX908798_13      1300001   1410000
     BX908798_14      1400001   1510000
     BX908798_15      1500001   1610000
     BX908798_16      1600001   1710000
     BX908798_17      1700001   1810000
     BX908798_18      1800001   1910000
     BX908798_19      1900001   2010000
     BX908798_20      2000001   2110000
     BX908798_21      2100001   2210000
     BX908798_22      2200001   2310000
     BX908798_23      2300001   2410000
     BX908798_24      2400001   2414465
LOCUS       BX908798             2414465 bp    DNA     circular BCT 17-APR-2005
DEFINITION  Parachlamydia-related symbiont UWE25, complete genome.
ACCESSION   BX908798
VERSION     BX908798.1  GI:46399275
KEYWORDS    complete genome.
SOURCE      Parachlamydia sp. UWE25
  ORGANISM  Parachlamydia sp. UWE25
            Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae;
            Parachlamydia.
```

```
REFERENCE   1
  AUTHORS   Horn,M., Collingro,A., Schmitz-Esser,S., Beier,C.L., Purkhold,U.,
            Fartmann,B., Brandt,P., Nyakatura,G.J., Droege,M., Frishman,D.,
            Rattei,T., Mewes,H.W. and Wagner,M.
  TITLE     Illuminating the evolutionary history of chlamydiae
  JOURNAL   Science 304 (5671), 728-730 (2004)
   PUBMED   15073324
REFERENCE   2  (bases 1 to 2414465)
  AUTHORS   Horn,M., Collingro,A., Schmitz-Esser,S., Beier,C.L., Purkhold,U.,
            Fartmann,B., Brandt,P., Nyakatura,G.J., Droege,M., Frishman,D.,
            Rattei,T., Mewes,H. and Wagner,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2003) Submitted by Matthias Horn, Department of
            Microbial Ecology, University of Vienna, Althanstr. 14, A-1090
            Wien, Austria, Email: horn@microbial-ecology.net Complete
            annotation available at: http://mips.gsf.de/services/genomes/uwe25
FEATURES             Location/Qualifiers
     source          1. .2414465
                     /organism="Parachlamydia sp. UWE25"
                     /mol_type="genomic DNA"
                     /strain="UWE25"
                     /db_xref="taxon:264201"
     gene            complement(123. .338)
                     /locus_tag="pc0001"
     CDS             complement(123. .338)
                     /locus_tag="pc0001"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown protein"
                     /protein_id="CAF22725.1"
                     /db_xref="GI:46399276"
                     /db_xref="UniProt/TrEMBL:Q6MFC4"
                     /translation="MFSPLLVFLKTCAVSSLSSWLLLQSWPAHLGTSCQAKDKKYLLI
                     AIYFFAANILIIYIFSSVIQSTLNLAL"
     gene            462. .746
                     /locus_tag="pc0002"
     CDS             462. .746
                     /locus_tag="pc0002"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown protein"
                     /protein_id="CAF22726.1"
                     /db_xref="GI:46399277"
                     /db_xref="UniProt/TrEMBL:Q6MFC3"
                     /translation="MANLILSLPLCSKNKHLILNTLTTPYLMMGLIITECHVKPDVLL
                     VYWIDEEIGSHANLFRNKMGKSLPDRVEINFGNLNLTLKAGLSALPLKLM"
     gene            1055. .4024
                     /locus_tag="pc0003"
     CDS             1055. .4024
                     /locus_tag="pc0003"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="CAF22727.1"
                     /db_xref="GI:46399278"
                     /db_xref="UniProt/TrEMBL:Q6MFC2"
                     /translation="MCEKVNKEKINTILSLVKAPLNHPFFYPRIFALIAKQKKQERLA
                     ADKICAYIVEEYEEISRRLDRSQIQESCSVRNILRTGHLAALLIDDKGEINFAVLPTL
                     IQKFKSHLYSLGPNRQYDSKRQEHILKVLQFLLTNKELVRSLKKISKPLSNKIAEDII
                     KQTLQIPSQTVITDAHAKRAVLAAWLCYLRQNVGSCFATAPAEIIQNEQPEQFLQDMT
                     DLLATGRLKRTFGGAEYSVPISASWGSGDLKKPLLLSISSKKISPEIWYSPGLIVAFE
                     AVGLLQKNEKLKDKIKRLKDWIDPIILQHSRHTNYCIVTAEVILRTILRQSIGLSEQQ
                     IKDYEERPRGMIQSNLMQIPAIGKQTGGIGEKCANFLYLFEIGKNAFKSLADNALLKT
                     WEFTLASFSETKLEFTTWNLYASLGMGSHEQGGIGHCIQEQIQYKLDQANKLVQDAQY
                     EYEMMYTHVKTLESRIRHASEKEAQWLKMEYQNRANELYFIQDQRDDAHQRASALVNL
                     YESLYKIYLELFREYFQEVYDADMQEVLTGPFDDSPAGFRLLYKHGRSNSSQWTRIKN
                     QTDFIDSLTSFFVATEPQIASMLQNLGLEKDLGEVVTAIVNHVKTKEFLETAFYRMAA
                     AHRMPIPQNPLDHLDQIEKKPWVYTSGGTMNTLVSCYYRIEDKPQEGAKWVESEMELL
                     VFFADTLKHIPRKQMEPYLEGRRQSMLIQSPTHAFLLKPNLTPFKDAWIGDEFTYTSV
                     RDKFVKPAEFFIEKIVLDHNMIHSFLQELIEKIPENFQPRFKAVFQNLTGPLNPLYFR
                     EIIVDTLQQDRGLNIGKKSILSADELDSLLYSRLPYFSIYELKEKIYQILILLPGITS
                     SKADEILSYFDQISLSKSQYVMSANQLQDICKALICLSEMQTTFSIDYHWHISFTCQK
                     LGFAMPAPIIFADTNWVKDEFGFVVNPGTGRLELWRVDYTGSIGYPMSIWKEWVNGTR
                     TDLKWGIYIKPTEYGQV"
     gene            4428. .5840
                     /locus_tag="pc0004"
     CDS             4428. .5840
                     /locus_tag="pc0004"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="CAF22728.1"
                     /db_xref="GI:46399279"
                     /db_xref="UniProt/TrEMBL:Q6MFC1"
                     /translation="MPILMAAGFTSATIFAQQQETSFERQLNENDDQSLREFVQSKDN
                     IDIKKKSKNLDISGDIRFEWRTIREKGVALYADDSSKFYEDYRSLRGGSHVDAKGLPI
                     SHNDFDVEFNFKLKYSFGRSWAQAHLQFDNPAGLRGRNDCSGKFPIYNKHGSEVIKSL
                     PRNMQCCGKGSGESNLINLKRAYIGYNVYADGKHRFDMEIGRRKLDDVFDSEIQFSSR
                     FDGVLLKFASAIDEVMDWYWEGAAFVVDERVNHFGYVTEIGILNLFDTGLDLRYSFID
                     WKKTGINRCHVHNPLGTQFTNSQISFAYTIEPEIYCKRIPVEFYGGFLINHAAKAKAI
                     THHKKKNLAAYGGVYIGQVHKEGDWALDIEYIYVQAQAVSDFDVGSIGRGNLVNDDFA
                     DYIETSHLPSTSDYSSQNGYYYPARGNANFKGWRFEFLYAITDNFSIDTIYEFSRAED
                     SRIGGRHSYSDFELEAIYAF"
     gene            complement(5907. .8303)
                     /locus_tag="pc0005"
     CDS             complement(5907. .8303)
                     /locus_tag="pc0005"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="CAF22729.1"
                     /db_xref="GI:46399280"
                     /db_xref="GOA:Q6MFC0"
                     /db_xref="InterPro:IPR005593"
                     /db_xref="UniProt/Swiss-Prot:Q6MFC0"
                     /translation="MITKLPPTSESNNSLSDELLEKINAYWRAANYLSVGQIYLLDNP
                     LLKKPLSIKDIKPRLLGHWGTTPGLNFIYVHLNRIIKKFDLDMIYLIGPGHGGPGIVA
                     NTYLEGTYSEYYPHISQDAEGMQKLFKQFSFPGGIPSHVAPETPGSINEGGELGYSLS
                     HAFGAAFNNPDLIVGCVIGDGEAETGPLATAWHSNKFLNPAIDGAVLPILHLNGYKIA
                     NPTILARISQEELKSLFVGYGYKPYFVEGSDPQLMHRLMAETLDMIIKEIQEIQTNAR
                     KNGIVKRPIWPMIILRTPKGWTGPKEVDGKKTENYWRSHQVPFSDMKEEHIKLLEEWL
                     KSYRPEELFDETGKLFPILAELAPKGKRRMGDNFHANGGLLRKNLRLPDFRSYQLEVV
                     TPGVVEAESTRIMGSFLRDVLKLNPHNFLVTGPDETASNRLTNLFEVTQRRWLAEYLP
                     EDTDGGNLSLEGRVLEILSEHTCLGWLEGYILTGRHGFFSTYEAFAHIIDSMFNQHAK
                     WLQTTLTIPWRAPIPSINILLSSHVWRQDHNGFSHQDPGFIDLVMNKKAAVVRVYLPP
                     DSNTLLSVTDHCLRSLNYINVIVAGKQPALQYLSMEAAIKHCTAGIGIWEWASNDKGG
                     EPDVVLACAGDIPTLEILAAVDILRQSIPTLKIRVINVVDLMTLQPKEEHPHGLSNTD
                     FDEMFTTNKPIIFAYHGYPWLIHRLTYRRSNHSNIHVRGYKEEGTTTTPFDMVVLNEV
                     DRFHLVGDVIDRVPKLGYLAAYTKQEMRHKLIEHKEYIHQFGEDMPSVKNWKWPYGKE
                     "
     gene            complement(9037. .9249)
                     /locus_tag="pc0006"
     CDS             complement(9037. .9249)
                     /locus_tag="pc0006"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown protein"
                     /protein_id="CAF22730.1"
                     /db_xref="GI:46399281"
                     /db_xref="UniProt/TrEMBL:Q6MFB9"
                     /translation="MYENLYFKLFSSVFLRIFYSPPMAQSYQNAFNHSHLAKSINTWL
                     PFYLLFLIYMKIQNGSAKIHLLKHFQ"
     gene            complement(9395. .11098)
                     /locus_tag="pc0007"
     CDS             complement(9395. .11098)
                     /locus_tag="pc0007"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="CAF22731.1"
                     /db_xref="GI:46399282"
                     /db_xref="GOA:Q6MFB8"
                     /db_xref="InterPro:IPR002123"
                     /db_xref="InterPro:IPR007114"
                     /db_xref="InterPro:IPR011701"
                     /db_xref="UniProt/TrEMBL:Q6MFB8"
```

```
                     /translation="MIVYFFIQLEGMEYSHQILASTGATFVIPFLLFSALSGTLADRF
                     SKRDIIVTTKILELVIMALGVCAFAFQSKWGSYAILFLLATQSAIFGPSKYGIVPELV
                     SKDKISKANGLMTSFTFLAIILGTFLSSFLIQISGRNFIVSSIFCTFIALVGLVASFC
                     IEHTPPAGSEKKIQIFFLKQIFQTLKFARQEISLLPAILGAAYFLFVGAFTQLNIIPF
                     AVQSLQLSDIQGGYLFFLTALGIGTGALIAGKISGRIVELGLVPIGALGISVCAFFLD
                     IYSENIWAVIPIVTLLGVFGGIFEIPLDSYIQVASPKQMRGQMVAATNFLSFIGVLLA
                     SGLIYLNSEIFQISADKGFSLIGFITFGIFIIFMFQFFDYLTRFIASILSKLHFSTIF
                     KGKENIPDCPAVYVCTHTAWNDTLLILGSQRRRIRFFIEQEQHHSPWLRRLYGLLRVV
                     FIPDIEPLEKNQACLTEIRNTLKKGISVCIFVQSWDIYSEVEKLKHSTPFNHIIEEAN
                     YPIIPVYIQKGIKSPRFQFLTNSLKKFRVPASVVFGDLIYTGHQSILKSKHEHDLYFE
                     D"
     gene            11759. .15346
                     /gene="recC"
                     /locus_tag="pc0008"
     CDS             11759. .15346
                     /gene="recC"
                     /locus_tag="pc0008"

Alignment Scores:
Pred. No.:              3.5e+04         Length:         110000
Score:                  37.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            97.4%           Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BX908798_00 (1-110000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db      55187 TCGCTCATAACGTGGAATGTT 55207

RESULT 73
AC138524_3/c
WPCOMMENT
  Sequence split into 6 fragments  LOCUS AC138524 Accession AC138524
     Fragment Name      Begin     End
     AC138524_0             1     110000
     AC138524_1         100001    210000
     AC138524_2         200001    310000
     AC138524_3         300001    410000
     AC138524_4         400001    510000
     AC138524_5         500001    550277
Continuation (4 of 6) of AC138524 from base 300001 (AC138524 Homo sapiens chromosome 5 c

Alignment Scores:
Pred. No.:              3.5e+04         Length:         110000
Score:                  37.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            97.4%           Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AC138524_3 (1-110000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db      84149 TCTCTGATAACATGGAATGTG 84129

RESULT 74
AP008208_093
WPCOMMENT
  Sequence split into 360 fragments  LOCUS AP008208 Accession AP008208
     Fragment Name      Begin     End
     AP008208_000           1     110000
     AP008208_001       100001    210000
     AP008208_002       200001    310000
     AP008208_003       300001    410000
     AP008208_004       400001    510000
     AP008208_005       500001    610000
     AP008208_006       600001    710000
     AP008208_007       700001    810000
     AP008208_008       800001    910000
```

AP008208_009	900001	1010000
AP008208_010	1000001	1110000
AP008208_011	1100001	1210000
AP008208_012	1200001	1310000
AP008208_013	1300001	1410000
AP008208_014	1400001	1510000
AP008208_015	1500001	1610000
AP008208_016	1600001	1710000
AP008208_017	1700001	1810000
AP008208_018	1800001	1910000
AP008208_019	1900001	2010000
AP008208_020	2000001	2110000
AP008208_021	2100001	2210000
AP008208_022	2200001	2310000
AP008208_023	2300001	2410000
AP008208_024	2400001	2510000
AP008208_025	2500001	2610000
AP008208_026	2600001	2710000
AP008208_027	2700001	2810000
AP008208_028	2800001	2910000
AP008208_029	2900001	3010000
AP008208_030	3000001	3110000
AP008208_031	3100001	3210000
AP008208_032	3200001	3310000
AP008208_033	3300001	3410000
AP008208_034	3400001	3510000
AP008208_035	3500001	3610000
AP008208_036	3600001	3710000
AP008208_037	3700001	3810000
AP008208_038	3800001	3910000
AP008208_039	3900001	4010000
AP008208_040	4000001	4110000
AP008208_041	4100001	4210000
AP008208_042	4200001	4310000
AP008208_043	4300001	4410000
AP008208_044	4400001	4510000
AP008208_045	4500001	4610000
AP008208_046	4600001	4710000
AP008208_047	4700001	4810000
AP008208_048	4800001	4910000
AP008208_049	4900001	5010000
AP008208_050	5000001	5110000
AP008208_051	5100001	5210000
AP008208_052	5200001	5310000
AP008208_053	5300001	5410000
AP008208_054	5400001	5510000
AP008208_055	5500001	5610000
AP008208_056	5600001	5710000
AP008208_057	5700001	5810000
AP008208_058	5800001	5910000
AP008208_059	5900001	6010000
AP008208_060	6000001	6110000
AP008208_061	6100001	6210000
AP008208_062	6200001	6310000
AP008208_063	6300001	6410000
AP008208_064	6400001	6510000
AP008208_065	6500001	6610000
AP008208_066	6600001	6710000
AP008208_067	6700001	6810000
AP008208_068	6800001	6910000
AP008208_069	6900001	7010000
AP008208_070	7000001	7110000
AP008208_071	7100001	7210000
AP008208_072	7200001	7310000
AP008208_073	7300001	7410000
AP008208_074	7400001	7510000
AP008208_075	7500001	7610000
AP008208_076	7600001	7710000
AP008208_077	7700001	7810000
AP008208_078	7800001	7910000
AP008208_079	7900001	8010000
AP008208_080	8000001	8110000
AP008208_081	8100001	8210000

AP008208_082	8200001	8310000
AP008208_083	8300001	8410000
AP008208_084	8400001	8510000
AP008208_085	8500001	8610000
AP008208_086	8600001	8710000
AP008208_087	8700001	8810000
AP008208_088	8800001	8910000
AP008208_089	8900001	9010000
AP008208_090	9000001	9110000
AP008208_091	9100001	9210000
AP008208_092	9200001	9310000
AP008208_093	9300001	9410000
AP008208_094	9400001	9510000
AP008208_095	9500001	9610000
AP008208_096	9600001	9710000
AP008208_097	9700001	9810000
AP008208_098	9800001	9910000
AP008208_099	9900001	10010000
AP008208_100	10000001	10110000
AP008208_101	10100001	10210000
AP008208_102	10200001	10310000
AP008208_103	10300001	10410000
AP008208_104	10400001	10510000
AP008208_105	10500001	10610000
AP008208_106	10600001	10710000
AP008208_107	10700001	10810000
AP008208_108	10800001	10910000
AP008208_109	10900001	11010000
AP008208_110	11000001	11110000
AP008208_111	11100001	11210000
AP008208_112	11200001	11310000
AP008208_113	11300001	11410000
AP008208_114	11400001	11510000
AP008208_115	11500001	11610000
AP008208_116	11600001	11710000
AP008208_117	11700001	11810000
AP008208_118	11800001	11910000
AP008208_119	11900001	12010000
AP008208_120	12000001	12110000
AP008208_121	12100001	12210000
AP008208_122	12200001	12310000
AP008208_123	12300001	12410000
AP008208_124	12400001	12510000
AP008208_125	12500001	12610000
AP008208_126	12600001	12710000
AP008208_127	12700001	12810000
AP008208_128	12800001	12910000
AP008208_129	12900001	13010000
AP008208_130	13000001	13110000
AP008208_131	13100001	13210000
AP008208_132	13200001	13310000
AP008208_133	13300001	13410000
AP008208_134	13400001	13510000
AP008208_135	13500001	13610000
AP008208_136	13600001	13710000
AP008208_137	13700001	13810000
AP008208_138	13800001	13910000
AP008208_139	13900001	14010000
AP008208_140	14000001	14110000
AP008208_141	14100001	14210000
AP008208_142	14200001	14310000
AP008208_143	14300001	14410000
AP008208_144	14400001	14510000
AP008208_145	14500001	14610000
AP008208_146	14600001	14710000
AP008208_147	14700001	14810000
AP008208_148	14800001	14910000
AP008208_149	14900001	15010000
AP008208_150	15000001	15110000
AP008208_151	15100001	15210000
AP008208_152	15200001	15310000
AP008208_153	15300001	15410000
AP008208_154	15400001	15510000
AP008208_155	15500001	15610000
AP008208_156	15600001	15710000
AP008208_157	15700001	15810000
AP008208_158	15800001	15910000
AP008208_159	15900001	16010000
AP008208_160	16000001	16110000
AP008208_161	16100001	16210000
AP008208_162	16200001	16310000
AP008208_163	16300001	16410000
AP008208_164	16400001	16510000
AP008208_165	16500001	16610000
AP008208_166	16600001	16710000
AP008208_167	16700001	16810000
AP008208_168	16800001	16910000
AP008208_169	16900001	17010000
AP008208_170	17000001	17110000
AP008208_171	17100001	17210000
AP008208_172	17200001	17310000
AP008208_173	17300001	17410000
AP008208_174	17400001	17510000
AP008208_175	17500001	17610000
AP008208_176	17600001	17710000
AP008208_177	17700001	17810000
AP008208_178	17800001	17910000
AP008208_179	17900001	18010000
AP008208_180	18000001	18110000
AP008208_181	18100001	18210000
AP008208_182	18200001	18310000
AP008208_183	18300001	18410000
AP008208_184	18400001	18510000
AP008208_185	18500001	18610000
AP008208_186	18600001	18710000
AP008208_187	18700001	18810000
AP008208_188	18800001	18910000
AP008208_189	18900001	19010000
AP008208_190	19000001	19110000
AP008208_191	19100001	19210000
AP008208_192	19200001	19310000
AP008208_193	19300001	19410000
AP008208_194	19400001	19510000
AP008208_195	19500001	19610000
AP008208_196	19600001	19710000

Alignment Scores:

Pred. No.:	3.5e+04	Length:	110000
Score:	37.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	15	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AP008208_093 (1-110000)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||:::|||||||||||||
Db      87185 TCTTTGGTTACCTGGAACATT 87205
```

RESULT 75
AP008209_123/c
WPCOMMENT
Sequence split into 362 fragments LOCUS AP008209 Accession AP008209

Fragment Name	Begin	End
AP008209_000	1	110000
AP008209_001	100001	210000
AP008209_002	200001	310000
AP008209_003	300001	410000
AP008209_004	400001	510000
AP008209_005	500001	610000
AP008209_006	600001	710000
AP008209_007	700001	810000
AP008209_008	800001	910000
AP008209_009	900001	1010000
AP008209_010	1000001	1110000

AP008209_011	1100001	1210000
AP008209_012	1200001	1310000
AP008209_013	1300001	1410000
AP008209_014	1400001	1510000
AP008209_015	1500001	1610000
AP008209_016	1600001	1710000
AP008209_017	1700001	1810000
AP008209_018	1800001	1910000
AP008209_019	1900001	2010000
AP008209_020	2000001	2110000
AP008209_021	2100001	2210000
AP008209_022	2200001	2310000
AP008209_023	2300001	2410000
AP008209_024	2400001	2510000
AP008209_025	2500001	2610000
AP008209_026	2600001	2710000
AP008209_027	2700001	2810000
AP008209_028	2800001	2910000
AP008209_029	2900001	3010000
AP008209_030	3000001	3110000
AP008209_031	3100001	3210000
AP008209_032	3200001	3310000
AP008209_033	3300001	3410000
AP008209_034	3400001	3510000
AP008209_035	3500001	3610000
AP008209_036	3600001	3710000
AP008209_037	3700001	3810000
AP008209_038	3800001	3910000
AP008209_039	3900001	4010000
AP008209_040	4000001	4110000
AP008209_041	4100001	4210000
AP008209_042	4200001	4310000
AP008209_043	4300001	4410000
AP008209_044	4400001	4510000
AP008209_045	4500001	4610000
AP008209_046	4600001	4710000
AP008209_047	4700001	4810000
AP008209_048	4800001	4910000
AP008209_049	4900001	5010000
AP008209_050	5000001	5110000
AP008209_051	5100001	5210000
AP008209_052	5200001	5310000
AP008209_053	5300001	5410000
AP008209_054	5400001	5510000
AP008209_055	5500001	5610000
AP008209_056	5600001	5710000
AP008209_057	5700001	5810000
AP008209_058	5800001	5910000
AP008209_059	5900001	6010000
AP008209_060	6000001	6110000
AP008209_061	6100001	6210000
AP008209_062	6200001	6310000
AP008209_063	6300001	6410000
AP008209_064	6400001	6510000
AP008209_065	6500001	6610000
AP008209_066	6600001	6710000
AP008209_067	6700001	6810000
AP008209_068	6800001	6910000
AP008209_069	6900001	7010000
AP008209_070	7000001	7110000
AP008209_071	7100001	7210000
AP008209_072	7200001	7310000
AP008209_073	7300001	7410000
AP008209_074	7400001	7510000
AP008209_075	7500001	7610000
AP008209_076	7600001	7710000
AP008209_077	7700001	7810000
AP008209_078	7800001	7910000
AP008209_079	7900001	8010000
AP008209_080	8000001	8110000
AP008209_081	8100001	8210000
AP008209_082	8200001	8310000
AP008209_083	8300001	8410000
AP008209_084	8400001	8510000
AP008209_085	8500001	8610000
AP008209_086	8600001	8710000
AP008209_087	8700001	8810000
AP008209_088	8800001	8910000
AP008209_089	8900001	9010000
AP008209_090	9000001	9110000
AP008209_091	9100001	9210000
AP008209_092	9200001	9310000
AP008209_093	9300001	9410000
AP008209_094	9400001	9510000
AP008209_095	9500001	9610000
AP008209_096	9600001	9710000
AP008209_097	9700001	9810000
AP008209_098	9800001	9910000
AP008209_099	9900001	10010000
AP008209_100	10000001	10110000
AP008209_101	10100001	10210000
AP008209_102	10200001	10310000
AP008209_103	10300001	10410000
AP008209_104	10400001	10510000
AP008209_105	10500001	10610000
AP008209_106	10600001	10710000
AP008209_107	10700001	10810000
AP008209_108	10800001	10910000
AP008209_109	10900001	11010000
AP008209_110	11000001	11110000
AP008209_111	11100001	11210000
AP008209_112	11200001	11310000
AP008209_113	11300001	11410000
AP008209_114	11400001	11510000
AP008209_115	11500001	11610000
AP008209_116	11600001	11710000
AP008209_117	11700001	11810000
AP008209_118	11800001	11910000
AP008209_119	11900001	12010000
AP008209_120	12000001	12110000
AP008209_121	12100001	12210000
AP008209_122	12200001	12310000
AP008209_123	12300001	12410000
AP008209_124	12400001	12510000
AP008209_125	12500001	12610000
AP008209_126	12600001	12710000
AP008209_127	12700001	12810000
AP008209_128	12800001	12910000
AP008209_129	12900001	13010000
AP008209_130	13000001	13110000
AP008209_131	13100001	13210000
AP008209_132	13200001	13310000
AP008209_133	13300001	13410000
AP008209_134	13400001	13510000
AP008209_135	13500001	13610000
AP008209_136	13600001	13710000
AP008209_137	13700001	13810000
AP008209_138	13800001	13910000
AP008209_139	13900001	14010000
AP008209_140	14000001	14110000
AP008209_141	14100001	14210000
AP008209_142	14200001	14310000
AP008209_143	14300001	14410000
AP008209_144	14400001	14510000
AP008209_145	14500001	14610000
AP008209_146	14600001	14710000
AP008209_147	14700001	14810000
AP008209_148	14800001	14910000
AP008209_149	14900001	15010000
AP008209_150	15000001	15110000
AP008209_151	15100001	15210000
AP008209_152	15200001	15310000
AP008209_153	15300001	15410000
AP008209_154	15400001	15510000
AP008209_155	15500001	15610000
AP008209_156	15600001	15710000

AP008209_157	15700001	15810000
AP008209_158	15800001	15910000
AP008209_159	15900001	16010000
AP008209_160	16000001	16110000
AP008209_161	16100001	16210000
AP008209_162	16200001	16310000
AP008209_163	16300001	16410000
AP008209_164	16400001	16510000
AP008209_165	16500001	16610000
AP008209_166	16600001	16710000
AP008209_167	16700001	16810000
AP008209_168	16800001	16910000
AP008209_169	16900001	17010000
AP008209_170	17000001	17110000
AP008209_171	17100001	17210000
AP008209_172	17200001	17310000
AP008209_173	17300001	17410000
AP008209_174	17400001	17510000
AP008209_175	17500001	17610000
AP008209_176	17600001	17710000
AP008209_177	17700001	17810000
AP008209_178	17800001	17910000
AP008209_179	17900001	18010000
AP008209_180	18000001	18110000
AP008209_181	18100001	18210000
AP008209_182	18200001	18310000
AP008209_183	18300001	18410000
AP008209_184	18400001	18510000
AP008209_185	18500001	18610000
AP008209_186	18600001	18710000
AP008209_187	18700001	18810000
AP008209_188	18800001	18910000
AP008209_189	18900001	19010000
AP008209_190	19000001	19110000
AP008209_191	19100001	19210000
AP008209_192	19200001	19310000
AP008209_193	19300001	19410000
AP008209_194	19400001	19510000
AP008209_195	19500001	19610000
AP008209_196	19600001	19710000

Alignment Scores:

Pred. No.:	3.5e+04	Length:	110000
Score:	37.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	15	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AP008209_123 (1-110000)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db      55734 TCGTTAGTTACCTGGAACATT 55714
```

Search completed: March 10, 2006, 22:15:35
Job time : 524.957 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 00:16:23 ; Search time 452.849 Seconds
(without alignments)
1588.688 Million cell updates/sec

Title: US-10-757-745-2_COPY_54_140
Perfect score: 448
Sequence: 1 MERALNSYFEPPVEESALER.....IDGLDLNNLSERARGVCSYL 87

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170337_4279/app_query.fasta_1
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss08
-USER=US10757745_@CGN_1_1_2715_@runat_09032006_170337_4279 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	448	100.0	1296	3	US-09-925-299-170	Sequence 170, App
2	448	100.0	1296	3	US-09-925-299-170	Sequence 170, App
3	448	100.0	1920	9	US-10-757-745-1	Sequence 1, Appli
4	448	100.0	1940	9	US-10-783-271-28	Sequence 28, Appl
5	448	100.0	1948	5	US-10-037-270-889	Sequence 889, App
6	448	100.0	1948	6	US-10-117-722-889	Sequence 889, App
7	448	100.0	1948	9	US-10-122-851-889	Sequence 889, App
8	448	100.0	3152	3	US-09-981-353-88	Sequence 88, Appl
c 9	414	92.4	391	5	US-10-066-543-826	Sequence 826, App
10	332.5	74.2	1227	9	US-10-450-763-22059	Sequence 22059, A
11	288	64.3	606	4	US-09-925-065A-783537	Sequence 783537,
12	288	64.3	634	4	US-09-925-065A-783536	Sequence 783536,
13	211.5	47.2	1312	9	US-10-757-745-3	Sequence 3, Appli
c 14	114	25.4	725	8	US-10-363-345A-21055	Sequence 21055, A
15	114	25.4	725	8	US-10-363-345A-21056	Sequence 21056, A
c 16	114	25.4	725	9	US-10-363-483A-21055	Sequence 21055, A
17	114	25.4	725	9	US-10-363-483A-21056	Sequence 21056, A
18	101	22.5	60	3	US-09-908-975-14608	Sequence 14608, A
19	84.5	18.9	1589	9	US-10-450-763-23304	Sequence 23304, A
20	84.5	18.9	2228	9	US-10-450-763-23303	Sequence 23303, A
c 21	84.5	18.9	2717	9	US-10-450-763-18712	Sequence 18712, A
22	84.5	18.9	4002	9	US-10-450-763-11799	Sequence 11799, A
c 23	84.5	18.9	4890	9	US-10-450-763-23293	Sequence 23293, A
24	84.5	18.9	6448	9	US-10-450-763-18715	Sequence 18715, A
25	84.5	18.9	7197	9	US-10-450-763-23289	Sequence 23289, A
c 26	84.5	18.9	7556	9	US-10-450-763-18697	Sequence 18697, A
c 27	84.5	18.9	7731	9	US-10-450-763-23290	Sequence 23290, A
28	84.5	18.9	11742	9	US-10-450-763-23318	Sequence 23318, A
29	84.5	18.9	12036	9	US-10-450-763-11800	Sequence 11800, A
30	84.5	18.9	23157	9	US-10-450-763-11802	Sequence 11802, A
31	84.5	18.9	32313	9	US-10-450-763-18716	Sequence 18716, A
32	83.5	18.6	819	7	US-10-424-599-4169	Sequence 4169, Ap
c 33	79	17.6	1473	5	US-10-291-737-1	Sequence 1, Appli
c 34	79	17.6	1473	6	US-10-365-564-1	Sequence 1, Appli
c 35	79	17.6	1703	6	US-10-225-810-4	Sequence 4, Appli
36	78	17.4	480	6	US-10-264-049-1311	Sequence 1311, Ap
37	78	17.4	1142	5	US-10-208-408-33	Sequence 33, Appl
38	78	17.4	3672	6	US-10-133-937-64	Sequence 64, Appl
39	78	17.4	3672	6	US-10-465-572-13	Sequence 13, Appl
40	78	17.4	3672	6	US-10-172-118-390	Sequence 390, App
41	78	17.4	3672	6	US-10-159-563-64	Sequence 64, Appl
42	78	17.4	3672	7	US-10-342-887-390	Sequence 390, App
43	78	17.4	3672	9	US-10-848-755A-76	Sequence 76, Appl
44	78	17.4	4194	9	US-10-494-940-9	Sequence 9, Appli
45	78	17.4	5511	3	US-09-974-298-13	Sequence 13, Appl
46	78	17.4	5511	5	US-10-044-090-285	Sequence 285, App
47	78	17.4	5511	5	US-10-116-802-238	Sequence 238, App
48	78	17.4	5511	5	US-10-084-817-152	Sequence 152, App
c 49	77.5	17.3	1620	7	US-10-424-599-112946	Sequence 112946,
c 50	77	17.2	894	8	US-10-363-345A-21005	Sequence 21005, A
51	77	17.2	894	8	US-10-363-345A-21006	Sequence 21006, A
c 52	77	17.2	894	9	US-10-363-483A-21005	Sequence 21005, A
53	77	17.2	894	9	US-10-363-483A-21006	Sequence 21006, A
54	76	17.0	2933	7	US-10-437-963-99213	Sequence 99213, A
55	75.5	16.9	10506	5	US-10-153-219-1	Sequence 1, Appli
56	75.5	16.9	10506	7	US-10-719-372-1	Sequence 1, Appli
57	75.5	16.9	10506	8	US-10-743-975-1	Sequence 1, Appli
58	75.5	16.9	10617	5	US-10-153-219-14	Sequence 14, Appl
59	75.5	16.9	10617	7	US-10-719-372-14	Sequence 14, Appl
60	75.5	16.9	10617	8	US-10-743-975-14	Sequence 14, Appl
c 61	74	16.5	2490	9	US-10-450-763-8911	Sequence 8911, Ap
62	73.5	16.4	893	10	US-11-037-601-1	Sequence 1, Appli
c 63	72	16.1	1339	8	US-10-363-345A-20611	Sequence 20611, A
64	72	16.1	1339	8	US-10-363-345A-20612	Sequence 20612, A
c 65	72	16.1	1339	9	US-10-363-483A-20611	Sequence 20611, A
66	72	16.1	1339	9	US-10-363-483A-20612	Sequence 20612, A
67	72	16.1	2555	10	US-11-097-143-34261	Sequence 34261, A
c 68	72	16.1	2658	7	US-10-437-963-5051	Sequence 5051, Ap
69	72	16.1	3748	10	US-11-097-143-13687	Sequence 13687, A
70	72	16.1	4964	6	US-10-369-493-26910	Sequence 26910, A
71	71.5	16.0	921	7	US-10-451-337-3	Sequence 3, Appli
72	71.5	16.0	4320	10	US-11-097-143-7073	Sequence 7073, Ap
c 73	71.5	16.0	139613	7	US-10-367-094-90	Sequence 90, Appl
c 74	71	15.8	498	5	US-10-027-632-288729	Sequence 288729,
c 75	71	15.8	498	6	US-10-027-632-288729	Sequence 288729,

ALIGNMENTS

RESULT 1

```
US-09-925-299-170
; Sequence 170, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:                6.54e-54        Length:          1296
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-299-170 (1-1296)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 242

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 302

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 362

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 422

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        423 CGAGGGGTGTGTTCCTACTTA 443

RESULT 2
US-09-925-299-170
; Sequence 170, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:                6.54e-54        Length:          1296
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-299-170 (1-1296)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 242

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 302

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 362

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 422

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        423 CGAGGGGTGTGTTCCTACTTA 443

RESULT 3
US-10-757-745-1
; Sequence 1, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1627)..(1627)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1108)
```

```
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1849)..(1849)
;   OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

Alignment Scores:
Pred. No.:              1.13e-53        Length:         1920
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-757-745-1 (1-1920)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 238

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 298

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 358

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 418

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        419 CGAGGGGTGTGTTCCTACTTA 439

RESULT 4
US-10-783-271-28
; Sequence 28, Application US/10783271
; Publication No. US20050186577A1
; GENERAL INFORMATION:
;  APPLICANT: Veridex, LLC
;  APPLICANT:  Wang, Yixin
;  TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
;  FILE REFERENCE: VDX-5003 USNP
;  CURRENT APPLICATION NUMBER: US/10/783,271
;  CURRENT FILING DATE:  2004-02-20
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: human
US-10-783-271-28

Alignment Scores:
Pred. No.:              1.14e-53        Length:         1940
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-783-271-28 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 5
US-10-037-270-889
; Sequence 889, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/10/037,270
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-037-270-889

Alignment Scores:
Pred. No.:              1.15e-53        Length:         1948
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-037-270-889 (1-1948)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317
```

```
Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 6
US-10-117-722-889
; Sequence 889, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BCIP
;  CURRENT APPLICATION NUMBER: US/10/117,722
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-117-722-889

Alignment Scores:
Pred. No.:                1.15e-53        Length:          1948
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-117-722-889 (1-1948)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 7
US-10-122-851-889
; Sequence 889, Application US/10122851
; Publication No. US20050239060A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BDV3
;  CURRENT APPLICATION NUMBER: US/10/122,851
;  CURRENT FILING DATE:  2002-04-12
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-122-851-889

Alignment Scores:
Pred. No.:                1.15e-53        Length:          1948
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-122-851-889 (1-1948)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 8
US-09-981-353-88
; Sequence 88, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE: PA-0038 US
;  CURRENT APPLICATION NUMBER: US/09/981,353
```

```
;  CURRENT FILING DATE:  2001-10-11
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020160382A1 232992.1
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

Alignment Scores:
Pred. No.:              2.23e-53      Length:         3152
Score:                  448.00        Matches:        87
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-981-353-88 (1-3152)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1411 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 1470

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1471 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 1530

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1531 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 1590

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1591 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 1650

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db       1651 CGAGGGGTGTGTTCCTACTTA 1671

RESULT 9
US-10-066-543-826/c
; Sequence 826, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 826
;   LENGTH: 391
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 373
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-826

Alignment Scores:
Pred. No.:              1.01e-49      Length:         391
Score:                  414.00        Matches:        81
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            92.4%         Indels:         0
DB:                     5             Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-066-543-826 (1-391)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 227

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 167

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 107

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 47

Qy         81 Arg 81
              |||
Db         46 CGA 44

RESULT 10
US-10-450-763-22059
; Sequence 22059, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22059
;   LENGTH: 1227
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (84)..(203)
;   OTHER INFORMATION: 85% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=164.
US-10-450-763-22059

Alignment Scores:
Pred. No.:              2.81e-37      Length:         1227
Score:                  332.50        Matches:        84
Percent Similarity:     89.5%         Conservative:   1
Best Local Similarity:  88.4%         Mismatches:     2
Query Match:            74.2%         Indels:         8
DB:                     9             Gaps:           1
```

```
US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-22059 (1-1227)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSer-AlaLeuGluAr 20
              ||||||||||||||||||||||||||||||||||    ||||||||| ||||||||||||
Db        235 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGCGGAGGAGAGCAGCCTTGGAACG 294

Qy         20 gArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrTh 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 CCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 354

Qy         40 rAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn-GlySerM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        355 TGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGGCAGCA 414

Qy         60 etPheSerLeu-IleThrTrp-AsnIle-AspGlyLeuAsp---LeuAsnAsnLeu-Ser 77
              |||||||||||| ||||||||| |||||| ||||||:::      |||||||||||| |||
Db        415 TGTTCTCTCTTCATTACCTGGGAATATTTGATGGGATTAGGATCTTAAACAATCTGTTCA 474

Qy         78 Glu-ArgAlaArgGlyValCysSerTyrLeu 87
              ||| |||||||||||||||||||||||||||
Db        475 GAGGAGGGCTCGAGGGGTGTGTTCCTACTTA 505

RESULT 11
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783537

Alignment Scores:
Pred. No.:                2.8e-31       Length:          606
Score:                    288.00        Matches:         56
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              64.3%         Indels:          0
DB:                       4             Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-783537 (1-606)

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 473

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 533

Qy         72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 581

RESULT 12
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783536

Alignment Scores:
Pred. No.:                2.98e-31      Length:          634
Score:                    288.00        Matches:         56
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              64.3%         Indels:          0
DB:                       4             Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-783536 (1-634)

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 324

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 384

Qy         72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        385 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 432

RESULT 13
US-10-757-745-3
; Sequence 3, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Mus musculus
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (122)..(1234)
;   OTHER INFORMATION:
US-10-757-745-3

Alignment Scores:
Pred. No.:              8.88e-20        Length:         1312
Score:                  211.50          Matches:        45
Percent Similarity:     66.3%           Conservative:   12
Best Local Similarity:  52.3%           Mismatches:     28
Query Match:            47.2%           Indels:         1
DB:                     9               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-757-745-3 (1-1312)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::||||||::::::|||||||||   |||      :::           |||:::
Db        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||   |||       :::   :::   ||||||||||||||||||||:::       |||
Db        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436

Qy         42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
              :::|||       :::   |||||||||       |||       |||:::   |||
Db        437 ACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGCAGCACTATT 493

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
              |||   ||||||||||||||||||||||||||       ||||||   ||||||||||||
Db        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553

Qy         82 GlyValCysSerTyrLeu 87
              ||||||||||||   |||
Db        554 GGGGTGTGTTCCTGCCTA 571

RESULT 14
US-10-363-345A-21055/c
; Sequence 21055, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21055
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21055
US-10-363-345A-21055

Alignment Scores:
Pred. No.:              4.59e-06        Length:         725
Score:                  114.00          Matches:        19
Percent Similarity:     96.6%           Conservative:   9
Best Local Similarity:  65.5%           Mismatches:     1
Query Match:            25.4%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-345A-21055 (1-725)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::   ||||||||||||||||||||||||:::::::::::::||||||:::||||||
Db        155 CAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAAACGCCTTAAAACGCCGA 96

Qy         22 ProGluThrIleSerGluProLysThr 30
              |||:::|||||||||:::|||||||||
Db         95 CCTAAAACCATCTCTAAACCCAAAACC 69

RESULT 15
US-10-363-345A-21056
; Sequence 21056, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21056
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21056
US-10-363-345A-21056

Alignment Scores:
Pred. No.:              4.59e-06        Length:         725
Score:                  114.00          Matches:        19
Percent Similarity:     96.6%           Conservative:   9
Best Local Similarity:  65.5%           Mismatches:     1
Query Match:            25.4%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-345A-21056 (1-725)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::   ||||||||||||||||||||||||:::::::::::::||||||:::||||||
Db        571 CAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAAACGCCTTAAAACGCCGA 630

Qy         22 ProGluThrIleSerGluProLysThr 30
              |||:::|||||||||:::|||||||||
Db        631 CCTAAAACCATCTCTAAACCCAAAACC 657

RESULT 16
US-10-363-483A-21055/c
; Sequence 21055, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21055
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21055
US-10-363-483A-21055

Alignment Scores:
Pred. No.:              4.59e-06        Length:         725
Score:                  114.00          Matches:        19
Percent Similarity:     96.6%           Conservative:   9
Best Local Similarity:  65.5%           Mismatches:     1
```

```
Query Match:            25.4%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-483A-21055 (1-725)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::  ||||||||||||||||||||||||:::::::::::::|||||:::||||||
Db        155 CAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAAACGCCTTAAAACGCCGA 96

Qy         22 ProGluThrIleSerGluProLysThr 30
              |||:::||||||||||:::|||||||||
Db         95 CCTAAAACCATCTCTAAACCCAAAACC 69

RESULT 17
US-10-363-483A-21056
; Sequence 21056, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21056
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21056
US-10-363-483A-21056

Alignment Scores:
Pred. No.:              4.59e-06       Length:        725
Score:                  114.00         Matches:       19
Percent Similarity:     96.6%          Conservative:  9
Best Local Similarity:  65.5%          Mismatches:    1
Query Match:            25.4%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-483A-21056 (1-725)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::  ||||||||||||||||||||||||:::::::::::::|||||:::||||||
Db        571 CAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAAACGCCTTAAAACGCCGA 630

Qy         22 ProGluThrIleSerGluProLysThr 30
              |||:::||||||||||:::|||||||||
Db        631 CCTAAAACCATCTCTAAACCCAAAACC 657

RESULT 18
US-09-908-975-14608
; Sequence 14608, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14608
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-14608

Alignment Scores:
Pred. No.:              1.1e-05        Length:        60
Score:                  101.00         Matches:       19
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            22.5%          Indels:        0
DB:                     3              Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-908-975-14608 (1-60)

Qy         48 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGG 58

RESULT 19
US-10-450-763-23304
; Sequence 23304, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 23304
;   LENGTH: 1589
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (368)..(1204)
;   OTHER INFORMATION: 23% homologous to Staphylococcus simulans
;   OTHER INFORMATION: lysostaphin,accession number U66883,Smith-Waterman Score=141.
US-10-450-763-23304

Alignment Scores:
Pred. No.:              0.245          Length:        1589
Score:                  84.50          Matches:       26
Percent Similarity:     47.2%          Conservative:  16
Best Local Similarity:  29.2%          Mismatches:    26
Query Match:            18.9%          Indels:        21
DB:                     9              Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-23304 (1-1589)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||           |||||||::   |||   ||||||
Db        638 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 685

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                          |||      |||:::||||||||||||:::|||            |||
Db        686 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 745

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                    :::      :::|||   :::   :::|||               |||   :::
Db        746 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 790
```

```
Qy          65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
               |||       :::   :::|||:::|||                  :::
Db         791 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 850

Qy          78 GluArgAlaArgGlyValCysSerTyr 86
                  :::|||   |||:::   :::|||
Db         851 TCTAAAGCAGAGGGTATTAAGAACTAT 877
```

RESULT 20
US-10-450-763-23303
; Sequence 23303, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23303
; LENGTH: 2228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (758)..(1594)
; OTHER INFORMATION: 23% homologous to Staphylococcus simulans
; OTHER INFORMATION: lysostaphin,accession number U66883,Smith-Waterman Score=141.
US-10-450-763-23303

Alignment Scores:

Pred. No.:	0.39	Length:	2228
Score:	84.50	Matches:	26
Percent Similarity:	47.2%	Conservative:	16
Best Local Similarity:	29.2%	Mismatches:	26
Query Match:	18.9%	Indels:	21
DB:	9	Gaps:	5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-23303 (1-2228)

```
Qy          10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
               ||||||      :::||||||             |||||:::   |||   ||||||
Db        1028 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 1075

Qy          29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                         |||      |||:::|||||||||||||:::|||             |||
Db        1076 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 1135

Qy          45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                  :::      :::|||   :::   :::|||                |||   :::
Db        1136 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 1180

Qy          65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
               |||       :::   :::|||:::|||                  :::
Db        1181 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 1240

Qy          78 GluArgAlaArgGlyValCysSerTyr 86
                  :::|||   |||:::   :::|||
Db        1241 TCTAAAGCAGAGGGTATTAAGAACTAT 1267
```

RESULT 21
US-10-450-763-18712/c
; Sequence 18712, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 18712
; LENGTH: 2717
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2379)..(1492)
; OTHER INFORMATION: 24% homologous to Staphylococcus simulans
; OTHER INFORMATION: lysostaphin,accession number U66883,Smith-Waterman Score=143.
US-10-450-763-18712

Alignment Scores:

Pred. No.:	0.513	Length:	2717
Score:	84.50	Matches:	26
Percent Similarity:	47.2%	Conservative:	16
Best Local Similarity:	29.2%	Mismatches:	26
Query Match:	18.9%	Indels:	21
DB:	9	Gaps:	5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-18712 (1-2717)

```
Qy          10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
               ||||||      :::||||||             |||||:::   |||   ||||||
Db        2109 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 2062

Qy          29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                         |||      |||:::|||||||||||||:::|||             |||
Db        2061 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 2002

Qy          45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                  :::      :::|||   :::   :::|||                |||   :::
Db        2001 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 1957

Qy          65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
               |||       :::   :::|||:::|||                  :::
Db        1956 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 1897

Qy          78 GluArgAlaArgGlyValCysSerTyr 86
                  :::|||   |||:::   :::|||
Db        1896 TCTAAAGCAGAGGGTATTAAGAACTAT 1870
```

RESULT 22
US-10-450-763-11799
; Sequence 11799, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 11799
; LENGTH: 4002

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2)..(370)
;   OTHER INFORMATION: 98% homologous to Cricetulus griseus alpha-tubulin
;   OTHER INFORMATION: III,accession number M12329,Smith-Waterman Score=644.
US-10-450-763-11799

Alignment Scores:
Pred. No.:                0.876          Length:        4002
Score:                    84.50          Matches:       26
Percent Similarity:       47.2%          Conservative:  16
Best Local Similarity:    29.2%          Mismatches:    26
Query Match:              18.9%          Indels:        21
DB:                       9              Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-11799 (1-4002)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||            ||||||:::   |||   ||||||
Db       1435 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 1482

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||     |||:::||||||||||||:::|||             |||
Db       1483 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 1542

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||               |||   :::
Db       1543 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 1587

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||     :::   :::|||:::|||                 :::
Db       1588 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 1647

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       1648 TCTAAAGCAGAGGGTATTAAGAACTAT 1674

RESULT 23
US-10-450-763-23293/c
; Sequence 23293, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 23293
;   LENGTH: 4890
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2451)..(1615)
;   OTHER INFORMATION: 23% homologous to Staphylococcus simulans
;   OTHER INFORMATION: lysostaphin,accession number U66883,Smith-Waterman Score=141.
US-10-450-763-23293

Alignment Scores:
Pred. No.:                1.16           Length:        4890
Score:                    84.50          Matches:       26
Percent Similarity:       47.2%          Conservative:  16
Best Local Similarity:    29.2%          Mismatches:    26
Query Match:              18.9%          Indels:        21
DB:                       9              Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-23293 (1-4890)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||            ||||||:::   |||   ||||||
Db       2181 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 2134

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||     |||:::||||||||||||:::|||             |||
Db       2133 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 2074

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||               |||   :::
Db       2073 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 2029

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||     :::   :::|||:::|||                 :::
Db       2028 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 1969

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       1968 TCTAAAGCAGAGGGTATTAAGAACTAT 1942

RESULT 24
US-10-450-763-18715
; Sequence 18715, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 18715
;   LENGTH: 6448
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (4037)..(4363)
;   OTHER INFORMATION: 91% homologous to Homo sapiens putative p150,accession number
;   OTHER INFORMATION: U93572,Smith-Waterman Score=499.
US-10-450-763-18715

Alignment Scores:
Pred. No.:                1.69           Length:        6448
Score:                    84.50          Matches:       26
Percent Similarity:       47.2%          Conservative:  16
Best Local Similarity:    29.2%          Mismatches:    26
Query Match:              18.9%          Indels:        21
DB:                       9              Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-18715 (1-6448)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||            ||||||:::   |||   ||||||
Db       1028 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 1075

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||     |||:::||||||||||||:::|||             |||
Db       1076 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 1135

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
```

```
                  :::      :::|||   :::   :::|||               |||   :::
Db       1136 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 1180

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                :::
Db       1181 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 1240

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       1241 TCTAAAGCAGAGGGTATTAAGAACTAT 1267

RESULT 25
US-10-450-763-23289
; Sequence 23289, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 23289
;   LENGTH: 7197
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2650)..(3408)
;   OTHER INFORMATION: 99% homologous to Bacteriophage N4 single-stranded DNA-
;   OTHER INFORMATION: binding protein,accession number U29728,Smith-Waterman Score=1298
US-10-450-763-23289

Alignment Scores:
Pred. No.:                1.97          Length:         7197
Score:                    84.50         Matches:        26
Percent Similarity:       47.2%         Conservative:   16
Best Local Similarity:    29.2%         Mismatches:     26
Query Match:              18.9%         Indels:         21
DB:                       9             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-23289 (1-7197)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||            |||||||:::   |||   ||||||
Db       5158 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 5205

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||       |||:::|||||||||||||:::|||              |||
Db       5206 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 5265

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||                |||   :::
Db       5266 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 5310

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                :::
Db       5311 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 5370

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       5371 TCTAAAGCAGAGGGTATTAAGAACTAT 5397

RESULT 26
US-10-450-763-18697/c
; Sequence 18697, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 18697
;   LENGTH: 7556
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (6132)..(3832)
;   OTHER INFORMATION: 20% homologous to Staphylococcus aureus FmtB,accession number
;   OTHER INFORMATION: AB015223,Smith-Waterman Score=154.
US-10-450-763-18697

Alignment Scores:
Pred. No.:                2.11          Length:         7556
Score:                    84.50         Matches:        26
Percent Similarity:       47.2%         Conservative:   16
Best Local Similarity:    29.2%         Mismatches:     26
Query Match:              18.9%         Indels:         21
DB:                       9             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-18697 (1-7556)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||            |||||||:::   |||   ||||||
Db       4740 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 4693

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||       |||:::|||||||||||||:::|||              |||
Db       4692 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 4633

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||                |||   :::
Db       4632 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 4588

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                :::
Db       4587 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 4528

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       4527 TCTAAAGCAGAGGGTATTAAGAACTAT 4501

RESULT 27
US-10-450-763-23290/c
; Sequence 23290, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
```

```
;  SOFTWARE: Custom
; SEQ ID NO 23290
;   LENGTH: 7731
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (792)..(4)
;   OTHER INFORMATION: 90% homologous to Homo sapiens putative p150,accession number
;   OTHER INFORMATION: U93569,Smith-Waterman Score=1271.
US-10-450-763-23290

Alignment Scores:
Pred. No.:              2.17          Length:         7731
Score:                  84.50         Matches:        26
Percent Similarity:     47.2%         Conservative:   16
Best Local Similarity:  29.2%         Mismatches:     26
Query Match:            18.9%         Indels:         21
DB:                     9             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-23290 (1-7731)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||           ||||||:::   |||   ||||||
Db       3801 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 3754

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::|||            |||
Db       3753 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 3694

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||                ||| :::
Db       3693 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 3649

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                 :::
Db       3648 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 3589

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                :::|||   |||:::   :::|||
Db       3588 TCTAAAGCAGAGGGTATTAAGAACTAT 3562

RESULT 28
US-10-450-763-23318
; Sequence 23318, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 23318
;   LENGTH: 11742
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (5518)..(6258)
;   OTHER INFORMATION: 27% homologous to Bacteriophage T4 DNA helicase,accession
;   OTHER INFORMATION: number M93048,Smith-Waterman Score=171.
US-10-450-763-23318

Alignment Scores:
Pred. No.:              3.87          Length:         11742
Score:                  84.50         Matches:        26
Percent Similarity:     47.2%         Conservative:   16
Best Local Similarity:  29.2%         Mismatches:     26
Query Match:            18.9%         Indels:         21
DB:                     9             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-23318 (1-11742)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||           ||||||:::   |||   ||||||
Db      10954 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 11001

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::|||            |||
Db      11002 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 11061

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||                ||| :::
Db      11062 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 11106

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                 :::
Db      11107 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 11166

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                :::|||   |||:::   :::|||
Db      11167 TCTAAAGCAGAGGGTATTAAGAACTAT 11193

RESULT 29
US-10-450-763-11800
; Sequence 11800, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 11800
;   LENGTH: 12036
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1544)..(714)
;   OTHER INFORMATION: 83% homologous to Xenopus laevis alpha-tubulin,accession
;   OTHER INFORMATION: number X07045,Smith-Waterman Score=1139.
US-10-450-763-11800

Alignment Scores:
Pred. No.:              4.01          Length:         12036
Score:                  84.50         Matches:        26
Percent Similarity:     47.2%         Conservative:   16
Best Local Similarity:  29.2%         Mismatches:     26
Query Match:            18.9%         Indels:         21
DB:                     9             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-11800 (1-12036)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||           ||||||:::   |||   ||||||
Db       4558 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 4605

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::|||            |||
```

```
Db       4606 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 4665

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              :::      :::|||   :::   :::|||               |||   :::
Db       4666 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 4710

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       4711 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 4770

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                :::|||  |||:::  :::|||
Db       4771 TCTAAAGCAGAGGGTATTAAGAACTAT 4797

RESULT 30
US-10-450-763-11802
; Sequence 11802, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 11802
;   LENGTH: 23157
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (489)..(3767)
;   OTHER INFORMATION: 100% homologous to Homo sapiens bcr,accession number
;   OTHER INFORMATION: X02596,Smith-Waterman Score=5702.
US-10-450-763-11802

Alignment Scores:
Pred. No.:              9.89          Length:        23157
Score:                  84.50         Matches:       26
Percent Similarity:     47.2%         Conservative:  16
Best Local Similarity:  29.2%         Mismatches:    26
Query Match:            18.9%         Indels:        21
DB:                     9             Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-11802 (1-23157)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||           ||||||:::   |||   ||||||
Db      15505 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 15552

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::|||||||||||||:::|||            |||
Db      15553 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 15612

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              :::      :::|||   :::   :::|||               |||   :::
Db      15613 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 15657

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db      15658 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 15717

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                :::|||  |||:::  :::|||
Db      15718 TCTAAAGCAGAGGGTATTAAGAACTAT 15744

RESULT 31
US-10-450-763-18716
; Sequence 18716, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 18716
;   LENGTH: 32313
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (19866)..(24821)
;   OTHER INFORMATION: 15% homologous to Streptococcus cristatus srpA,accession
;   OTHER INFORMATION: number U96166,Smith-Waterman Score=184.
US-10-450-763-18716

Alignment Scores:
Pred. No.:              15.7          Length:        32313
Score:                  84.50         Matches:       26
Percent Similarity:     47.2%         Conservative:  16
Best Local Similarity:  29.2%         Mismatches:    26
Query Match:            18.9%         Indels:        21
DB:                     9             Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-18716 (1-32313)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||           ||||||:::   |||   ||||||
Db      24660 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 24707

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::|||||||||||||:::|||            |||
Db      24708 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 24767

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              :::      :::|||   :::   :::|||               |||   :::
Db      24768 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 24812

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db      24813 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 24872

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                :::|||  |||:::  :::|||
Db      24873 TCTAAAGCAGAGGGTATTAAGAACTAT 24899

RESULT 32
US-10-424-599-4169
; Sequence 4169, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
```

```
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 4169
;   LENGTH: 819
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_103771C.1
US-10-424-599-4169

Alignment Scores:
Pred. No.:                0.137          Length:         819
Score:                    83.50          Matches:        20
Percent Similarity:       49.2%          Conservative:   10
Best Local Similarity:    32.8%          Mismatches:     26
Query Match:              18.6%          Indels:         5
DB:                       7              Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-424-599-4169 (1-819)

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||   |||:::   |||||||||:::      :::||||||:::      :::
Db        385 CCTGCTACCTTA---GAGCCTAAAAGTCTGGCATCTATCACCAATAAAAGAAATGACGAG 441

Qy         42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
                 |||                |||||||||     ::::::   |||:::
Db        442 GATGACACAGAA-----------AGTGAGGATAGGTCTGAGGATGAAGGAGCTGCAGGA 489

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
                       |||||||||   :::      |||:::||||||   |||
Db        490 CCACAAAAAACATGGAATAAGAATGCTACTGATATGAATAACCACAGCTCTGATGATGAA 549

Qy         82 Gly 82
              |||
Db        550 GGA 552

RESULT 33
US-10-291-737-1/c
; Sequence 1, Application US/10291737
; Publication No. US20030087299A1
; GENERAL INFORMATION:
;  APPLICANT: WEBSTER, Marion et al.
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL001017-CON
;  CURRENT APPLICATION NUMBER: US/10/291,737
;  CURRENT FILING DATE:  2002-11-12
;  PRIOR APPLICATION NUMBER: US 09/822,859
;  PRIOR FILING DATE: 2001-04-02
;  PRIOR APPLICATION NUMBER: US 60/254,588
;  PRIOR FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1473
;   TYPE: DNA
;   ORGANISM: Human
US-10-291-737-1

Alignment Scores:
Pred. No.:                1.37           Length:         1473
Score:                    79.00          Matches:        23
Percent Similarity:       43.3%          Conservative:   6
Best Local Similarity:    34.3%          Mismatches:     36
Query Match:              17.6%          Indels:         2
DB:                       5              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-291-737-1 (1-1473)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                 |||      |||              |||
Db       1289 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1230

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||          ::::::|||      |||   |||      |||      |||   |||
Db       1229 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1176

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::              |||:::      :::      ||||||   |||
Db       1175 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1116

Qy         61 PheSerLeuIleThrTrpAsn 67
                 |||      |||||||||
Db       1115 ATCTCAGCTGGGACGTGGAAC 1095

RESULT 34
US-10-365-564-1/c
; Sequence 1, Application US/10365564
; Publication No. US20030143623A1
; GENERAL INFORMATION:
;  APPLICANT: WEBSTER, Marion et al.
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL001017-CIPCON
;  CURRENT APPLICATION NUMBER: US/10/365,564
;  CURRENT FILING DATE:  2003-02-13
;  PRIOR APPLICATION NUMBER: US 09/822,859
;  PRIOR FILING DATE: 2001-04-02
;  PRIOR APPLICATION NUMBER: US 60/254,588
;  PRIOR FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1473
;   TYPE: DNA
;   ORGANISM: Human
US-10-365-564-1

Alignment Scores:
Pred. No.:                1.37           Length:         1473
Score:                    79.00          Matches:        23
Percent Similarity:       43.3%          Conservative:   6
Best Local Similarity:    34.3%          Mismatches:     36
Query Match:              17.6%          Indels:         2
DB:                       6              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-365-564-1 (1-1473)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                 |||      |||              |||
Db       1289 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1230

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||          ::::::|||      |||   |||      |||      |||   |||
Db       1229 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1176

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::              |||:::      :::      ||||||   |||
Db       1175 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1116

Qy         61 PheSerLeuIleThrTrpAsn 67
                 |||      |||||||||
Db       1115 ATCTCAGCTGGGACGTGGAAC 1095

RESULT 35
US-10-225-810-4/c
; Sequence 4, Application US/10225810
; Publication No. US20030157512A1
; GENERAL INFORMATION:
;  APPLICANT: Bermingham, Jr., John R.
;  TITLE OF INVENTION: Tramdorins and Methods of Using Tramdorin
;  FILE REFERENCE: McLaugh-07165
;  CURRENT APPLICATION NUMBER: US/10/225,810
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 76
```

```
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
;   LENGTH: 1703
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-810-4

Alignment Scores:
Pred. No.:                1.67          Length:          1703
Score:                    79.00         Matches:         23
Percent Similarity:       43.3%         Conservative:    6
Best Local Similarity:    34.3%         Mismatches:      36
Query Match:              17.6%         Indels:          2
DB:                       6             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-225-810-4 (1-1703)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               |||   ||||||                 |||      |||             |||
Db        1483 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1424

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               |||          ::::::|||      |||   |||      |||      |||   |||
Db        1423 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1370

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               |||    :::|||:::           |||:::      :::      ||||||   |||
Db        1369 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1310

Qy          61 PheSerLeuIleThrTrpAsn 67
                 |||      |||||||||
Db        1309 ATCTCAGCTGGGACGTGGAAC 1289

RESULT 36
US-10-264-049-1311
; Sequence 1311, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA133P1
;  CURRENT APPLICATION NUMBER: US/10/264,049
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/18569
;  PRIOR FILING DATE: 2001-06-07
;  PRIOR APPLICATION NUMBER: US 60/209,467
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 4360
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1311
;   LENGTH: 480
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-264-049-1311

Alignment Scores:
Pred. No.:                0.406         Length:          480
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       6             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-264-049-1311 (1-480)

Qy           7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
               |||  |||   ||||||:::           ||||||:::    |||      |||
Db         113 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 172

Qy          26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
                :::      |||   |||   |||      |||:::|||         |||||||||
Db         173 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 232

Qy          46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                     |||   ||||||  |||          |||   ||||||:::|||:::
Db         233 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 292

Qy          64 eThrTrpAsnIleAspGlyLeu 71
                  |||   :::      |||
Db         293 TGTCTGGACTTTGAGCAGGCTT 314

RESULT 37
US-10-208-408-33
; Sequence 33, Application US/10208408
; Publication No. US20030096272A1
; GENERAL INFORMATION:
;  APPLICANT: Schebye, Xiao Min
;  TITLE OF INVENTION: GENES REGULATED BY PEROXISOME PROLIFERATOR-ACTIVATED RECEPTOR
;  FILE REFERENCE: PA-0048-1 US
;  CURRENT APPLICATION NUMBER: US/10/208,408
;  CURRENT FILING DATE:  2002-07-29
;  PRIOR APPLICATION NUMBER: 60/308,868
;  PRIOR FILING DATE: 2001-07-30
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: PERL Program
; SEQ ID NO 33
;   LENGTH: 1142
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030096272A1 1453266.4
US-10-208-408-33

Alignment Scores:
Pred. No.:                1.34          Length:          1142
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       5             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-208-408-33 (1-1142)

Qy           7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
               |||  |||   ||||||:::           ||||||:::    |||      |||
Db         836 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 895

Qy          26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
                :::      |||   |||   |||      |||:::|||         |||||||||
Db         896 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 955

Qy          46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                     |||   ||||||  |||          |||   ||||||:::|||:::
Db         956 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1015

Qy          64 eThrTrpAsnIleAspGlyLeu 71
                  |||   :::      |||
Db        1016 TGTCTGGACTTTGAGCAGGCTT 1037

RESULT 38
US-10-133-937-64
; Sequence 64, Application US/10133937
; Publication No. US20030207278A1
; GENERAL INFORMATION:
;  APPLICANT: Khan, Javed
;  APPLICANT:  Ringner, Markus
;  APPLICANT:  Peterson, Carsten
;  APPLICANT:  Meltzer, Paul
;  TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
;  TITLE OF INVENTION:  DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
;  TITLE OF INVENTION:  OTHER BIOLOGICAL STATES
;  FILE REFERENCE: 11613.56US01
;  CURRENT APPLICATION NUMBER: US/10/133,937
;  CURRENT FILING DATE:  2002-11-04
;  NUMBER OF SEQ ID NOS: 99
```

```
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
;   LENGTH: 3672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-133-937-64

Alignment Scores:
Pred. No.:              6.74          Length:         3672
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     6             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-133-937-64 (1-3672)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::     |||   |||   |||      |||:::|||         |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||          |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::     |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 39
US-10-465-572-13
; Sequence 13, Application US/10465572
; Publication No. US20030207840A1
; GENERAL INFORMATION:
;  APPLICANT: Riggins, Gregory
;  APPLICANT:  Lal, Anita
;  TITLE OF INVENTION: GENES INDUCED BY HYPOXIA
;  FILE REFERENCE: 000250.00012
;  CURRENT APPLICATION NUMBER: US/10/465,572
;  CURRENT FILING DATE:  2003-06-20
;  PRIOR APPLICATION NUMBER: US/10/201,642
;  PRIOR FILING DATE: 2002-07-24
;  PRIOR APPLICATION NUMBER: 60/307,600
;  PRIOR FILING DATE: 2001-07-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;   LENGTH: 3672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-465-572-13

Alignment Scores:
Pred. No.:              6.74          Length:         3672
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     6             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-465-572-13 (1-3672)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::     |||   |||   |||      |||:::|||         |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||          |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::     |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 40
US-10-172-118-390
; Sequence 390, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Chris
;  APPLICANT:  Van 't Veer, Laura
;  APPLICANT:  Van de Vijver, Marc
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-175-999
;  CURRENT APPLICATION NUMBER: US/10/172,118
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: 60/380,770
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 390
;   LENGTH: 3672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: L27560
;   DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-390

Alignment Scores:
Pred. No.:              6.74          Length:         3672
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     6             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-172-118-390 (1-3672)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::     |||   |||   |||      |||:::|||         |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||          |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::     |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 41
US-10-159-563-64
; Sequence 64, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
;  APPLICANT: Khan, Javed
;  APPLICANT:  Ringner, Markus
;  APPLICANT:  Peterson, Carsten
;  APPLICANT:  Meltzer, Paul
```

```
;  TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
;  TITLE OF INVENTION:  DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
;  FILE REFERENCE: 11613.56USI1
;  CURRENT APPLICATION NUMBER: US/10/159,563
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US 10/133,937
;  PRIOR FILING DATE: 2002-04-25
;  NUMBER OF SEQ ID NOS: 444
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
;   LENGTH: 3672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-159-563-64

Alignment Scores:
Pred. No.:              6.74          Length:        3672
Score:                  78.00         Matches:       26
Percent Similarity:     48.5%         Conservative:  7
Best Local Similarity:  38.2%         Mismatches:    32
Query Match:            17.4%         Indels:        3
DB:                     6             Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-159-563-64 (1-3672)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||          |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||          |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 42
US-10-342-887-390
; Sequence 390, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter S.
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Christopher J.
;  APPLICANT:  Van 't Veer, Laura Johanna
;  APPLICANT:  Van de Vijver, Marc J.
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-188-999
;  CURRENT APPLICATION NUMBER: US/10/342,887
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: 60/298,918
;  PRIOR FILING DATE: 2001-06-18
;  PRIOR APPLICATION NUMBER: 60/380,710
;  PRIOR FILING DATE: 2002-05-14
;  PRIOR APPLICATION NUMBER: 10/172,118
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 390
;   LENGTH: 3672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-390

Alignment Scores:
Pred. No.:              6.74          Length:        3672
Score:                  78.00         Matches:       26
Percent Similarity:     48.5%         Conservative:  7
Best Local Similarity:  38.2%         Mismatches:    32
Query Match:            17.4%         Indels:        3
DB:                     7             Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-342-887-390 (1-3672)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||          |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||          |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 43
US-10-848-755A-76
; Sequence 76, Application US/10848755A
; Publication No. US20050054826A1
; GENERAL INFORMATION:
;  APPLICANT: Mao, Mao
;  TITLE OF INVENTION: HUMAN DIAPHANOUS-3 GENE AND METHODS OF USE THEREFOR
;  FILE REFERENCE: 9301-196-999
;  CURRENT APPLICATION NUMBER: US/10/848,755A
;  CURRENT FILING DATE:  2004-05-18
;  PRIOR APPLICATION NUMBER: 60/471,842
;  PRIOR FILING DATE: 2003-05-11
;  NUMBER OF SEQ ID NOS: 275
;  SOFTWARE: PatentIn version 3.2  CAM: 301891-999188
; SEQ ID NO 76
;   LENGTH: 3672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-848-755A-76

Alignment Scores:
Pred. No.:              6.74          Length:        3672
Score:                  78.00         Matches:       26
Percent Similarity:     48.5%         Conservative:  7
Best Local Similarity:  38.2%         Mismatches:    32
Query Match:            17.4%         Indels:        3
DB:                     9             Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-848-755A-76 (1-3672)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||          |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||          |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 44
US-10-494-940-9
; Sequence 9, Application US/10494940
```

```
; Publication No. US20050069886A1
; GENERAL INFORMATION:
;  APPLICANT: OriGene Technologies, Inc
;  TITLE OF INVENTION: REGULATED PROSTATE CANCER GENES
;  FILE REFERENCE: OGT 9U 803 PCT
;  CURRENT APPLICATION NUMBER: US/10/494,940
;  CURRENT FILING DATE:  2004-05-07
;  PRIOR APPLICATION NUMBER: 60/331,042
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: 60/331,041
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: 60/340,251
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: 60/344,791
;  PRIOR FILING DATE: 2002-01-07
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 4194
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1588)..(2019)
;   OTHER INFORMATION:
US-10-494-940-9

Alignment Scores:
Pred. No.:              8.1          Length:        4194
Score:                  78.00        Matches:       26
Percent Similarity:     48.5%        Conservative:  7
Best Local Similarity:  38.2%        Mismatches:    32
Query Match:            17.4%        Indels:        3
DB:                     9            Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-494-940-9 (1-4194)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::    |||       |||
Db       2254 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 2313

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||       |||:::|||         |||||||||
Db       2314 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 2373

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||          |||   ||||||:::|||:::
Db       2374 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 2433

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       2434 TGTCTGGACTTTGAGCAGGCTT 2455

RESULT 45
US-09-974-298-13
; Sequence 13, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Huei-Mei
;  TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
;  FILE REFERENCE: PA-0037 P
;  CURRENT APPLICATION NUMBER: US/09/974,298
;  CURRENT FILING DATE:  2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,331
;  PRIOR FILING DATE: 2000-05-10
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 13
;   LENGTH: 5511
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020156263A1 1075592.6
;   NAME/KEY: unsure
;   LOCATION: 5496
;   OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-13

Alignment Scores:
Pred. No.:              11.8         Length:        5511
Score:                  78.00        Matches:       26
Percent Similarity:     48.5%        Conservative:  7
Best Local Similarity:  38.2%        Mismatches:    32
Query Match:            17.4%        Indels:        3
DB:                     3            Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-974-298-13 (1-5511)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::    |||       |||
Db       4092 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4151

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||       |||:::|||         |||||||||
Db       4152 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4211

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||          |||   ||||||:::|||:::
Db       4212 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4271

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       4272 TGTCTGGACTTTGAGCAGGCTT 4293

RESULT 46
US-10-044-090-285
; Sequence 285, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 285
;   LENGTH: 5511
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 1075592.6
;   NAME/KEY: unsure
;   LOCATION: 5496
;   OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-285

Alignment Scores:
Pred. No.:              11.8         Length:        5511
Score:                  78.00        Matches:       26
Percent Similarity:     48.5%        Conservative:  7
Best Local Similarity:  38.2%        Mismatches:    32
Query Match:            17.4%        Indels:        3
DB:                     5            Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-044-090-285 (1-5511)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::    |||       |||
Db       4092 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4151

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||       |||:::|||         |||||||||
Db       4152 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4211
```

```
Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                     |||   ||||||   |||         |||   ||||||:::|||:::
Db       4212 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4271

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                ||| :::     |||
Db       4272 TGTCTGGACTTTGAGCAGGCTT 4293

RESULT 47
US-10-116-802-238
; Sequence 238, Application US/10116802
; Publication No. US20030065157A1
; GENERAL INFORMATION:
;  APPLICANT: Amy Lasek
;  TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
;  FILE REFERENCE: PA-0045 US
;  CURRENT APPLICATION NUMBER: US/10/116,802
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 60/281,593
;  PRIOR FILING DATE: 2001-04-04
;  NUMBER OF SEQ ID NOS: 519
;  SOFTWARE: PERL Program
; SEQ ID NO 238
;   LENGTH: 5511
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 1075592.6
;   NAME/KEY: unsure
;   LOCATION: 5496
;   OTHER INFORMATION: a, t, c, g, or other
US-10-116-802-238

Alignment Scores:
Pred. No.:              11.8          Length:         5511
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     5             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-116-802-238 (1-5511)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::   |||       |||
Db       4092 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4151

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::     |||   |||   |||     |||:::|||         |||||||||
Db       4152 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4211

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||         |||   ||||||:::|||:::
Db       4212 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4271

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                ||| :::     |||
Db       4272 TGTCTGGACTTTGAGCAGGCTT 4293

RESULT 48
US-10-084-817-152
; Sequence 152, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
;  APPLICANT: Susan Stuart
;  APPLICANT:  Jed G. Nuchtern
;  APPLICANT:  Sharon E. Plon
;  APPLICANT:  Jason M. Shohet
;  TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
;  FILE REFERENCE: PA-0046 US
;  CURRENT APPLICATION NUMBER: US/10/084,817
;  CURRENT FILING DATE:  2002-02-25
;  PRIOR APPLICATION NUMBER: 60/270,784
;  PRIOR FILING DATE: 2001-02-23
;  NUMBER OF SEQ ID NOS: 365
;  SOFTWARE: PERL Program
; SEQ ID NO 152
;   LENGTH: 5511
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030119009A1 1075592.6
;   NAME/KEY: unsure
;   LOCATION: 5496
;   OTHER INFORMATION: a, t, c, g, or other
US-10-084-817-152

Alignment Scores:
Pred. No.:              11.8          Length:         5511
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     5             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-084-817-152 (1-5511)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::   |||       |||
Db       4092 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4151

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::     |||   |||   |||     |||:::|||         |||||||||
Db       4152 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4211

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||         |||   ||||||:::|||:::
Db       4212 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4271

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                ||| :::     |||
Db       4272 TGTCTGGACTTTGAGCAGGCTT 4293

RESULT 49
US-10-424-599-112946/c
; Sequence 112946, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 112946
;   LENGTH: 1620
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_72999C.1
US-10-424-599-112946

Alignment Scores:
Pred. No.:              2.57          Length:         1620
Score:                  77.50         Matches:        25
Percent Similarity:     47.4%         Conservative:   11
Best Local Similarity:  32.9%         Mismatches:     35
Query Match:            17.3%         Indels:         5
DB:                     7             Gaps:           2
```

```
US-10-757-745-2_COPY_54_140 (1-87) x US-10-424-599-112946 (1-1620)

Qy          12 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
               |||       :::|||   :::   |||  |||:::||||||                 |||
Db        1194 CCCTCTTCCCAATCACCCATGCGGCGACCACCAAAAACAATTTGTAGCCTTCCATGATAT 1135

Qy          32 ValAspLeuThrAsnGluGluThrThrAspSerThrThr----------SerLysIle 47
               :::|||                          ||||||:::          |||   :::
Db        1134 TCCGAGCTCAGGGCAGTTAGGCAGTTGAGGAGCACATCAATTGCAAAATAATCAGATGTA 1075

Qy          48 SerProSerGluAspThrGln---GlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
                   |||       |||:::   |||  :::   :::      |||||||||:::|||
Db        1074 CCCAAGTCCACCCAAACACGTCCCCAGTTTTCCTGGAACTCAATATCACTGATATCATGG 1015

Qy          67 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGly 82
                  ||||||:::   |||       ||||||         ||||||
Db        1014 AAAGATGACGGCATCACTTTGAAACCTTTATCTGCATCATATAGAGGA 967

RESULT 50
US-10-363-345A-21005/c
; Sequence 21005, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21005
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21005
US-10-363-345A-21005

Alignment Scores:
Pred. No.:                1.34          Length:          894
Score:                    77.00         Matches:         24
Percent Similarity:       39.8%         Conservative:    9
Best Local Similarity:    28.9%         Mismatches:      24
Query Match:              17.2%         Indels:          26
DB:                       8             Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-345A-21005 (1-894)

Qy           2 GluArgAlaLeuAsnSerTyr---------------------------------------- 8
               |||   |||||||||:::|||
Db         693 GAAAACGCGCTTAACGCGTACGTACGCGACGACGATTAACGACGCGCGAACAACGTAAAA 634

Qy           9 ------------------PheGluProProValGluGluSer------------------ 16
                                 ||||||||        ::::::
Db         633 CGAAACGAAACAAAACTTTTCGAACCCACGAAACGACAAAACGAACGACCCCTACGACCG 574

Qy          17 AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsn 36
               :::      ||||||||||||   :::|||:::   ||||||              |||
Db         573 TCGACGACCCGACGACCCGAAATATTATCTAAAAAAAAAACG---ACAACCGCGACGACG 517

Qy          37 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 56
                  ||||||   :::|||   :::|||   :::|||                    |||
Db         516 ACGACTACAACGACAACAACCGAAACGAAAACTAACCCTAAAACAACAAACGACTCCGAA 457

Qy          57 AsnGlySer 59
               |||   |||
Db         456 AACGAATCT 448

RESULT 51
```

```
US-10-363-345A-21006
; Sequence 21006, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21006
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21006
US-10-363-345A-21006

Alignment Scores:
Pred. No.:                1.34          Length:          894
Score:                    77.00         Matches:         24
Percent Similarity:       39.8%         Conservative:    9
Best Local Similarity:    28.9%         Mismatches:      24
Query Match:              17.2%         Indels:          26
DB:                       8             Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-345A-21006 (1-894)

Qy           2 GluArgAlaLeuAsnSerTyr---------------------------------------- 8
               |||   |||||||||:::|||
Db         202 GAAAACGCGCTTAACGCGTACGTACGCGACGACGATTAACGACGCGCGAACAACGTAAAA 261

Qy           9 ------------------PheGluProProValGluGluSer------------------ 16
                                 ||||||||        ::::::
Db         262 CGAAACGAAACAAAACTTTTCGAACCCACGAAACGACAAAACGAACGACCCCTACGACCG 321

Qy          17 AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsn 36
               :::      ||||||||||||   :::|||:::   ||||||              |||
Db         322 TCGACGACCCGACGACCCGAAATATTATCTAAAAAAAAAACG---ACAACCGCGACGACG 378

Qy          37 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 56
                  ||||||   :::|||   :::|||   :::|||                    |||
Db         379 ACGACTACAACGACAACAACCGAAACGAAAACTAACCCTAAAACAACAAACGACTCCGAA 438

Qy          57 AsnGlySer 59
               |||   |||
Db         439 AACGAATCT 447

RESULT 52
US-10-363-483A-21005/c
; Sequence 21005, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21005
;   LENGTH: 894
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 21005
```

US-10-363-483A-21005

Alignment Scores:

Pred. No.:	1.34	Length:	894
Score:	77.00	Matches:	24
Percent Similarity:	39.8%	Conservative:	9
Best Local Similarity:	28.9%	Mismatches:	24
Query Match:	17.2%	Indels:	26
DB:	9	Gaps:	3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-483A-21005 (1-894)

```
Qy          2 GluArgAlaLeuAsnSerTyr---------------------------------------- 8
              |||   |||||||||:::|||
Db        693 GAAAACGCGCTTAACGCGTACGTACGCGACGACGATTAACGACGCGCGAACAACGTAAAA 634

Qy          9 ------------------PheGluProProValGluGluSer------------------ 16
                              |||||||||         ::::::
Db        633 CGAAACGAAACAAAACTTTTCGAACCCACGAAACGACAAAACGAACGACCCCTACGACCG 574

Qy         17 AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsn 36
              :::     ||||||||||||   :::|||:::   ||||||              |||
Db        573 TCGACGACCCGACGACCCGAAATATTATCTAAAAAAAAAACG---ACAACCGCGACGACG 517

Qy         37 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 56
                    ||||||   :::|||   :::|||   :::|||                  |||
Db        516 ACGACTACAACGACAACAACCGAAACGAAAACTAACCCTAAAACAACAAACGACTCCGAA 457

Qy         57 AsnGlySer 59
              |||   |||
Db        456 AACGAATCT 448
```

RESULT 53
US-10-363-483A-21006
; Sequence 21006, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21006
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 21006
US-10-363-483A-21006

Alignment Scores:

Pred. No.:	1.34	Length:	894
Score:	77.00	Matches:	24
Percent Similarity:	39.8%	Conservative:	9
Best Local Similarity:	28.9%	Mismatches:	24
Query Match:	17.2%	Indels:	26
DB:	9	Gaps:	3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-483A-21006 (1-894)

```
Qy          2 GluArgAlaLeuAsnSerTyr---------------------------------------- 8
              |||   |||||||||:::|||
Db        202 GAAAACGCGCTTAACGCGTACGTACGCGACGACGATTAACGACGCGCGAACAACGTAAAA 261

Qy          9 ------------------PheGluProProValGluGluSer------------------ 16
                              |||||||||         ::::::
Db        262 CGAAACGAAACAAAACTTTTCGAACCCACGAAACGACAAAACGAACGACCCCTACGACCG 321

Qy         17 AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsn 36
              :::     ||||||||||||   :::|||:::   ||||||              |||
Db        322 TCGACGACCCGACGACCCGAAATATTATCTAAAAAAAAAACG---ACAACCGCGACGACG 378

Qy         37 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 56
                    ||||||   :::|||   :::|||   :::|||                  |||
Db        379 ACGACTACAACGACAACAACCGAAACGAAAACTAACCCTAAAACAACAAACGACTCCGAA 438

Qy         57 AsnGlySer 59
              |||   |||
Db        439 AACGAATCT 447
```

RESULT 54
US-10-437-963-99213
; Sequence 99213, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 99213
; LENGTH: 2933
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_97045C.1
US-10-437-963-99213

Alignment Scores:

Pred. No.:	9.61	Length:	2933
Score:	76.00	Matches:	28
Percent Similarity:	50.0%	Conservative:	12
Best Local Similarity:	35.0%	Mismatches:	28
Query Match:	17.0%	Indels:	12
DB:	7	Gaps:	5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-437-963-99213 (1-2933)

```
Qy          8 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 27
              |||       |||||||||   |||:::|||       |||||||||       :::|||
Db       1122 TATTGTCGTCCACCTGTGACGGAGAATGCAGCTCCCCGTAGACCGACGGTTGTTTCTAAC 1181

Qy         28 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 47
                        :::   |||||||||         |||                  |||:::
Db       1182 ------------ATGGCATTAACCAAC---------ACGCCAGATAATGTTGAAAAATTA 1220

Qy         48 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 67
                    :::   |||   :::||||||                  |||::::::      |||
Db       1221 ------ACCACTGATGATAAGCAAGAGCTTAGTATAGATCCAAGTATTATG------AAT 1268

Qy         68 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              :::   |||:::|||       |||:::   |||   |||:::          |||||||||
Db       1269 TTACCAGGAGTAGATTGTTTCAATATAAAGGAAATTGCAAAGAAGGAG---AGTTATTTA 1325
```

RESULT 55
US-10-153-219-1
; Sequence 1, Application US/10153219
; Publication No. US20030096349A1
; GENERAL INFORMATION:
; APPLICANT: KAZMIERCZAK, KRYSTYNA M.
; APPLICANT: DAVYDOVA K.
; APPLICANT: ROTHMAN-DENES LUCIA B.

```
;  TITLE OF INVENTION: N4 VIRION SINGLE STRANDED DNA DEPENDENT RNA POLYMERASE
;  FILE REFERENCE: ARCD:375US
;  CURRENT APPLICATION NUMBER: US/10/153,219
;  CURRENT FILING DATE:  2002-05-22
;  PRIOR APPLICATION NUMBER: 60/292,845
;  PRIOR FILING DATE: 2001-05-22
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 10506
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
US-10-153-219-1

Alignment Scores:
Pred. No.:                66              Length:         10506
Score:                    75.50           Matches:        24
Percent Similarity:       46.1%           Conservative:   17
Best Local Similarity:    27.0%           Mismatches:     27
Query Match:              16.9%           Indels:         21
DB:                       5               Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-153-219-1 (1-10506)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||:::            |||:::   |||   ||||||
Db       2851 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 2898

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::|||||||||||||:::                 |||
Db       2899 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 2958

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::     :::|||  :::  :::|||                |||   :::
Db       2959 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3003

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       3004 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3063

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                :::|||   |||:::   :::|||
Db       3064 TCTAAAGCAGAGGGTATTAAGAACTAT 3090

RESULT 56
US-10-719-372-1
; Sequence 1, Application US/10719372
; Publication No. US20040171041A1
; GENERAL INFORMATION:
;  APPLICANT: Davydova, Elena K.
;  APPLICANT:  Rothman-Denes, Lucia B.
;  APPLICANT:  Dahl, Gary A.
;  APPLICANT:  Gerdes, Svetlana Y.
;  APPLICANT:  Jendrisak, Jerome J.
;  TITLE OF INVENTION: PREPARATION AND USE OF SINGLE-STRANDED TRANSCRIPTION SUBSTRATES
;  TITLE OF INVENTION:  FOR SYNTHESIS OF TRANSCRIPTION PRODUCTS CORRESPONDING TO TARGET
;  TITLE OF INVENTION:  SEQUENCES
;  FILE REFERENCE: ARCD:
;  CURRENT APPLICATION NUMBER: US/10/719,372
;  CURRENT FILING DATE:  2003-11-22
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 10506
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
US-10-719-372-1

Alignment Scores:
Pred. No.:                66              Length:         10506
Score:                    75.50           Matches:        24
Percent Similarity:       46.1%           Conservative:   17
Best Local Similarity:    27.0%           Mismatches:     27
Query Match:              16.9%           Indels:         21
DB:                       7               Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-719-372-1 (1-10506)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||:::            |||:::   |||   ||||||
Db       2851 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 2898

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::|||||||||||||:::                 |||
Db       2899 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 2958

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::     :::|||  :::  :::|||                |||   :::
Db       2959 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3003

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       3004 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3063

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                :::|||   |||:::   :::|||
Db       3064 TCTAAAGCAGAGGGTATTAAGAACTAT 3090

RESULT 57
US-10-743-975-1
; Sequence 1, Application US/10743975
; Publication No. US20040191812A1
; GENERAL INFORMATION:
;  APPLICANT: Davydova, Elena K.
;  APPLICANT:  Rothman-Denes, Lucia B.
;  APPLICANT:  Dahl, Gary A.
;  APPLICANT:  Gerdes, Svetlana Y.
;  APPLICANT:  Jendrisak, Jerome J.
;  TITLE OF INVENTION: TARGET-DEPENDENT TRANSCRIPTION USING DELETION MUTANTS OF N4 RNA
;  TITLE OF INVENTION:  POLYMERASE
;  FILE REFERENCE: ARCD:?
;  CURRENT APPLICATION NUMBER: US/10/743,975
;  CURRENT FILING DATE:  2003-12-23
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 10506
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
US-10-743-975-1

Alignment Scores:
Pred. No.:                66              Length:         10506
Score:                    75.50           Matches:        24
Percent Similarity:       46.1%           Conservative:   17
Best Local Similarity:    27.0%           Mismatches:     27
Query Match:              16.9%           Indels:         21
DB:                       8               Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-743-975-1 (1-10506)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||:::            |||:::   |||   ||||||
Db       2851 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 2898

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::|||||||||||||:::                 |||
```

```
Db       2899 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 2958

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                :::      :::|||   :::   :::|||               |||   :::
Db       2959 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3003

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                 :::
Db       3004 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3063

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                ::::|||   |||:::   :::|||
Db       3064 TCTAAAGCAGAGGGTATTAAGAACTAT 3090

RESULT 58
US-10-153-219-14
; Sequence 14, Application US/10153219
; Publication No. US20030096349A1
; GENERAL INFORMATION:
;  APPLICANT: KAZMIERCZAK, KRYSTYNA M.
;  APPLICANT:  DAVYDOVA K.
;  APPLICANT:  ROTHMAN-DENES LUCIA B.
;  TITLE OF INVENTION: N4 VIRION SINGLE STRANDED DNA DEPENDENT RNA POLYMERASE
;  FILE REFERENCE: ARCD:375US
;  CURRENT APPLICATION NUMBER: US/10/153,219
;  CURRENT FILING DATE:  2002-05-22
;  PRIOR APPLICATION NUMBER: 60/292,845
;  PRIOR FILING DATE: 2001-05-22
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 10617
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
US-10-153-219-14

Alignment Scores:
Pred. No.:              67              Length:         10617
Score:                  75.50           Matches:        24
Percent Similarity:     46.1%           Conservative:   17
Best Local Similarity:  27.0%           Mismatches:     27
Query Match:            16.9%           Indels:         21
DB:                     5               Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-153-219-14 (1-10617)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::            |||:::   |||   ||||||
Db       2962 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 3009

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||      |||:::|||||||||||||:::                |||
Db       3010 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 3069

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                :::      :::|||   :::   :::|||               |||   :::
Db       3070 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3114

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                 :::
Db       3115 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3174

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                ::::|||   |||:::   :::|||
Db       3175 TCTAAAGCAGAGGGTATTAAGAACTAT 3201

RESULT 59
US-10-719-372-14
; Sequence 14, Application US/10719372
; Publication No. US20040171041A1
; GENERAL INFORMATION:
;  APPLICANT: Davydova, Elena K.
;  APPLICANT:  Rothman-Denes, Lucia B.
;  APPLICANT:  Dahl, Gary A.
;  APPLICANT:  Gerdes, Svetlana Y.
;  APPLICANT:  Jendrisak, Jerome J.
;  TITLE OF INVENTION: PREPARATION AND USE OF SINGLE-STRANDED TRANSCRIPTION SUBSTRATES
;  TITLE OF INVENTION:  FOR SYNTHESIS OF TRANSCRIPTION PRODUCTS CORRESPONDING TO TARGET
;  TITLE OF INVENTION:  SEQUENCES
;  FILE REFERENCE: ARCD:
;  CURRENT APPLICATION NUMBER: US/10/719,372
;  CURRENT FILING DATE:  2003-11-22
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 10617
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
US-10-719-372-14

Alignment Scores:
Pred. No.:              67              Length:         10617
Score:                  75.50           Matches:        24
Percent Similarity:     46.1%           Conservative:   17
Best Local Similarity:  27.0%           Mismatches:     27
Query Match:            16.9%           Indels:         21
DB:                     7               Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-719-372-14 (1-10617)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::            |||:::   |||   ||||||
Db       2962 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 3009

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||      |||:::|||||||||||||:::                |||
Db       3010 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 3069

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                :::      :::|||   :::   :::|||               |||   :::
Db       3070 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3114

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                 :::
Db       3115 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3174

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                ::::|||   |||:::   :::|||
Db       3175 TCTAAAGCAGAGGGTATTAAGAACTAT 3201

RESULT 60
US-10-743-975-14
; Sequence 14, Application US/10743975
; Publication No. US20040191812A1
; GENERAL INFORMATION:
;  APPLICANT: Davydova, Elena K.
;  APPLICANT:  Rothman-Denes, Lucia B.
;  APPLICANT:  Dahl, Gary A.
;  APPLICANT:  Gerdes, Svetlana Y.
;  APPLICANT:  Jendrisak, Jerome J.
;  TITLE OF INVENTION: TARGET-DEPENDENT TRANSCRIPTION USING DELETION MUTANTS OF N4 RNA
;  TITLE OF INVENTION:  POLYMERASE
;  FILE REFERENCE: ARCD:?
;  CURRENT APPLICATION NUMBER: US/10/743,975
;  CURRENT FILING DATE:  2003-12-23
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 10617
;   TYPE: DNA
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
US-10-743-975-14

Alignment Scores:
Pred. No.:                67              Length:          10617
Score:                    75.50           Matches:         24
Percent Similarity:       46.1%           Conservative:    17
Best Local Similarity:    27.0%           Mismatches:      27
Query Match:              16.9%           Indels:          21
DB:                       8               Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-743-975-14 (1-10617)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::            |||:::   |||   ||||||
Db       2962 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 3009

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||     |||:::|||||||||||||:::           |||
Db       3010 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 3069

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                :::      :::|||  :::  :::|||                |||   :::
Db       3070 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3114

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                  :::
Db       3115 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3174

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
              :::|||   |||:::   :::|||
Db       3175 TCTAAAGCAGAGGGTATTAAGAACTAT 3201

RESULT 61
US-10-450-763-8911/c
; Sequence 8911, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 8911
;   LENGTH: 2490
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2488)..(2219)
;   OTHER INFORMATION: 93% homologous to Homo sapiens putative p150,accession number
;   OTHER INFORMATION: U93563,Smith-Waterman Score=429.
US-10-450-763-8911

Alignment Scores:
Pred. No.:                14.9            Length:          2490
Score:                    74.00           Matches:         28
Percent Similarity:       44.6%           Conservative:    13
Best Local Similarity:    30.4%           Mismatches:      25
Query Match:              16.5%           Indels:          26
DB:                       9               Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-8911 (1-2490)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGlu--- 19
              |||::::::   ||||||:::|||      |||   :::   :::      ||||||
Db       1994 ATGGATAAATTCCTCAACACATACACCCTCCCAAGACTAAACCAGGAAGAACTTGAATCT 1935

Qy         20 ---ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 38
                  ||||||   |||   ||||||   :::|||  :::
Db       1934 CTGAATAGACCAATAACAGGCTCTGAAATTGAGACAATAATT----------------- 1893

Qy         39 ThrThrAspSerThrThrSerLysIleSerProSerGluAsp----------------- 52
                 :::|||      :::|||   ||||||      |||
Db       1892 ------AATAGCTTACCAACCAAAAAAAGTCCAGGATCAGATGGATTCACAGCCGAATTC 1839

Qy         53 ------------------------ThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                                       |||::::::||||||             ||||||:::
Db       1838 TACCAGAGGCCTGGCAGAGACACAACAAAAAAAGAGAATTTTAGACCAATATCCTTGATG 1779

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 76
                 |||||||||         ||||||   :::
Db       1778 ------AACATTGATGCAAAAATCCTCAATAAAATA 1749

RESULT 62
US-11-037-601-1
; Sequence 1, Application US/11037601
; Publication No. US20050158345A1
; GENERAL INFORMATION:
;  APPLICANT: Actinova Limited
;  APPLICANT:  Bjorck, Lars
;  APPLICANT:  Janulczyk, Robert
;  TITLE OF INVENTION: STREPTOCOCCAL ABC TRANSPORTER PROTEIN
;  FILE REFERENCE: KCO1004US
;  CURRENT APPLICATION NUMBER: US/11/037,601
;  CURRENT FILING DATE:  2005-01-18
;  PRIOR APPLICATION NUMBER: US/09/869,677
;  PRIOR FILING DATE: 2001-10-09
;  PRIOR APPLICATION NUMBER: GB 9828880.6
;  PRIOR FILING DATE: 1998-12-31
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;   LENGTH: 893
;   TYPE: DNA
;   ORGANISM: Streptococcus pyogenes
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(861)
US-11-037-601-1

Alignment Scores:
Pred. No.:                4.27            Length:          893
Score:                    73.50           Matches:         21
Percent Similarity:       54.2%           Conservative:    18
Best Local Similarity:    29.2%           Mismatches:      18
Query Match:              16.4%           Indels:          15
DB:                       10              Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x US-11-037-601-1 (1-893)

Qy         15 GluSerAlaLeuGluArgArgPro---GluThrIleSerGluProLysThrTyrValAsp 33
              ||||||:::::::::|||||||   ||||||:::|||:::      :::   :::
Db        700 GAGTCAAGTGTCGATAGACGCCCTATGGAAACTGTTTCTAAA---GATAGTGGTATTCCT 756

Qy         34 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
              :::   :::|||      |||||||||          |||      |||   :::
Db        757 ATTTATTCTGAGATCTTTACAGATTCAATTGCTAAAAAAGGTAAACCTGGCGAT------ 810

Qy         54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 73
                        |||   ::::::::::::   ||||||:::|||   :::
Db        811 --------------AGTTATTATGCTATGATGAAATGGAACCTTGACAAAATTTCTGAA 855

Qy         74 ------------------AsnAsnLeuSerGluArg 79
                             ||||||   ::::::|||
```

```
Db        856 GGTCTAGCAAAATAAGATAACAATAAAGCACAGCGA 891

RESULT 63
US-10-363-345A-20611/c
; Sequence 20611, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 20611
;   LENGTH: 1339
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 20611
US-10-363-345A-20611

Alignment Scores:
Pred. No.:              12.3          Length:         1339
Score:                  72.00         Matches:        24
Percent Similarity:     48.5%         Conservative:   8
Best Local Similarity:  36.4%         Mismatches:     18
Query Match:            16.1%         Indels:         16
DB:                     8             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-345A-20611 (1-1339)

Qy         19 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 38
              |||   |||||||||||||:::   |||||||||
Db        819 GAAACGCGACCCGAAACCGTACGCGAACCGAAA--------------------CACCGA 781

Qy         39 ThrThrAspSerThrThr--------------SerLysIleSerProSerGluAspThr 53
              ||||||   :::||||||               :::||||||:::             |||
Db        780 ACGACGACGACGACGACGCGCGCCATATCGTTCAATAAAATAAACCGCAAAACGCTAACG 721

Qy         54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 73
                 ::::::|||           |||||||||   |||   |||
Db        720 TTCCGAAAAAAC-----------GAATTAATCACCGCTAACGACGACGACTCCACGAAC 673

Qy         74 AsnAsnLeuSerGluArg 79
              ||||||   :::|||:::
Db        672 AATAACGCTAACGAAAAA 655

RESULT 64
US-10-363-345A-20612
; Sequence 20612, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Method for determining the degree of methylation of defined
;  TITLE OF INVENTION:  cytosines in genomic DNA in the sequence context of 5'-CpG-3
;  FILE REFERENCE: E01/1227
;  CURRENT APPLICATION NUMBER: US/10/363,345A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 20612
;   LENGTH: 1339
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 20612
US-10-363-345A-20612

Alignment Scores:
Pred. No.:              12.3          Length:         1339
Score:                  72.00         Matches:        24
Percent Similarity:     48.5%         Conservative:   8
Best Local Similarity:  36.4%         Mismatches:     18
Query Match:            16.1%         Indels:         16
DB:                     8             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-345A-20612 (1-1339)

Qy         19 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 38
              |||   |||||||||||||:::   |||||||||
Db        521 GAAACGCGACCCGAAACCGTACGCGAACCGAAA--------------------CACCGA 559

Qy         39 ThrThrAspSerThrThr--------------SerLysIleSerProSerGluAspThr 53
              ||||||   :::||||||               :::||||||:::             |||
Db        560 ACGACGACGACGACGACGCGCGCCATATCGTTCAATAAAATAAACCGCAAAACGCTAACG 619

Qy         54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 73
                 ::::::|||           |||||||||   |||   |||
Db        620 TTCCGAAAAAAC-----------GAATTAATCACCGCTAACGACGACGACTCCACGAAC 667

Qy         74 AsnAsnLeuSerGluArg 79
              ||||||   :::|||:::
Db        668 AATAACGCTAACGAAAAA 685

RESULT 65
US-10-363-483A-20611/c
; Sequence 20611, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 20611
;   LENGTH: 1339
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 20611
US-10-363-483A-20611

Alignment Scores:
Pred. No.:              12.3          Length:         1339
Score:                  72.00         Matches:        24
Percent Similarity:     48.5%         Conservative:   8
Best Local Similarity:  36.4%         Mismatches:     18
Query Match:            16.1%         Indels:         16
DB:                     9             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-483A-20611 (1-1339)

Qy         19 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 38
              |||   |||||||||||||:::   |||||||||
Db        819 GAAACGCGACCCGAAACCGTACGCGAACCGAAA--------------------CACCGA 781

Qy         39 ThrThrAspSerThrThr--------------SerLysIleSerProSerGluAspThr 53
              ||||||   :::||||||               :::||||||:::             |||
Db        780 ACGACGACGACGACGACGCGCGCCATATCGTTCAATAAAATAAACCGCAAAACGCTAACG 721

Qy         54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 73
                 ::::::|||           |||||||||   |||   |||
Db        720 TTCCGAAAAAAC-----------GAATTAATCACCGCTAACGACGACGACTCCACGAAC 673

Qy         74 AsnAsnLeuSerGluArg 79
```

```
                ||||||   :::|||:::
Db        672 AATAACGCTAACGAAAAA 655

RESULT 66
US-10-363-483A-20612
; Sequence 20612, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
;  APPLICANT: Alexander Olek
;  APPLICANT:  Christian Piepenbrock
;  APPLICANT:  Kurt Berlin
;  TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
;  TITLE OF INVENTION:  illnesses
;  FILE REFERENCE: 82011
;  CURRENT APPLICATION NUMBER: US/10/363,483A
;  CURRENT FILING DATE:  2003-03-03
;  NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 20612
;   LENGTH: 1339
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   OTHER INFORMATION: CpG-island No: 20612
US-10-363-483A-20612

Alignment Scores:
Pred. No.:                12.3          Length:         1339
Score:                    72.00         Matches:        24
Percent Similarity:       48.5%         Conservative:   8
Best Local Similarity:    36.4%         Mismatches:     18
Query Match:              16.1%         Indels:         16
DB:                       9             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-483A-20612 (1-1339)

Qy         19 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 38
              |||   |||||||||||:::   ||||||||
Db        521 GAAACGCGACCCGAAACCGTACGCGAACCGAAA--------------------CACCGA 559

Qy         39 ThrThrAspSerThrThr---------------SerLysIleSerProSerGluAspThr 53
              ||||||   :::|||||||               :::|||||:::           |||
Db        560 ACGACGACGACGACGACGCGCGCCATATCGTTCAATAAAATAAACCGCAAAACGCTAACG 619

Qy         54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 73
              :::::||             |||||||||   |||   |||
Db        620 TTCCGAAAAAAC------------GAATTAATCACCGCTAACGACGACGACTCCACGAAC 667

Qy         74 AsnAsnLeuSerGluArg 79
              ||||||   :::|||:::
Db        668 AATAACGCTAACGAAAAA 685

RESULT 67
US-11-097-143-34261
; Sequence 34261, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34261
;   LENGTH: 2555
;   TYPE: DNA
;   ORGANISM: DROSOPHILA
US-11-097-143-34261

Alignment Scores:
Pred. No.:                30            Length:         2555
Score:                    72.00         Matches:        20
Percent Similarity:       50.8%         Conservative:   13
Best Local Similarity:    30.8%         Mismatches:     26
Query Match:              16.1%         Indels:         6
DB:                       10            Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-097-143-34261 (1-2555)

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              :::|||   |||   :::|||         ||||||  ||||||:::      :::
Db        548 AAGCCGCATACATGCACAGAAAGGTACGCGTATGTA---CTTACATCTTCCTCGCGATCT 604

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
                ||||||   :::            |||               :::   ||||||
Db        605 GCAAGCACAATCACATCGGCGGACTCGTCC--------------GATAGGGGCTCATCC 649

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              |||:::::::::     |||         |||:::|||      :::   :::|||
Db        650 TTCACCATTGTGTTAGCAAATTAATTATCACTAAATTTAACCGAAATTAAGAAACGAAAA 709

Qy         81 ArgGlyValCysSer 85
              |||   ||||||:::
Db        710 CGACCAGTTTGCACG 724

RESULT 68
US-10-437-963-5051/c
; Sequence 5051, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 5051
;   LENGTH: 2658
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_11877C.1
US-10-437-963-5051

Alignment Scores:
Pred. No.:                31.7          Length:         2658
Score:                    72.00         Matches:        28
Percent Similarity:       35.5%         Conservative:   15
```

```
Best Local Similarity:  23.1%          Mismatches:    28
Query Match:            16.1%          Indels:        50
DB:                     7              Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-437-963-5051 (1-2658)

Qy         11 ProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThr 30
              ||||||:::   :::|||::::::   :::             :::   ||||||:::|||
Db       1941 CCACCTATGTTGGATAGTTCTGTTGGTAAAACAACAACTGATGTAAAAGAGCCACGGACG 1882

Qy         31 TyrVal---------------------------------AspLeu---ThrAsnGluGlu 38
                 :::                               ||||||   |||      |||
Db       1881 TCAATTCGAGGACCGGCTGTAACACCAAATAGCAGCCCTGACTTGGAGACAACTTTTGAA 1822

Qy         39 ThrThrAspSerThrThrSerLysIleSerProSerGlu-----------------Asp 52
              :::|||::::::   ||||||   :::|||   |||:::
Db       1821 AGTACTGAAGCTGGAACATCCTCTGTCTCAAGTTCAGATCCAGGGACTTCTCCATATAGT 1762

Qy         53 ThrGlnGlnGluAsnGlySerMet---------------------------------- 60
                        ||||||||||||
Db       1761 GCTTCTGACACAAATGGTTCTATGAAGAAAGAGGCACCAGCAACAAAGGATAGGGTGCAG 1702

Qy         61 ------------------------------------------------------------Phe 61
                                                                          |||
Db       1701 CATTTGGACATCAATATTTCTGATTCTGACAGCGAAACATTGAGTCCTCCAGCATCCTTC 1642

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
              ||||||      |||   :::|||   |||         :::   |||   |||
Db       1641 TCGCTTCAACCATGGATGGTTGACATTCTACAG---GGATCAGCATCGTCGAGAACACAT 1585

Qy         82 Gly 82
              |||
Db       1584 GGA 1582
```

RESULT 69
US-11-097-143-13687
; Sequence 13687, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13687
; LENGTH: 3748
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-13687

```
Alignment Scores:
Pred. No.:              50.9           Length:        3748
Score:                  72.00          Matches:       20
Percent Similarity:     50.8%          Conservative:  13
Best Local Similarity:  30.8%          Mismatches:    26
Query Match:            16.1%          Indels:        6
DB:                     10             Gaps:          2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-097-143-13687 (1-3748)

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              :::|||   |||   :::|||          ||||||   ||||||:::         :::
Db       2639 AAGCCGCATACATGCACAGAAAGGTACGCGTATGTA---CTTACATCTTCCTCGCGATCT 2695

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
                 ||||||   :::              |||                  :::   ||||||
Db       2696 GCAAGCACAATCACATCGGCGGACTCGTCC--------------GATAGGGGCTCATCC 2740

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              |||:::::::::      |||         |||:::|||      :::   :::|||
Db       2741 TTCACCATTGTGTTAGCAAATTAATTATCACTAAATTTAACCGAAATTAAGAAACGAAAA 2800

Qy         81 ArgGlyValCysSer 85
              |||   ||||||:::
Db       2801 CGACCAGTTTGCACG 2815
```

RESULT 70
US-10-369-493-26910
; Sequence 26910, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26910
; LENGTH: 4964
; TYPE: DNA
; ORGANISM: Neurospora crassa
US-10-369-493-26910

```
Alignment Scores:
Pred. No.:              75             Length:        4964
Score:                  72.00          Matches:       16
Percent Similarity:     50.0%          Conservative:  15
Best Local Similarity:  25.8%          Mismatches:    15
Query Match:            16.1%          Indels:        16
DB:                     6              Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-369-493-26910 (1-4964)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              ||||||             :::::::::      :::      |||:::::::::||||||
Db       4052 GAGCCCGTTACTACCATCGCCAGTGTGCGCGCCAAGCGTGCCACCGTAGCGAAGCCTAAG 4111

Qy         30 ThrTyr-------------------------------------------------ValAsp 33
              ||||||                                                 |||
Db       4112 ACTTATGCGCTCAGTGATGAAGATTCGGACGATGAGTTCTTGGACATCGCAAAGGTTACC 4171

Qy         34 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
                 |||:::|||:::||||||   |||   ::::::   :::|||   :::|||   :::
Db       4172 ACAACTTCTGAGAAAACGACCACGTCGGCTTCAGCTTCGGTTTCAGCGGCAGAACAATCT 4231

Qy         54 GlnGln 55
              :::|||
Db       4232 GAGCAA 4237
```

```
RESULT 71
US-10-451-337-3
; Sequence 3, Application US/10451337
; Publication No. US20040097706A1
; GENERAL INFORMATION:
;  APPLICANT: SHIRE BIOCHEM INC.
;  APPLICANT:  MARTIN, Denis
;  APPLICANT:  BRODEUR, Bernard R.
;  APPLICANT:  HAMEL, Josee
;  APPLICANT:  RIOUX, Stephane
;  APPLICANT:  RHEAULT, Patrick
;  TITLE OF INVENTION: STREPTOCOCCUS PYOGENES ANTIGENS AND
;  TITLE OF INVENTION:  CORRESPONDING DNA FRAGMENTS
;  FILE REFERENCE: 12806-24PCT
;  CURRENT APPLICATION NUMBER: US/10/451,337
;  CURRENT FILING DATE:  2003-11-18
;  PRIOR APPLICATION NUMBER: US 60/256,940
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 921
;   TYPE: DNA
;   ORGANISM: S. pyogenes
US-10-451-337-3

Alignment Scores:
Pred. No.:              8.66          Length:         921
Score:                  71.50         Matches:        18
Percent Similarity:     58.9%         Conservative:   15
Best Local Similarity:  32.1%         Mismatches:     14
Query Match:            16.0%         Indels:         9
DB:                     7             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-451-337-3 (1-921)

Qy          15 GluSerAlaLeuGluArgArgPro---GluThrIleSerGluProLysThrTyrValAsp 33
               ||||||:::::::::|||||||||   ||||||:::|||:::      :::   :::
Db         751 GAGTCAAGTGTCGATAGACGCCCTATGGAAACTGTTTCTAAA---GATAGTGGTATTCCT 807

Qy          34 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
               :::   :::|||      |||||||||          |||       |||   :::
Db         808 ATTTATTCTGAGATCTTTACAGATTCAATTGCTAAAAAAGGTAAACCTGGCGAT------ 861

Qy          54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 69
                        |||   ::::::::::::   ||||||:::|||
Db         862 ----------------AGTTATTATGCTATGATGAAATGGAACCTTGAC 894

RESULT 72
US-11-097-143-7073
; Sequence 7073, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7073
;   LENGTH: 4320
;   TYPE: DNA
;   ORGANISM: DROSOPHILA
US-11-097-143-7073

Alignment Scores:
Pred. No.:              73.1          Length:         4320
Score:                  71.50         Matches:        27
Percent Similarity:     42.0%         Conservative:   15
Best Local Similarity:  27.0%         Mismatches:     35
Query Match:            16.0%         Indels:         23
DB:                     10            Gaps:           6

US-10-757-745-2_COPY_54_140 (1-87) x US-11-097-143-7073 (1-4320)

Qy           4 AlaLeuAsnSerTyrPhe-----------GluProProValGluGluSerAlaLeuGlu 19
               |||||||||   ||||||           :::||||||      :::      |||:::
Db        2215 GCCTTAAACGAGTACTTTACGCAGAACAACCAACCTCCGGAGAACAAGCAGCAGTTGAAG 2274

Qy          20 ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 39
               :::     |||           |||   :::|||:::      :::|||||||||
Db        2275 AAGGCCGTGGAA------GAGGAATATCGAACCTGGACGTCGATGACCAACGAG------ 2322

Qy          40 ThrAspSerThrThrSerLysIleSerProSerGluAspThr----------------- 53
                  |||               :::   |||:::   :::
Db        2323 AATGACATAATTGCGCACTTCTTGCGTCCAACTACGAATCCTTTATTCCTTTGCCTTCTG 2382

Qy          54 --------------GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle 68
                           :::   ||||||           |||   :::   :::   |||
Db        2383 TTCAAAATAATTTGGGAGACGGAGAAC---------ATAAGCCCAGTGGCTTACAAAATT 2433

Qy          69 ---AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
                  :::|||:::          ||||||      |||      |||    ||||||
Db        2434 TTGGAGGGTATAAGCGCGAGAGCCTTGTCCACTCATCTGCGTAAATTTTGTGACTACTTG 2493

RESULT 73
US-10-367-094-90/c
; Sequence 90, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
;  FILE REFERENCE: 529452001500
;  CURRENT APPLICATION NUMBER: US/10/367,094
;  CURRENT FILING DATE:  2003-02-14
;  NUMBER OF SEQ ID NOS: 203
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 90
;   LENGTH: 139613
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(139613)
;   OTHER INFORMATION: n = A,T,C or G
US-10-367-094-90

Alignment Scores:
Pred. No.:              8.88e+03      Length:         139613
Score:                  71.50         Matches:        20
Percent Similarity:     48.3%         Conservative:   8
Best Local Similarity:  34.5%         Mismatches:     27
Query Match:            16.0%         Indels:         3
DB:                     7             Gaps:           2
```

```
US-10-757-745-2_COPY_54_140 (1-87) x US-10-367-094-90 (1-139613)

Qy          12 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
               ||||||:::||||||   :::   :::   |||   :::          |||   |||
Db        5077 CCTGTTGATGAGAGCGTGGTACTAAAGTTACCAAATAGTTTTGGCAATCCGTCAACAAGA 5018

Qy          32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSer------LysIleSerPro 49
                 |||   |||                 :::      |||      |||   |||
Db        5017 TTCTCCCTA---AATGTATCAATTGCATCCAATGCACTGAGTATTTTAAAGTTTTCATTT 4961

Qy          50 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 67
               :::   ::::::||||||          |||   |||||||||      ||||||
Db        4960 AATTTGAATTCTCAGCAGAGTGGAAATAGTTATTTCTCCTTGTGAAACTGGAAC 4907

RESULT 74
US-10-027-632-288729/c
; Sequence 288729, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 288729
;   LENGTH: 498
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-288729

Alignment Scores:
Pred. No.:                4.37          Length:          498
Score:                    71.00         Matches:         24
Percent Similarity:       41.4%         Conservative:    17
Best Local Similarity:    24.2%         Mismatches:      23
Query Match:              15.8%         Indels:          35
DB:                       5             Gaps:            2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-027-632-288729 (1-498)

Qy          15 GluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThr------------ 30
               |||   :::||||||      |||:::   :::          :::|||
Db        407 GAAGAAAGCCTTGAAGAGTTCCCTGACAAGTTAGAAAGTTCAAGAACTGGAGGCAGAATT 348

Qy          31 --------------------TyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
                                   :::::::::::::::::::::      |||||||||
Db        347 GGGCCCTGAGGCTTTAAGCTCCCTCTTAACATTTCCCACCAACAATATTTGGACTCCACT 288

Qy          44 ThrSerLysIleSerPro---------------------------------------- 49
               :::      |||   |||
Db        287 AGTAAGCCCATAGGTCCCTCATTTACTGCTGTTTTATGGGGCCTAGGAAGTAGAGAGCAG 228

Qy          50 --------------------------SerGluAspThrGlnGlnGluAsnGlySerMet 60
                                        |||   |||   :::::::::||||||
Db        227 ATATAGTTCAAATTTGCTCCTTAAAACTCAGGAGACTACGAGGAGAAAAATGGCAGGGAC 168

Qy          61 PheSerLeuIleThr-TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 78
               |||||||||:::|||    ||||||   |||:::|||          |||   |||
Db        167 TTCTCTCTAGTGACCTGAGAATATTGGTGGCATAGACAAATGTGCACTAAAGGAG 113

RESULT 75
US-10-027-632-288729/c
; Sequence 288729, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 288729
;   LENGTH: 498
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-288729

Alignment Scores:
Pred. No.:                4.37          Length:          498
Score:                    71.00         Matches:         24
Percent Similarity:       41.4%         Conservative:    17
Best Local Similarity:    24.2%         Mismatches:      23
Query Match:              15.8%         Indels:          35
DB:                       6             Gaps:            2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-027-632-288729 (1-498)

Qy          15 GluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThr------------ 30
               |||   :::||||||      |||:::   :::          :::|||
Db        407 GAAGAAAGCCTTGAAGAGTTCCCTGACAAGTTAGAAAGTTCAAGAACTGGAGGCAGAATT 348

Qy          31 --------------------TyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
                                   :::::::::::::::::::::      |||||||||
Db        347 GGGCCCTGAGGCTTTAAGCTCCCTCTTAACATTTCCCACCAACAATATTTGGACTCCACT 288

Qy          44 ThrSerLysIleSerPro---------------------------------------- 49
               :::      |||   |||
Db        287 AGTAAGCCCATAGGTCCCTCATTTACTGCTGTTTTATGGGGCCTAGGAAGTAGAGAGCAG 228

Qy          50 --------------------------SerGluAspThrGlnGlnGluAsnGlySerMet 60
                                        |||   |||   :::::::::||||||
Db        227 ATATAGTTCAAATTTGCTCCTTAAAACTCAGGAGACTACGAGGAGAAAAATGGCAGGGAC 168

Qy          61 PheSerLeuIleThr-TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 78
               |||||||||:::|||    ||||||   |||:::|||          |||   |||
Db        167 TTCTCTCTAGTGACCTGAGAATATTGGTGGCATAGACAAATGTGCACTAAAGGAG 113


Search completed: March 11, 2006, 05:42:41
Job time : 469.849 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 20:06:44 ; Search time 98.9389 Seconds
(without alignments)
1563.065 Million cell updates/sec

Title: US-10-757-745-2_COPY_54_140
Perfect score: 448
Sequence: 1 MERALNSYFEPPVEESALER.....IDGLDLNNLSERARGVCSYL 87

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170331_4111/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_417_@runat_09032006_170331_4111 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	448	100.0	1920	3	US-09-697-863A-1	Sequence 1, Appli
	2	448	100.0	1948	3	US-09-620-312D-889	Sequence 889, App
	3	211.5	47.2	1312	3	US-09-697-863A-3	Sequence 3, Appli
	4	145	32.4	379	3	US-09-621-976-8403	Sequence 8403, Ap
	5	73.5	16.4	893	3	US-09-869-677B-1	Sequence 1, Appli
	6	72.5	16.2	573	3	US-09-248-796A-1008	Sequence 1008, Ap
	7	71.5	16.0	999	3	US-09-248-796A-12164	Sequence 12164, A
c	8	71	15.8	24459	3	US-09-902-540-5004	Sequence 5004, Ap
	9	71	15.8	32241	3	US-09-902-540-1247	Sequence 1247, Ap
c	10	70.5	15.7	15213	3	US-08-961-527-26	Sequence 26, Appl
	11	70	15.6	835	3	US-08-961-083-93	Sequence 93, Appl
	12	70	15.6	835	3	US-09-536-784-93	Sequence 93, Appl
	13	70	15.6	835	3	US-09-765-271-93	Sequence 93, Appl
	14	70	15.6	835	3	US-09-765-272A-93	Sequence 93, Appl
	15	70	15.6	1714	3	US-08-961-083-3	Sequence 3, Appli
	16	70	15.6	1714	3	US-09-536-784-3	Sequence 3, Appli
	17	70	15.6	1714	3	US-09-765-271-3	Sequence 3, Appli
	18	70	15.6	1714	3	US-09-765-272A-3	Sequence 3, Appli
	19	70	15.6	3640	3	US-09-976-594-888	Sequence 888, App
	20	70	15.6	4061	3	US-09-620-312D-363	Sequence 363, App
	21	69.5	15.5	486	3	US-09-248-796A-13914	Sequence 13914, A
	22	69	15.4	231	3	US-09-543-681A-1380	Sequence 1380, Ap
	23	69	15.4	291	3	US-08-894-699-1	Sequence 1, Appli
	24	69	15.4	291	3	US-09-444-410-1	Sequence 1, Appli
	25	68	15.2	2817	6	PCT-US93-05944-1	Sequence 1, Appli
	26	68	15.2	3356	3	US-09-379-523-4	Sequence 4, Appli
	27	68	15.2	4358	3	US-09-308-453-1	Sequence 1, Appli
	28	67.5	15.1	3033	2	US-08-095-737-3	Sequence 3, Appli
	29	67.5	15.1	3033	2	US-08-480-145-3	Sequence 3, Appli
	30	67.5	15.1	3033	2	US-08-477-389-3	Sequence 3, Appli
c	31	67.5	15.1	9623	3	US-08-956-171E-166	Sequence 166, App
c	32	67.5	15.1	9623	3	US-08-781-986A-166	Sequence 166, App
	33	67	15.0	610	3	US-09-280-116-22	Sequence 22, Appl
	34	67	15.0	938	3	US-09-270-767-1642	Sequence 1642, Ap
	35	67	15.0	938	3	US-09-270-767-16924	Sequence 16924, A
	36	67	15.0	1230	3	US-09-248-796A-542	Sequence 542, App
	37	67	15.0	1488	3	US-08-216-592A-7	Sequence 7, Appli
	38	67	15.0	1518	2	US-08-820-170A-20	Sequence 20, Appl
	39	67	15.0	1518	3	US-09-055-699-20	Sequence 20, Appl
	40	67	15.0	1518	3	US-09-273-565-20	Sequence 20, Appl
	41	67	15.0	1518	3	US-09-565-538-20	Sequence 20, Appl
	42	67	15.0	1518	3	US-09-661-468-20	Sequence 20, Appl
	43	67	15.0	1518	3	US-09-976-165-20	Sequence 20, Appl
	44	67	15.0	1662	2	US-08-336-408B-5	Sequence 5, Appli
	45	67	15.0	1662	6	PCT-US91-00399-5	Sequence 5, Appli
	46	67	15.0	2636	2	US-08-820-170A-21	Sequence 21, Appl
	47	67	15.0	2636	3	US-09-055-699-21	Sequence 21, Appl
	48	67	15.0	2636	3	US-09-273-565-21	Sequence 21, Appl
	49	67	15.0	2636	3	US-09-565-538-21	Sequence 21, Appl
	50	67	15.0	2636	3	US-09-661-468-21	Sequence 21, Appl
	51	67	15.0	2636	3	US-09-976-165-21	Sequence 21, Appl
c	52	67	15.0	4031	2	US-08-993-118-1	Sequence 1, Appli
c	53	67	15.0	4031	3	US-08-845-528C-1	Sequence 1, Appli
c	54	67	15.0	4031	3	US-09-066-281B-1	Sequence 1, Appli
c	55	67	15.0	4031	3	US-09-468-433C-1	Sequence 1, Appli
c	56	67	15.0	4225	2	US-08-993-118-9	Sequence 9, Appli
c	57	67	15.0	4225	3	US-08-845-528C-9	Sequence 9, Appli
c	58	67	15.0	4225	3	US-09-066-281B-9	Sequence 9, Appli
c	59	67	15.0	4225	3	US-09-468-433C-9	Sequence 9, Appli
c	60	67	15.0	4265	3	US-09-061-709-1	Sequence 1, Appli
c	61	67	15.0	4265	3	US-09-899-651-1	Sequence 1, Appli
c	62	67	15.0	4265	3	US-09-392-714-15	Sequence 15, Appl
c	63	67	15.0	4265	3	US-09-270-437D-1	Sequence 1, Appli
	64	67	15.0	10160	3	US-09-949-016-13215	Sequence 13215, A
	65	66.5	14.8	891	3	US-09-248-796A-9100	Sequence 9100, Ap
	66	66.5	14.8	3003	3	US-09-248-796A-6805	Sequence 6805, Ap
	67	66.5	14.8	3119	3	US-09-112-450-1	Sequence 1, Appli
	68	66.5	14.8	3119	3	US-09-419-291A-1	Sequence 1, Appli
	69	66.5	14.8	3119	3	US-10-116-048-1	Sequence 1, Appli
	70	66.5	14.8	8561	3	US-09-112-450-3	Sequence 3, Appli
	71	66.5	14.8	8561	3	US-09-419-291A-3	Sequence 3, Appli
	72	66.5	14.8	8561	3	US-10-116-048-3	Sequence 3, Appli
c	73	66	14.7	762	3	US-09-199-637A-153	Sequence 153, App
c	74	66	14.7	801	3	US-09-199-637A-154	Sequence 154, App
c	75	66	14.7	891	3	US-09-252-991A-13439	Sequence 13439, A

ALIGNMENTS

RESULT 1
US-09-697-863A-1
; Sequence 1, Application US/09697863A

```
; Patent No. 6812203
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/09/697,863A
;  CURRENT FILING DATE:  2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1627)..(1627)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1108)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1849)..(1849)
;   OTHER INFORMATION: N stands for any nucleotide.
US-09-697-863A-1

Alignment Scores:
Pred. No.:                4.52e-54        Length:          1920
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-697-863A-1 (1-1920)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 238

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 298

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 358

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 418

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        419 CGAGGGGTGTGTTCCTACTTA 439

RESULT 2
US-09-620-312D-889
; Sequence 889, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-09-620-312D-889

Alignment Scores:
Pred. No.:                4.61e-54        Length:          1948
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-620-312D-889 (1-1948)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 3
US-09-697-863A-3
; Sequence 3, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/09/697,863A
;  CURRENT FILING DATE:  2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
```

```
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;    LENGTH: 1312
;    TYPE: DNA
;    ORGANISM: Mus musculus
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (122)..(1234)
;    OTHER INFORMATION:
US-09-697-863A-3

Alignment Scores:
Pred. No.:              2.07e-20        Length:         1312
Score:                  211.50          Matches:        45
Percent Similarity:     66.3%           Conservative:   12
Best Local Similarity:  52.3%           Mismatches:     28
Query Match:            47.2%           Indels:         1
DB:                     3               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-697-863A-3 (1-1312)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::||||||::::::|||||||||   |||      :::            |||:::
Db        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||   |||      :::   :::   |||||||||||||||||||:::      |||
Db        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436

Qy         42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
              :::|||      :::   |||||||||      |||      |||:::   |||
Db        437 ACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGCAGCACTATT 493

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
              |||   ||||||||||||||||||||||||||      ||||||   ||||||||||||
Db        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553

Qy         82 GlyValCysSerTyrLeu 87
              ||||||||||||   |||
Db        554 GGGGTGTGTTCCTGCCTA 571

RESULT 4
US-09-621-976-8403
; Sequence 8403, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 8403
;    LENGTH: 379
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-621-976-8403

Alignment Scores:
Pred. No.:              1.24e-11        Length:         379
Score:                  145.00          Matches:        28
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  96.6%           Mismatches:     0
Query Match:            32.4%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-621-976-8403 (1-379)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 78

Qy         22 ProGluThrIleSerGluProLysThr 30
              |||||||||||||||||||||||||||
Db         79 CCTGAAACCATCTCTGAGCCCAAGACC 105

RESULT 5
US-09-869-677B-1
; Sequence 1, Application US/09869677B
; Patent No. 6843999
; GENERAL INFORMATION:
;  APPLICANT: Actinova Limited
;  APPLICANT:  Bjorck, Lars
;  APPLICANT:  Janulczyk, Robert
;  TITLE OF INVENTION: STREPTOCOCCAL ABC TRANSPORTER PROTEIN
;  FILE REFERENCE: KCO1004US
;  CURRENT APPLICATION NUMBER: US/09/869,677B
;  CURRENT FILING DATE:  2001-10-09
;  PRIOR APPLICATION NUMBER: GB 9828880.6
;  PRIOR FILING DATE: 1998-12-31
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;    LENGTH: 893
;    TYPE: DNA
;    ORGANISM: Streptococcus pyogenes
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (1)..(861)
US-09-869-677B-1

Alignment Scores:
Pred. No.:              0.721           Length:         893
Score:                  73.50           Matches:        21
Percent Similarity:     54.2%           Conservative:   18
Best Local Similarity:  29.2%           Mismatches:     18
Query Match:            16.4%           Indels:         15
DB:                     3               Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-09-869-677B-1 (1-893)

Qy         15 GluSerAlaLeuGluArgArgPro---GluThrIleSerGluProLysThrTyrValAsp 33
              ||||||::::::::::||||||||   ||||||:::|||:::      :::   :::
Db        700 GAGTCAAGTGTCGATAGACGCCCTATGGAAACTGTTTCTAAA---GATAGTGGTATTCCT 756

Qy         34 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
              :::   :::|||      |||||||||      |||      |||   :::
Db        757 ATTTATTCTGAGATCTTTACAGATTCAATTGCTAAAAAAGGTAAACCTGGCGAT------ 810

Qy         54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 73
                          |||   ::::::::::::   ||||||:::|||   :::
Db        811 --------------AGTTATTATGCTATGATGAAATGGAACCTTGACAAAATTTCTGAA 855

Qy         74 ------------------AsnAsnLeuSerGluArg 79
                               ||||||   ::::::|||
Db        856 GGTCTAGCAAAATAAGATAACAATAAAGCACAGCGA 891

RESULT 6
US-09-248-796A-1008
; Sequence 1008, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
```

```
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1008
;   LENGTH: 573
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-1008

Alignment Scores:
Pred. No.:              0.541          Length:         573
Score:                  72.50          Matches:        24
Percent Similarity:     50.7%          Conservative:   13
Best Local Similarity:  32.9%          Mismatches:     17
Query Match:            16.2%          Indels:         19
DB:                     3              Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-248-796A-1008 (1-573)

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||:::|||   |||:::|||  |||      |||        :::|||  ||||||
Db        262 CCTAAAACAGCAAGTAAACCA---ACAAGTAGTGATGCTGACGATGAACCAACCACAGGA 318

Qy         42 SerThrThrSerLysIle----------------------------SerProSerGlu 51
              :::||||||   |||                                |||   |||
Db        319 ACTACTACTGATAAGTCAGACACTACCAACGAAGATATTGAACCACAACAACCCCAAGAA 378

Qy         52 AspThrGlnGlnGluAsnGlySerMet--------------PheSerLeuIleThr--- 65
              |||:::::|||:::   :::                   |||   ||||||
Db        379 AAAACTGAGGAAGAAAGTGATAATCAAGATATACCTAATCCAATTTCCTCAATTACCAGT 438

Qy         66 ---TrpAsnIleAspGlyLeuAsp---LeuAsnAsnLeu 76
                 ||||||::::::|||   :::   :::|||:::|||
Db        439 TGGTGGAACCTGGAAGGGTCTAATAAAGTTAATTCACTT 477

RESULT 7
US-09-248-796A-12164
; Sequence 12164, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 12164
;   LENGTH: 999
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-12164

Alignment Scores:
Pred. No.:              1.63           Length:         999
Score:                  71.50          Matches:        20
Percent Similarity:     40.3%          Conservative:   7
Best Local Similarity:  29.9%          Mismatches:     25
Query Match:            16.0%          Indels:         15
DB:                     3              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-248-796A-12164 (1-999)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              ||||||     :::|||:::      |||      ||||||      :::      |||
Db        697 GAACCAGTTAAGAAGGAAACTGGCAAAGAAAGCACACCCGAAAACACTAATTTCAGTAAA 756

Qy         30 ThrTyrVal----------------------------------------------AspLeu 34
              |||                                                    |||
Db        757 GAAAATGTTTTCCAATCACCTCCTGGTTCGTCAACGAAAAAAAGAAAACATCAAGATGCC 816

Qy         35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
              ||||||        :::      ||||||   ||||||:::   |||
Db        817 GAAGAAGAAGAAGATGATAGCACTCTTGTGAAAAAGATTAAACCTTCTCAACGTACTGCT 876

Qy         55 GlnGluAsnGlySerMetPhe 61
              :::||||||   |||:::|||
Db        877 GAAGAAAACAGATCATTATTT 897

RESULT 8
US-09-902-540-5004/c
; Sequence 5004, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5004
;   LENGTH: 24459
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-5004

Alignment Scores:
Pred. No.:              164            Length:         24459
Score:                  71.00          Matches:        27
Percent Similarity:     40.0%          Conservative:   9
Best Local Similarity:  30.0%          Mismatches:     36
Query Match:            15.8%          Indels:         18
DB:                     3              Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-902-540-5004 (1-24459)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              |||||||||          ||||||        |||::::::      :::|||
Db      15374 TTGAACTCG------AGGCCTCCAGAAAGGCCCTCATCCGTCTCACTCAAGCCGAGCGTC 15321

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                 |||             :::   |||                  ||||||
Db      15320 AAATCGAACTTCGACGTCTTCGCAGAC--------------------GACTCCGCCGCC 15282

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              |||         ||||||        |||                        :::   |||
Db      15281 TCCAGTTGCAGCCCCGGCAGGCTCACCGCTCCTCCCGGCGCGTTCTGCAGCACCAACATC 15222

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer------------------Glu 78
              |||||||||          ||||||   |||    :::    |||
Db      15221 ACCTGGAAC---AGCGGACTGCGGCTCAAGTCCCGCTCTGGCTGCAGCTCCTCCACCAAC 15165

Qy         79 ArgAlaArgGlyVal------CysSerTyr 86
              |||:::::|||          |||:::|||
Db      15164 CGCTCGAAGGGCACTTCCTGGTGCGCGTAC 15135

RESULT 9
US-09-902-540-1247
; Sequence 1247, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
```

```
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1247
;   LENGTH: 32241
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(32241)
;   OTHER INFORMATION: unsure at all n locations
US-09-902-540-1247

Alignment Scores:
Pred. No.:              241          Length:         32241
Score:                  71.00        Matches:        27
Percent Similarity:     40.0%        Conservative:   9
Best Local Similarity:  30.0%        Mismatches:     36
Query Match:            15.8%        Indels:         18
DB:                     3            Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-902-540-1247 (1-32241)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              |||||||||          ||||||          |||::::::      :::|||
Db      16868 TTGAACTCG------AGGCCTCCAGAAAGGCCCTCATCCGTCTCACTCAAGCCGAGCGTC 16921

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                |||          :::   |||                          ||||||
Db      16922 AAATCGAACTTCGACGTCTTCGCAGAC--------------------GACTCCGCCGCC 16960

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              |||       ||||||         |||                        :::   |||
Db      16961 TCCAGTTGCAGCCCCGGCAGGCTCACCGCTCCTCCCGGCGCGTTCTGCAGCACCAACATC 17020

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer-----------------Glu 78
              |||||||||       ||||||   |||   :::   |||
Db      17021 ACCTGGAAC---AGCGGACTGCGGCTCAAGTCCCGCTCTGGCTGCAGCTCCTCCACCAAC 17077

Qy         79 ArgAlaArgGlyVal------CysSerTyr 86
              |||::::::|||         |||:::|||
Db      17078 CGCTCGAAGGGCACTTCCTGGTGCGCGTAC 17107

RESULT 10
US-08-961-527-26/c
; Sequence 26, Application US/08961527
; Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,527
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15213 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-527-26

Alignment Scores:
Pred. No.:              100          Length:         15213
Score:                  70.50        Matches:        24
Percent Similarity:     44.1%        Conservative:   17
Best Local Similarity:  25.8%        Mismatches:     27
Query Match:            15.7%        Indels:         25
DB:                     3            Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-08-961-527-26 (1-15213)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              :::||||||||||||||||||          :::      |||   ::::::|||:::
Db      3550 AAACCACCAGTTGAAGAATCAAATCAACCAGAAAAAAACGGAACTGCAACAAAACCAGAA 3491

Qy         30 ThrTyrValAspLeuThrAsnGluGlu-------------------------------- 38
                    :::   |||:::|||
Db      3490 AATTCAGGTAATACAACATCAGAGAATGGACAAACAGAACCAGAACCATCAAACGGAAAT 3431

Qy         39 --------------ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
                           :::|||:::|||   ||||||      :::   :::|||:::
Db      3430 TCAACTGAGGATGTTTCAACCGAATCAAACACATCCAATTCAAATGGAAACGAAGAAATT 3371

Qy         54 GlnGlnGluAsn--------------------------GlySerMetPheSerLeuIle 64
              :::|||||||||                                      :::|||
Db      3370 AAACAAGAAAATGAACTAGACCCTGATAAAAAGGTAGAAGAACCAGAGAAAACACTTGAA 3311

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 77
                 |||:::      |||:::|||   :::||||||
Db      3310 TTAAGAAATGTTTCCGACCTAGAGTTATACAGTTTGTCA 3272

RESULT 11
US-08-961-083-93
; Sequence 93, Application US/08961083
; Patent No. 6159469
;  GENERAL INFORMATION:
;    APPLICANT:  Choi et. al.
;    TITLE OF INVENTION:  Streptococcus pneumoniae Antigens and Vaccines
;    NUMBER OF SEQUENCES:  452
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,083
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
```

```
;      REFERENCE/DOCKET NUMBER:  PB340P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  93:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  835 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-083-93

Alignment Scores:
Pred. No.:              2.08          Length:          835
Score:                  70.00         Matches:         18
Percent Similarity:     48.4%         Conservative:    13
Best Local Similarity:  28.1%         Mismatches:      17
Query Match:            15.6%         Indels:          16
DB:                     3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-08-961-083-93 (1-835)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              :::|||||||||||||||||||          :::      |||   ::::::|||::: 
Db        587 AAACCACCAGTTGAAGAATCAAATCAACCAGAAAAAAACGGAACTGCAACAAAACCAGAA 646

Qy         30 ThrTyrValAspLeuThrAsnGluGlu-------------------------------- 38
                  :::   |||:::|||
Db        647 AATTCAGGTAATACAACATCAGAGAATGGACAAACAGAACCAGAACCATCAAACGGAAAT 706

Qy         39 --------------ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
                           :::|||:::|||   ||||||     :::   :::|||::: 
Db        707 TCAACTGAGGATGTTTCAACCGAATCAAACACATCCAATTCAAATGGAAACGAAGAAATT 766

Qy         54 GlnGlnGluAsn 57
              :::|||||||||
Db        767 AAACAAGAAAAT 778

RESULT 12
US-09-536-784-93
; Sequence 93, Application US/09536784
; Patent No. 6573082
;   GENERAL INFORMATION:
;        APPLICANT: Choi et. al.
;        TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;        NUMBER OF SEQUENCES: 452
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/536,784
;             FILING DATE: 30-Oct-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/961,083
;             FILING DATE: OCT-30-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Michelle S. Marks
;             REGISTRATION NUMBER: 41,971
;             REFERENCE/DOCKET NUMBER: PB340P3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 93:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 835 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 93:
US-09-536-784-93

Alignment Scores:
Pred. No.:              2.08          Length:          835
Score:                  70.00         Matches:         18
Percent Similarity:     48.4%         Conservative:    13
Best Local Similarity:  28.1%         Mismatches:      17
Query Match:            15.6%         Indels:          16
DB:                     3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-536-784-93 (1-835)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              :::|||||||||||||||||||          :::      |||   ::::::|||::: 
Db        587 AAACCACCAGTTGAAGAATCAAATCAACCAGAAAAAAACGGAACTGCAACAAAACCAGAA 646

Qy         30 ThrTyrValAspLeuThrAsnGluGlu-------------------------------- 38
                  :::   |||:::|||
Db        647 AATTCAGGTAATACAACATCAGAGAATGGACAAACAGAACCAGAACCATCAAACGGAAAT 706

Qy         39 --------------ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
                           :::|||:::|||   ||||||     :::   :::|||::: 
Db        707 TCAACTGAGGATGTTTCAACCGAATCAAACACATCCAATTCAAATGGAAACGAAGAAATT 766

Qy         54 GlnGlnGluAsn 57
              :::|||||||||
Db        767 AAACAAGAAAAT 778

RESULT 13
US-09-765-271-93
; Sequence 93, Application US/09765271
; Patent No. 6887663
;   GENERAL INFORMATION:
;        APPLICANT: Choi et. al.
;        TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;        NUMBER OF SEQUENCES: 452
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,271
;             FILING DATE: 22-Jan-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/536,784
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/961,083
;             FILING DATE: OCT-30-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Michelle S. Marks
;             REGISTRATION NUMBER: 41,971
;             REFERENCE/DOCKET NUMBER: PB340P3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 93:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 835 base pairs
```

```
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 93:
US-09-765-271-93

Alignment Scores:
Pred. No.:              2.08          Length:        835
Score:                  70.00         Matches:       18
Percent Similarity:     48.4%         Conservative:  13
Best Local Similarity:  28.1%         Mismatches:    17
Query Match:            15.6%         Indels:        16
DB:                     3             Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-765-271-93 (1-835)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              :::|||||||||||||||||||          :::      |||   ::::::|||::: 
Db        587 AAACCACCAGTTGAAGAATCAAATCAACCAGAAAAAAACGGAACTGCAACAAAACCAGAA 646

Qy         30 ThrTyrValAspLeuThrAsnGluGlu-------------------------------- 38
                      :::   |||:::|||
Db        647 AATTCAGGTAATACAACATCAGAGAATGGACAAACAGAACCAGAACCATCAAACGGAAAT 706

Qy         39 ---------------ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
                         :::|||:::|||   ||||||     :::   :::|||:::
Db        707 TCAACTGAGGATGTTTCAACCGAATCAAACACATCCAATTCAAATGGAAACGAAGAAATT 766

Qy         54 GlnGlnGluAsn 57
              :::|||||||||
Db        767 AAACAAGAAAAT 778

RESULT 14
US-09-765-272A-93
; Sequence 93, Application US/09765272A
; Patent No. 6929930
;   GENERAL INFORMATION:
;        APPLICANT: Choi et. al.
;        TITLE OF INVENTION: Streptococcus pneumoniae Antigens and
;                            Vaccines
;        NUMBER OF SEQUENCES: 454
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: Dell Latitude C610
;             OPERATING SYSTEM: Windows 2000
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,272A
;             FILING DATE: 22-Jan-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/961,083
;             FILING DATE: OCT-30-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lin J. Hymel
;             REGISTRATION NUMBER: 45,414
;             REFERENCE/DOCKET NUMBER: PB340P2C2
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 610-5790
;             TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 93:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 835 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 93:
US-09-765-272A-93

Alignment Scores:
Pred. No.:              2.08          Length:        835
Score:                  70.00         Matches:       18
Percent Similarity:     48.4%         Conservative:  13
Best Local Similarity:  28.1%         Mismatches:    17
Query Match:            15.6%         Indels:        16
DB:                     3             Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-765-272A-93 (1-835)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              :::|||||||||||||||||||          :::      |||   ::::::|||::: 
Db        587 AAACCACCAGTTGAAGAATCAAATCAACCAGAAAAAAACGGAACTGCAACAAAACCAGAA 646

Qy         30 ThrTyrValAspLeuThrAsnGluGlu-------------------------------- 38
                      :::   |||:::|||
Db        647 AATTCAGGTAATACAACATCAGAGAATGGACAAACAGAACCAGAACCATCAAACGGAAAT 706

Qy         39 ---------------ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
                         :::|||:::|||   ||||||     :::   :::|||:::
Db        707 TCAACTGAGGATGTTTCAACCGAATCAAACACATCCAATTCAAATGGAAACGAAGAAATT 766

Qy         54 GlnGlnGluAsn 57
              :::|||||||||
Db        767 AAACAAGAAAAT 778

RESULT 15
US-08-961-083-3
; Sequence 3, Application US/08961083
; Patent No. 6159469
;  GENERAL INFORMATION:
;    APPLICANT:  Choi et. al.
;    TITLE OF INVENTION:  Streptococcus pneumoniae Antigens and Vaccines
;    NUMBER OF SEQUENCES:  452
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,083
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1714 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-083-3

Alignment Scores:
Pred. No.:              5.67          Length:        1714
```

```
Score:                  70.00          Matches:        18
Percent Similarity:     48.4%          Conservative:   13
Best Local Similarity:  28.1%          Mismatches:     17
Query Match:            15.6%          Indels:         16
DB:                     3              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-08-961-083-3 (1-1714)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              :::|||||||||||||||||||          :::      |||   ::::::|||::: 
Db       1448 AAACCACCAGTTGAAGAATCAAATCAACCAGAAAAAAACGGAACTGCAACAAAACCAGAA 1507

Qy         30 ThrTyrValAspLeuThrAsnGluGlu-------------------------------- 38
                    :::   |||:::|||
Db       1508 AATTCAGGTAATACAACATCAGAGAATGGACAAACAGAACCAGAACCATCAAACGGAAAT 1567

Qy         39 --------------ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
                           :::|||:::|||   ||||||      :::   :::|||::: 
Db       1568 TCAACTGAGGATGTTTCAACCGAATCAAACACATCCAATTCAAATGGAAACGAAGAAATT 1627

Qy         54 GlnGlnGluAsn 57
              :::|||||||||
Db       1628 AAACAAGAAAAT 1639

RESULT 16
US-09-536-784-3
; Sequence 3, Application US/09536784
; Patent No. 6573082
;   GENERAL INFORMATION:
;        APPLICANT: Choi et. al.
;        TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;        NUMBER OF SEQUENCES: 452
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/536,784
;             FILING DATE: 30-Oct-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/961,083
;             FILING DATE: OCT-30-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Michelle S. Marks
;             REGISTRATION NUMBER: 41,971
;             REFERENCE/DOCKET NUMBER: PB340P3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1714 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-536-784-3

Alignment Scores:
Pred. No.:              5.67           Length:         1714
Score:                  70.00          Matches:        18
Percent Similarity:     48.4%          Conservative:   13
Best Local Similarity:  28.1%          Mismatches:     17
Query Match:            15.6%          Indels:         16
DB:                     3              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-536-784-3 (1-1714)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              :::|||||||||||||||||||          :::      |||   ::::::|||::: 
Db       1448 AAACCACCAGTTGAAGAATCAAATCAACCAGAAAAAAACGGAACTGCAACAAAACCAGAA 1507

Qy         30 ThrTyrValAspLeuThrAsnGluGlu-------------------------------- 38
                    :::   |||:::|||
Db       1508 AATTCAGGTAATACAACATCAGAGAATGGACAAACAGAACCAGAACCATCAAACGGAAAT 1567

Qy         39 --------------ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
                           :::|||:::|||   ||||||      :::   :::|||::: 
Db       1568 TCAACTGAGGATGTTTCAACCGAATCAAACACATCCAATTCAAATGGAAACGAAGAAATT 1627

Qy         54 GlnGlnGluAsn 57
              :::|||||||||
Db       1628 AAACAAGAAAAT 1639

RESULT 17
US-09-765-271-3
; Sequence 3, Application US/09765271
; Patent No. 6887663
;   GENERAL INFORMATION:
;        APPLICANT: Choi et. al.
;        TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;        NUMBER OF SEQUENCES: 452
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,271
;             FILING DATE: 22-Jan-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/536,784
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/961,083
;             FILING DATE: OCT-30-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Michelle S. Marks
;             REGISTRATION NUMBER: 41,971
;             REFERENCE/DOCKET NUMBER: PB340P3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1714 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-765-271-3

Alignment Scores:
Pred. No.:              5.67           Length:         1714
Score:                  70.00          Matches:        18
Percent Similarity:     48.4%          Conservative:   13
Best Local Similarity:  28.1%          Mismatches:     17
Query Match:            15.6%          Indels:         16
DB:                     3              Gaps:           1
```

```
US-10-757-745-2_COPY_54_140 (1-87) x US-09-765-271-3 (1-1714)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              :::||||||||||||||||||           :::      |||   ::::::|||:::
Db       1448 AAACCACCAGTTGAAGAATCAAATCAACCAGAAAAAAACGGAACTGCAACAAAACCAGAA 1507

Qy         30 ThrTyrValAspLeuThrAsnGluGlu------------------------------- 38
                    :::   |||:::|||
Db       1508 AATTCAGGTAATACAACATCAGAGAATGGACAAACAGAACCAGAACCATCAAACGGAAAT 1567

Qy         39 --------------ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
                           :::|||:::|||   ||||||      :::   :::|||:::
Db       1568 TCAACTGAGGATGTTTCAACCGAATCAAACACATCCAATTCAAATGGAAACGAAGAAATT 1627

Qy         54 GlnGlnGluAsn 57
              :::|||||||||
Db       1628 AAACAAGAAAAT 1639

RESULT 18
US-09-765-272A-3
; Sequence 3, Application US/09765272A
; Patent No. 6929930
;   GENERAL INFORMATION:
;        APPLICANT: Choi et. al.
;        TITLE OF INVENTION: Streptococcus pneumoniae Antigens and
;                            Vaccines
;        NUMBER OF SEQUENCES: 454
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: Dell Latitude C610
;             OPERATING SYSTEM: Windows 2000
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,272A
;             FILING DATE: 22-Jan-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/961,083
;             FILING DATE: OCT-30-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lin J. Hymel
;             REGISTRATION NUMBER: 45,414
;             REFERENCE/DOCKET NUMBER: PB340P2C2
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 610-5790
;             TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1714 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-765-272A-3

Alignment Scores:
Pred. No.:              5.67          Length:         1714
Score:                  70.00         Matches:        18
Percent Similarity:     48.4%         Conservative:   13
Best Local Similarity:  28.1%         Mismatches:     17
Query Match:            15.6%         Indels:         16
DB:                     3             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-765-272A-3 (1-1714)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              :::||||||||||||||||||           :::      |||   ::::::|||:::
Db       1448 AAACCACCAGTTGAAGAATCAAATCAACCAGAAAAAAACGGAACTGCAACAAAACCAGAA 1507

Qy         30 ThrTyrValAspLeuThrAsnGluGlu------------------------------- 38
                    :::   |||:::|||
Db       1508 AATTCAGGTAATACAACATCAGAGAATGGACAAACAGAACCAGAACCATCAAACGGAAAT 1567

Qy         39 --------------ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
                           :::|||:::|||   ||||||      :::   :::|||:::
Db       1568 TCAACTGAGGATGTTTCAACCGAATCAAACACATCCAATTCAAATGGAAACGAAGAAATT 1627

Qy         54 GlnGlnGluAsn 57
              :::|||||||||
Db       1628 AAACAAGAAAAT 1639

RESULT 19
US-09-976-594-888
; Sequence 888, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 888
;   LENGTH: 3640
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 5202390CB1
US-09-976-594-888

Alignment Scores:
Pred. No.:              16.2          Length:         3640
Score:                  70.00         Matches:        33
Percent Similarity:     39.8%         Conservative:   12
Best Local Similarity:  29.2%         Mismatches:     32
Query Match:            15.6%         Indels:         36
DB:                     3             Gaps:           8

US-10-757-745-2_COPY_54_140 (1-87) x US-09-976-594-888 (1-3640)

Qy          6 AsnSerTyrPheGluPro--------------------------------ProValGlu 14
              |||   ::::::||||||                                |||||:::
Db        827 AATCCTTTTTATGAACCTAAATCAACTCCTCCTCCAAATAATTTGGTAAATCCTGTTCAA 886

Qy         15 GluSerAlaLeuGluArgArg--------------------ProGluThrIleSerGlu 27
              |||          ||||||||                    |||      :::|||
Db        887 GAACTAGAAACTGAAAGGCGAGTGAAAAGAAAGGCCCCGGCTCCACCAGTCCTCTCA--- 943

Qy         28 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp--------------Ser 42
              |||||||||   |||         ||||||   |||                    |||
Db        944 CCAAAAACAGGAGTA------TTAAATGAAAACACAGTTTCTGCAGGAAAAGATCTCTCT 997

Qy         43 ThrThrSerLysIleSerProSerGluAspThr---------GlnGlnGluAsnGlySer 59
              |||:::   |||   ||||||                    ::::::   |||   |||
Db        998 ACTTCTCCTAAGCCAAGCCCTATACCAAGTCCTGTTTTGGGGCGAAAGCCAAATGCTAGT 1057

Qy         60 MetPheSerLeuIleThrTrp-----------------AsnIleAspGlyLeuAspLeu 73
                    ||||||:::   |||                 |||     |||:::   :::
Db       1058 CAG---TCTTTGCTTGTATGGTGTAAAGAAGTTACAAAGAACTACCGAGGAGTAAAAATC 1114

Qy         74 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
              |||   :::          |||          |||:::
Db       1115 ACCAATTTTACTACATCGTGGAGAAATGGTTTATCTTTT 1153
```

```
RESULT 20
US-09-620-312D-363
; Sequence 363, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 363
;   LENGTH: 4061
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (72)..(3062)
US-09-620-312D-363

Alignment Scores:
Pred. No.:                18.8          Length:          4061
Score:                    70.00         Matches:         33
Percent Similarity:       39.8%         Conservative:    12
Best Local Similarity:    29.2%         Mismatches:      32
Query Match:              15.6%         Indels:          36
DB:                       3             Gaps:            8

US-10-757-745-2_COPY_54_140 (1-87) x US-09-620-312D-363 (1-4061)

Qy          6 AsnSerTyrPheGluPro---------------------------------ProValGlu 14
              |||   ::::::||||||                                 ||||||:::
Db        474 AATCCTTTTTATGAACCTAAATCAACTCCTCCTCCAAATAATTTGGTAAATCCTGTTCAA 533

Qy         15 GluSerAlaLeuGluArgArg--------------------ProGluThrIleSerGlu 27
              |||         |||||||||                    |||     :::|||
Db        534 GAACTAGAAACTGAAAGGCGAGTGAAAAGAAAGGCCCCGGCTCCACCAGTCCTCTCA--- 590

Qy         28 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp--------------Ser 42
              |||||||||   |||        ||||||   |||                     |||
Db        591 CCAAAAACAGGAGTA------TTAAATGAAAACACAGTTTCTGCAGGAAAAGATCTCTCT 644

Qy         43 ThrThrSerLysIleSerProSerGluAspThr---------GlnGlnGluAsnGlySer 59
              |||:::   |||   ||||||                    ::::::   |||   |||
Db        645 ACTTCTCCTAAGCCAAGCCCTATACCAAGTCCTGTTTTGGGGCGAAAGCCAAATGCTAGT 704

Qy         60 MetPheSerLeuIleThrTrp-----------------AsnIleAspGlyLeuAspLeu 73
                 ||||||:::   |||                   |||     |||:::   :::
Db        705 CAG---TCTTTGCTTGTATGGTGTAAAGAAGTTACAAAGAACTACCGAGGAGTAAAAATC 761

Qy         74 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                |||   :::          |||         |||:::
Db        762 ACCAATTTTACTACATCGTGGAGAAATGGTTTATCTTTT 800

RESULT 21
US-09-248-796A-13914
; Sequence 13914, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 13914
;   LENGTH: 486
;   TYPE: DNA
;   ORGANISM: Candida albicans
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (486)
;   OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno.
US-09-248-796A-13914

Alignment Scores:
Pred. No.:                1.16          Length:          486
Score:                    69.50         Matches:         22
Percent Similarity:       47.0%         Conservative:    9
Best Local Similarity:    33.3%         Mismatches:      30
Query Match:              15.5%         Indels:          5
DB:                       3             Gaps:            2

US-10-757-745-2_COPY_54_140 (1-87) x US-09-248-796A-13914 (1-486)

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||   ||||||||||||                |||     |||   :::|||   ||||||
Db         46 CCTCCTACTATTTCCGAA------------GTTTCATCCACTACCAAAGAGAATACAGAT 93

Qy         42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
              ::::::   ||||||          :::|||     :::|||     |||
Db         94 AATTCTAATTCCAAA---CAAGAGGAGAAGGATAATGATAAGGAGATATCATCTACAATT 150

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
                 |||:::     |||     |||   |||     |||
Db        151 GAAAATCCCACATTTGTTGACGATGAAGAAGATAATAATCCATTTTCACACCATGGTGAA 210

Qy         82 GlyValCysSerTyrLeu 87
              |||:::   |||::::::
Db        211 GGATTAACTAGTTTTATG 228

RESULT 22
US-09-543-681A-1380
; Sequence 1380, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1380
;   LENGTH: 231
;   TYPE: DNA
;   ORGANISM: Proteus mirabilis
```

```
US-09-543-681A-1380

Alignment Scores:
Pred. No.:                0.486          Length:         231
Score:                    69.00          Matches:        23
Percent Similarity:       46.9%          Conservative:   15
Best Local Similarity:    28.4%          Mismatches:     29
Query Match:              15.4%          Indels:         14
DB:                       3              Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-09-543-681A-1380 (1-231)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||   |||   |||                     |||:::   |||   |||
Db         10 ATGAATCCTTACCATGAA--------------------CTTGACGCAAGGGAAGAGCGC 48

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                 |||     :::::::::|||     :::   |||   ::::::
Db         49 AAACAAGAAGAAGCATCATGGATTGATGCAAAGGATGCTGAGTTAAGCAATGTCGCATTT 108

Qy         45 SerLysIleSer------ProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
              :::   :::         |||   :::   |||   |||   :::   |||:::|||
Db        109 AATGTGGTTGATGGATTACCAAAGGATATTACCAGTCAATGGAGTGACAGCGTTTTCGAT 168

Qy         63 LeuIleThrTrpAsnIleAspGlyLeu------AspLeuAsnAsnLeuSerGluArgAla 80
              :::           ||||||||||||      :::|||   |||      ||||||
Db        169 ATG---------ACAATTGATGGTCTTTACAAGGAGTTAAAGAACTATCAGGAACGTAGG 219

Qy         81 Arg 81
              |||
Db        220 AGG 222

RESULT 23
US-08-894-699-1
; Sequence 1, Application US/08894699
; Patent No. 6030769
;  GENERAL INFORMATION:
;    APPLICANT:  SIMON, FRANCOIS
;    APPLICANT:  SARAGOSTI, SENTOB
;    APPLICANT:  LOUSSERT-AJAKA, IBITISSAM
;    APPLICANT:  LY, THOAI-DUONG
;    APPLICANT:  CHAIX-BAUDIER, MARIE-LAURE
;    TITLE OF INVENTION:  GROUP O HIV-1, FRAGMENTS OF SUCH
;    TITLE OF INVENTION:  VIRUSES, AND USES THEREOF
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
;      STREET:  FLOOR
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/894,699
;      FILING DATE:  01-DEC-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR96/00294
;      FILING DATE:  26-FEB-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 95/02236
;      FILING DATE:  27-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,614
;      REFERENCE/DOCKET NUMBER:  0354-0020-0PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  291 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-894-699-1

Alignment Scores:
Pred. No.:                0.669          Length:         291
Score:                    69.00          Matches:        26
Percent Similarity:       37.5%          Conservative:   10
Best Local Similarity:    27.1%          Mismatches:     31
Query Match:              15.4%          Indels:         29
DB:                       3              Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-08-894-699-1 (1-291)

Qy          6 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg---------- 21
              ||||||   :::         ::::::   :::   |||   |||:::
Db         30 AACAGTAAGTACTCAGCTAATATTAAATGGAACACTCTCAGAAGGAAAGATAAGAATGAT 89

Qy         22 ------------------------ProGluThrIleSerGluProLysThrTyrVal--- 32
                                        |||:::     |||:::||||||   |||
Db         90 GGCAAAAAATATTTCGGATAGTGGCCAAAATATCATAGTGACCCTAAATACTACTATAAA 149

Qy         33 ---AspLeuThrAsnGluGluThrThrAspSerThrThrSerLys-IleSerProSerGl 51
                 ||||||           |||::::::||||||      |||  |||   |||
Db        150 CATGACCTGCCAGAGACCAGGACATCAAACAGTACAAGAGATAAGGATAGGTCCA----- 204

Qy         51 uAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLe 71
                                                 :::   |||         |||||
Db        205 ----------------------------------ATGGCCTGGTACAGCATGGGCTT 227

Qy         71 uAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
              |         |||   ||||||      |||     |||   |||
Db        228 AGCGGCAGGAAACGGATCTGAGTCAAGAAGAGCTTATTGTGAATAT 273

RESULT 24
US-09-444-410-1
; Sequence 1, Application US/09444410
; Patent No. 6270975
;  GENERAL INFORMATION:
;    APPLICANT:  SIMON, FRANCOIS
;    APPLICANT:  SARAGOSTI, SENTOB
;    APPLICANT:  LOUSSERT-AJAKA, IBITISSAM
;    APPLICANT:  LY, THOAI-DUONG
;    APPLICANT:  CHAIX-BAUDIER, MARIE-LAURE
;    TITLE OF INVENTION:  GROUP O HIV-1, FRAGMENTS OF SUCH
;    TITLE OF INVENTION:  VIRUSES, AND USES THEREOF
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
;      STREET:  FLOOR
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/444,410
;      FILING DATE:
;      CLASSIFICATION:
```

```
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/894,699
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 95/02236
;       FILING DATE:  27-FEB-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  OBLON, NORMAN F.
;       REGISTRATION NUMBER:  24,614
;       REFERENCE/DOCKET NUMBER:  0354-0020-0PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-413-3000
;       TELEFAX:  703-413-2220
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  291 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-09-444-410-1

Alignment Scores:
Pred. No.:                0.669          Length:          291
Score:                    69.00          Matches:         26
Percent Similarity:       37.5%          Conservative:    10
Best Local Similarity:    27.1%          Mismatches:      31
Query Match:              15.4%          Indels:          29
DB:                       3              Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-444-410-1 (1-291)

Qy           6 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg------------ 21
               ||||||    :::      ::::::   :::   |||   |||:::
Db          30 AACAGTAAGTACTCAGCTAATATTAAATGGAACACTCTCAGAAGGAAAGATAAGAATGAT 89

Qy          22 ------------------------ProGluThrIleSerGluProLysThrTyrVal--- 32
                                       |||:::      |||:::||||||   |||
Db          90 GGCAAAAAATATTTCGGATAGTGGCCAAAATATCATAGTGACCCTAAATACTACTATAAA 149

Qy          33 ---AspLeuThrAsnGluGluThrThrAspSerThrThrSerLys-IleSerProSerGl 51
                  ||||||          |||::::::||||||       |||  |||   |||
Db         150 CATGACCTGCCAGAGACCAGGACATCAAACAGTACAAGAGATAAGGATAGGTCCA----- 204

Qy          51 uAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLe 71
                                                 :::      |||      |||||
Db         205 -----------------------------------ATGGCCTGGTACAGCATGGGCTT 227

Qy          71 uAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                 |   |||   ||||||       |||       |||   |||
Db         228 AGCGGCAGGAAACGGATCTGAGTCAAGAAGAGCTTATTGTGAATAT 273

RESULT 25
PCT-US93-05944-1
; Sequence 1, Application PC/TUS9305944
;  GENERAL INFORMATION:
;    APPLICANT:  Lin et al., Hun-Chi
;    TITLE OF INVENTION:  Molecular cloning of the genes
;    TITLE OF INVENTION:  responsible for collagenase product
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Harris Brotman
;      STREET:  401 B. St Ste 1700
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101-4297
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/05944
;      FILING DATE:  19930622
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brotman, Harris F.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 699-3630
;      TELEFAX:  (619) 236-1048
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2817 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..2808
PCT-US93-05944-1

Alignment Scores:
Pred. No.:                21.9           Length:          2817
Score:                    68.00          Matches:         17
Percent Similarity:       60.8%          Conservative:    14
Best Local Similarity:    33.3%          Mismatches:      16
Query Match:              15.2%          Indels:          4
DB:                       6              Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x PCT-US93-05944-1 (1-2817)

Qy          24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
               |||   :::|||       |||   :::::::::   ||||||:::||||||
Db        2059 ACAGCTACAGAAAGCTTTACT---ATAGAAATAAAGAACGAAGATACAACAACACCTATA 2115

Qy          44 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 63
               |||    ::::::   |||::::::|||   ::::::   ||||||   :::
Db        2116 ACTAAAGAAATGGAACCTAATGATGATATAAAAGAGGCTAATGGTCCAATAGTTGAAGGT 2175

Qy          64 IleThrTrpAsnIleAspGlyLeuAspLeuAsn 74
               :::|||       :::   |||   |||||||||
Db        2176 GTTACT------GTAAAAGGT---GATTTAAAT 2199

RESULT 26
US-09-379-523-4
; Sequence 4, Application US/09379523
; Patent No. 6280993
;  GENERAL INFORMATION:
;    APPLICANT:  YAMATO, Ichiro
;    APPLICANT:  HOSAKA, Toshiaki
;    TITLE OF INVENTION:  GENE ENCODING CLASS I COLLAGENASE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burgess, Ryan and Wayne
;      STREET:  370 Lexington Avenue, Suite 2105
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  10017
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3 1/2 inch disk
;      COMPUTER:  PC'S LIMITED SYSTEM 200
;      OPERATING SYSTEM:  WINDOWS 98
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/379,523
;      FILING DATE:  August 24, 1999
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wayne, Milton J.
;      REGISTRATION NUMBER:  17,906
```

```
;      REFERENCE/DOCKET NUMBER:  U-Wp-5462
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-683-8150
;      TELEFAX:  212-532-4285
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3356
;      TYPE:  nucleotide
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-379-523-4

Alignment Scores:
Pred. No.:              27.9          Length:         3356
Score:                  68.00         Matches:        17
Percent Similarity:     60.8%         Conservative:   14
Best Local Similarity:  33.3%         Mismatches:     16
Query Match:            15.2%         Indels:         4
DB:                     3             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-379-523-4 (1-3356)

Qy         24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
              |||   :::|||      |||   :::::::::   ||||||:::||||||
Db       2605 ACAGCTACAGAAAGCTTTACT---ATAGAAATAAAGAACGAAGATACAACAACACCTATA 2661

Qy         44 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 63
              |||   ::::::   |||:::::|||   ::::::   ||||||   :::
Db       2662 ACTAAAGAAATGGAACCTAATGATGATATAAAAGAGGCTAATGGTCCAATAGTTGAAGGT 2721

Qy         64 IleThrTrpAsnIleAspGlyLeuAspLeuAsn 74
              :::|||      :::   |||   |||||||||
Db       2722 GTTACT------GTAAAAGGT---GATTTAAAT 2745

RESULT 27
US-09-308-453-1
; Sequence 1, Application US/09308453
; Patent No. 6475764
; GENERAL INFORMATION:
;  APPLICANT: Roche Diagnostics GMBH
;  TITLE OF INVENTION: Recombinant collagenase type I from Clostridium histolyticum and
;  TITLE OF INVENTION:  for isolating cells and groups of cells
;  FILE REFERENCE: BMID9924US
;  CURRENT APPLICATION NUMBER: US/09/308,453
;  CURRENT FILING DATE:  1999-09-09
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 4358
;   TYPE: DNA
;   ORGANISM: Clostridium histolyticum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1332)..(4358)
US-09-308-453-1

Alignment Scores:
Pred. No.:              40.1          Length:         4358
Score:                  68.00         Matches:        17
Percent Similarity:     60.8%         Conservative:   14
Best Local Similarity:  33.3%         Mismatches:     16
Query Match:            15.2%         Indels:         4
DB:                     3             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-308-453-1 (1-4358)

Qy         24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
              |||   :::|||      |||   :::::::::   ||||||:::||||||
Db       3606 ACAGCTACAGAAAGCTTTACT---ATAGAAATAAAGAACGAAGATACAACAACACCTATA 3662

Qy         44 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 63
              |||   ::::::   |||:::::|||   ::::::   ||||||   :::
Db       3663 ACTAAAGAAATGGAACCTAATGATGATATAAAAGAGGCTAATGGTCCAATAGTTGAAGGT 3722

Qy         64 IleThrTrpAsnIleAspGlyLeuAspLeuAsn 74
              :::|||      :::   |||   |||||||||
Db       3723 GTTACT------GTAAAAGGT---GATTTAAAT 3746

RESULT 28
US-08-095-737-3
; Sequence 3, Application US/08095737
; Patent No. 5487979
;  GENERAL INFORMATION:
;    APPLICANT:  DiFiore, Pier P
;    APPLICANT:  Fazioli, Francesca
;    TITLE OF INVENTION:  A Substrate for the Epidermal Growth
;    TITLE OF INVENTION:  Factor Receptor Kinase
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  620 Newport Center Drive, Sixteenth Floor
;      CITY:  Newport Beach
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/095,737
;      FILING DATE:  19930722
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned A
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  NIH060.001A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3033 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  111..2802
US-08-095-737-3

Alignment Scores:
Pred. No.:              28.6          Length:         3033
Score:                  67.50         Matches:        27
Percent Similarity:     37.5%         Conservative:   12
Best Local Similarity:  26.0%         Mismatches:     36
Query Match:            15.1%         Indels:         29
DB:                     2             Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-08-095-737-3 (1-3033)

Qy          5 LeuAsnSerTyrPheGluPro---ProValGluGluSerAlaLeuGluArgArgProGlu 23
              ||||||      |||:::|||   |||      :::|||      |||:::   |||:::
Db       2520 CTGAATGATCCATTTCAGCCTTTCCCAGGCAATGATAGTCCCAAAGAAAAAGATCCTGAT 2579

Qy         24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
                    :::|||   |||            ||||||:::|||         |||
Db       2580 ATGTTTTGTGATCCATTCACTTCTTCTACCACTACCAATAAAGAGGCTGACCCAAGCAAT 2639

Qy         44 ThrSerLysIleSer------ProSerGluAsp-------------------------- 52
                 :::      |||      |||||||||:::
```

```
Db       2640 TTTGCTAACTTCAGTGCTTATCCCTCTGAAGAAGATATGATTGAATGGGCAAAAAGGGAA 2699

Qy         53 ------------------------------------ThrGlnGlnGluAsnGlySerMet 60
                                                    ||||||||          :::
Db       2700 AGTGAGCGGGAAGAAGAACAGAGGCTTGCCAGACTAAATCAGCAGGAGCAAGAAGACTTG 2759

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                  |||                  |||      :::   :::||||||
Db       2760 GAACTGGCCATT--------------GCACTTAGCAAATCTGAGATCTCAGAAGCATGA 2804

Qy         81 ArgGlyValCys 84
              |||   :::|||
Db       2805 AGAGTTATCTGT 2816

RESULT 29
US-08-480-145-3
; Sequence 3, Application US/08480145
; Patent No. 5717067
;  GENERAL INFORMATION:
;    APPLICANT:  DiFiore, Pier P
;    APPLICANT:  Fazioli, Francesca
;    TITLE OF INVENTION:  A Substrate for the Epidermal Growth
;    TITLE OF INVENTION:  Factor Receptor Kinase
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  620 Newport Center Drive, Sixteenth Floor
;      CITY:  Newport Beach
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,145
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/095,737
;      FILING DATE:  22-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned A
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  NIH060.001A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3033 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  111..2802
US-08-480-145-3

Alignment Scores:
Pred. No.:                28.6          Length:         3033
Score:                    67.50         Matches:        27
Percent Similarity:       37.5%         Conservative:   12
Best Local Similarity:    26.0%         Mismatches:     36
Query Match:              15.1%         Indels:         29
DB:                       2             Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-08-480-145-3 (1-3033)
```

```
Qy          5 LeuAsnSerTyrPheGluPro---ProValGluGluSerAlaLeuGluArgArgProGlu 23
              ||||||      |||:::|||   |||      :::|||      |||:::   |||:::
Db       2520 CTGAATGATCCATTTCAGCCTTTCCCAGGCAATGATAGTCCCAAAGAAAAAGATCCTGAT 2579

Qy         24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
                  :::|||   |||             ||||||:::|||          |||
Db       2580 ATGTTTTGTGATCCATTCACTTCTTCTACCACTACCAATAAAGAGGCTGACCCAAGCAAT 2639

Qy         44 ThrSerLysIleSer------ProSerGluAsp-------------------------- 52
                 :::      |||      |||||||||:::
Db       2640 TTTGCTAACTTCAGTGCTTATCCCTCTGAAGAAGATATGATTGAATGGGCAAAAAGGGAA 2699

Qy         53 ------------------------------------ThrGlnGlnGluAsnGlySerMet 60
                                                    |||||||||          :::
Db       2700 AGTGAGCGGGAAGAAGAACAGAGGCTTGCCAGACTAAATCAGCAGGAGCAAGAAGACTTG 2759

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                  |||                  |||      :::   :::||||||
Db       2760 GAACTGGCCATT--------------GCACTTAGCAAATCTGAGATCTCAGAAGCATGA 2804

Qy         81 ArgGlyValCys 84
              |||   :::|||
Db       2805 AGAGTTATCTGT 2816

RESULT 30
US-08-477-389-3
; Sequence 3, Application US/08477389
; Patent No. 5872219
;  GENERAL INFORMATION:
;    APPLICANT:  DiFiore, Pier P
;    APPLICANT:  Fazioli, Francesca
;    TITLE OF INVENTION:  A Substrate for the Epidermal Growth
;    TITLE OF INVENTION:  Factor Receptor Kinase
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  620 Newport Center Drive, Sixteenth Floor
;      CITY:  Newport Beach
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,389
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/095,737
;      FILING DATE:  22-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned A
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  NIH060.001A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3033 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  111..2802
```

```
US-08-477-389-3

Alignment Scores:
Pred. No.:                28.6          Length:         3033
Score:                    67.50         Matches:        27
Percent Similarity:       37.5%         Conservative:   12
Best Local Similarity:    26.0%         Mismatches:     36
Query Match:              15.1%         Indels:         29
DB:                       2             Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-08-477-389-3 (1-3033)

Qy          5 LeuAsnSerTyrPheGluPro---ProValGluGluSerAlaLeuGluArgArgProGlu 23
              ||||||      |||:::|||   |||      :::|||      |||:::   |||:::
Db       2520 CTGAATGATCCATTTCAGCCTTTCCCAGGCAATGATAGTCCCAAAGAAAAAGATCCTGAT 2579

Qy         24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
                :::|||  |||           ||||||:::|||         |||
Db       2580 ATGTTTTGTGATCCATTCACTTCTTCTACCACTACCAATAAAGAGGCTGACCCAAGCAAT 2639

Qy         44 ThrSerLysIleSer------ProSerGluAsp-------------------------- 52
                :::      |||      |||||||||:::
Db       2640 TTTGCTAACTTCAGTGCTTATCCCTCTGAAGAAGATATGATTGAATGGGCAAAAAGGGAA 2699

Qy         53 ----------------------------------ThrGlnGlnGluAsnGlySerMet 60
                                                |||||||||          :::
Db       2700 AGTGAGCGGGAAGAAGAACAGAGGCTTGCCAGACTAAATCAGCAGGAGCAAGAAGACTTG 2759

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                  |||                 |||      :::   :::||||||
Db       2760 GAACTGGCCATT--------------GCACTTAGCAAATCTGAGATCTCAGAAGCATGA 2804

Qy         81 ArgGlyValCys 84
              |||   :::|||
Db       2805 AGAGTTATCTGT 2816

RESULT 31
US-08-956-171E-166/c
; Sequence 166, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/956,171E
;             FILING DATE: 20-Oct-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mark J. Hyman
;             REGISTRATION NUMBER: 46,789
;             REFERENCE/DOCKET NUMBER: PB248P1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (240) 314-1224
;             TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 166:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9623 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 166:
US-08-956-171E-166

Alignment Scores:
Pred. No.:                142           Length:         9623
Score:                    67.50         Matches:        19
Percent Similarity:       47.5%         Conservative:   9
Best Local Similarity:    32.2%         Mismatches:     22
Query Match:              15.1%         Indels:         9
DB:                       3             Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-08-956-171E-166 (1-9623)

Qy         24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
              |||:::      |||      |||   |||          ::::::||||||      |||
Db       1843 ACTTTAGTTTCACCTACCGCTTATGCTGATACACCTCAAAAAGATACTACAGCTAAGACA 1784

Qy         44 Thr---------SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||         ||||||   |||      ::::::|||   ::::::      |||
Db       1783 ACATCTCATGATTCAAAAAAATCTAATGACGATGAAACTTCTAAGGATACTACAAGTAAA 1724

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                              :::||||||      |||   ||||||   |||   :::
Db       1723 -----------------GATATTGATAAAGCAGACAAAAATAATACAAGTAACCAA 1685

RESULT 32
US-08-781-986A-166/c
; Sequence 166, Application US/08781986A
; Patent No. 6737248
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  166:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9623 base pairs
;      TYPE:  nucleic acid
```

```
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-166

Alignment Scores:
Pred. No.:                142             Length:          9623
Score:                    67.50           Matches:         19
Percent Similarity:       47.5%           Conservative:    9
Best Local Similarity:    32.2%           Mismatches:      22
Query Match:              15.1%           Indels:          9
DB:                       3               Gaps:            2

US-10-757-745-2_COPY_54_140 (1-87) x US-08-781-986A-166 (1-9623)

Qy          24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
               |||:::      |||      |||   |||         ::::::||||||      |||
Db        1843 ACTTTAGTTTCACCTACCGCTTATGCTGATACACCTCAAAAAGATACTACAGCTAAGACA 1784

Qy          44 Thr---------SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               |||         ||||||  |||     :::::::|||   ::::::       |||
Db        1783 ACATCTCATGATTCAAAAAAATCTAATGACGATGAAACTTCTAAGGATACTACAAGTAAA 1724

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                           :::||||||      |||   ||||||   |||    :::
Db        1723 ------------------GATATTGATAAAGCAGACAAAAATAATACAAGTAACCAA 1685

RESULT 33
US-09-280-116-22
; Sequence 22, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
;  FILE REFERENCE: 5800-24, 035800/176965
;  CURRENT APPLICATION NUMBER: US/09/280,116A
;  CURRENT FILING DATE: 1999-03-26
;  NUMBER OF SEQ ID NOS: 268
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: trypsin-like serine proteases
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(610)
;   OTHER INFORMATION: n = a, t, c, or g
US-09-280-116-22

Alignment Scores:
Pred. No.:                3.63            Length:          610
Score:                    67.00           Matches:         26
Percent Similarity:       41.8%           Conservative:    12
Best Local Similarity:    28.6%           Mismatches:      35
Query Match:              15.0%           Indels:          18
DB:                       3               Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x US-09-280-116-22 (1-610)

Qy           4 AlaLeuAsnSerTyrPheGluPro----------ProValGluGluSerAlaLeuGlu 19
               ||||||          |||:::|||          ||||||   ||||||||||||
Db         341 GCCCTCATGCGGCTGTTCAAGCCCCTTGACCCTGTTCCGGTGAGGGAATCTGCACTCCCC 400

Qy          20 Arg----------ArgProGluThrIleSerGluProLysThrTyrValAspLeuThr 35
                           |||   |||:::     |||:::|||        |||
Db         401 GCTCTCCTGCCCCCCAGCCCCAGCACCCTCTGCAGCCCTCGCACTTGTCAGCATCTG--- 457

Qy          36 AsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGln 55
                           ::::::  |||      :::      |||
Db         458 ----------------------TCAACTCATATCCGGGCCCCAAAGCTTCTGCAGGGC 493

Qy          56 GluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsn 75
                        :::      :::      ::::::|||||||||          |||   |||
Db         494 AGAAGTCAAAGACTCTTAAAGATCCTTACATGGAACACTTCTGTTTTATAATTAGGGAAA 553

Qy          76 LeuSerGluArgAla---ArgGlyValCysSer 85
               |||      :::      :::   |||||||||
Db         554 CTGAAGCCCAAGGGTTATAAATAAGTTTGCTCC 586

RESULT 34
US-09-270-767-1642
; Sequence 1642, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1642
;   LENGTH: 938
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-1642

Alignment Scores:
Pred. No.:                6.59            Length:          938
Score:                    67.00           Matches:         21
Percent Similarity:       49.1%           Conservative:    5
Best Local Similarity:    39.6%           Mismatches:      25
Query Match:              15.0%           Indels:          3
DB:                       3               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-270-767-1642 (1-938)

Qy          12 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
               |||   ||||||||||      |||          |||:::   :::|||      |||
Db         463 CCAACGGAGGAAAGTACTACCGAAGGCTCCACTGAAAGCACAACCGAAGGATCCACA--- 519

Qy          32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
                  |||  |||                 ||||||||||||||   |||      :::
Db         520 ---GACAGCACCCAATCCACCGATCTGGACAGCACCACCAGTGATATCTGGAGCACCAGT 576

Qy          52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
               |||         |||   |||       :::|||||:::
Db         577 GATAAGGATGACGAATC-GGAGAGCAGTACTCCTTACTC 614

RESULT 35
US-09-270-767-16924
; Sequence 16924, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16924
;   LENGTH: 938
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-16924

Alignment Scores:
Pred. No.:                6.59            Length:          938
Score:                    67.00           Matches:         21
Percent Similarity:       49.1%           Conservative:    5
Best Local Similarity:    39.6%           Mismatches:      25
Query Match:              15.0%           Indels:          3
DB:                       3               Gaps:            1
```

```
US-10-757-745-2_COPY_54_140 (1-87) x US-09-270-767-16924 (1-938)

Qy          12 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
               |||   |||||||||       |||          |||:::   :::|||      |||
Db         463 CCAACGGAGGAAAGTACTACCGAAGGCTCCACTGAAAGCACAACCGAAGGATCCACA--- 519

Qy          32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
                  |||   |||                 ||||||||||||||   |||      :::
Db         520 ---GACAGCACCCAATCCACCGATCTGGACAGCACCACCAGTGATATCTGGAGCACCAGT 576

Qy          52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
               |||         |||  |||      :::||||||:::
Db         577 GATAAGGATGACGAATC-GGAGAGCAGTACTCCTTACTC 614

RESULT 36
US-09-248-796A-542
; Sequence 542, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 542
;   LENGTH: 1230
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-542

Alignment Scores:
Pred. No.:              9.61          Length:         1230
Score:                  67.00         Matches:        25
Percent Similarity:     47.8%         Conservative:   18
Best Local Similarity:  27.8%         Mismatches:     31
Query Match:            15.0%         Indels:         16
DB:                     3             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-248-796A-542 (1-1230)

Qy           3 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 22
               :::||||||   ::::::         |||   :::   |||   ::::::   |||
Db         745 AAAGCACTTGCCAATTTC---------CCAGATCAATATTCTACAATTGATTATCGTGAA 795

Qy          23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr---ThrAsp 41
               ::::::|||::::::   ||||||         ||||||:::   :::|||      |||
Db         796 AAATCAATTAATGAT---AAAACTGGCGCTCCCTTGACTGATTTCAAAACCATTGTTGAT 852

Qy          42 SerThrThrSerLysIleSerProSerGluAspThr------------GlnGlnGluAsn 57
                        :::::::::   |||   |||:::                  :::|||
Db         853 ATTTTGAAATTTAGAGTAGCTAAATCTGGTGATTCCATTGCATTTCAAAATGTTGACAAC 912

Qy          58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 77
               |||              :::||||||                                |||
Db         913 AATTCGAAAAGTAAACCATTAACTTGG--------------------AAGAAATTAGAG 951

Qy          78 GluArgAlaArgGlyValCysSerTyrLeu 87
               |||||||      |||||||   |||||||
Db         952 AATCGTGCTTATGCTGTATGTCAATATTTA 981

RESULT 37
US-08-216-592A-7
; Sequence 7, Application US/08216592A
; Patent No. 6635429
;  GENERAL INFORMATION:
;    APPLICANT:  LEID, MARK
;    APPLICANT:  KASTNER, PHILIPPE
;    APPLICANT:  CHAMBON, PIERRE
;    TITLE OF INVENTION:  NOVEL HETERODIMERIC STEROID RECEPTOR
;    TITLE OF INVENTION:  PROTEINS, GENES ENCODING SAME, AND USAGE THEREOF
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1100 New York Avenue NW Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/216,592A
;      FILING DATE:  23-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/825,667
;      FILING DATE:  24-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GOLDSTEIN, JORGE A
;      REGISTRATION NUMBER:  29,021
;      REFERENCE/DOCKET NUMBER:  1383.0060002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1488 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  36..1427
US-08-216-592A-7

Alignment Scores:
Pred. No.:              12.5          Length:         1488
Score:                  67.00         Matches:        23
Percent Similarity:     50.7%         Conservative:   11
Best Local Similarity:  34.3%         Mismatches:     23
Query Match:            15.0%         Indels:         10
DB:                     3             Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-08-216-592A-7 (1-1488)

Qy           2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
               |||   |||   :::|||:::   :::   |||||||||         ||||||
Db         708 GAATGTGCCAGTAGTAGCCACGAAGACATGCCCGTGGAGAGG---ATTCTAGAA------ 758

Qy          22 ProGluThrIleSerGluProLysThr------TyrValAspLeuThrAsnGluGluThr 39
                  |||          ||||||||||||      |||   |||:::      |||   :::
Db         759 GCCGAACTTGCTGTGGAACCAAAGACAGAATCCTACGGTGACATGAACGTGGAGAACTCA 818

Qy          40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 59
               |||:::          :::   |||      :::   |||   |||
Db         819 ACAAATGACCCTGTTACCAACATATGTCATGCTGCAGATAAGCAA--------------- 863

Qy          60 MetPheSerLeuIleThrTrp 66
               :::|||:::|||:::   |||
Db         864 CTTTTCACCCTCGTTGAGTGG 884

RESULT 38
US-08-820-170A-20
; Sequence 20, Application US/08820170A
; Patent No. 5831058
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Tsutomu, FUJIWARA
;    APPLICANT:  Takeshi, WATANABE
;    APPLICANT:  Masato, HORIE
;    APPLICANT:  Toyomasa, KATAGIRI
;    TITLE OF INVENTION:  HUMAN GENE
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak & Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/820,170A
;      FILING DATE:
;      CLASSIFICATION:  536
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1518 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA(genomic)
US-08-820-170A-20

Alignment Scores:
Pred. No.:                12.9          Length:          1518
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       2             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-08-820-170A-20 (1-1518)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::  :::||||||     ||||||                          |||
Db        451 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 510

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              |||                        ::::::||||||     |||:::|||
Db        511 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 546

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        547 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 585

RESULT 39
US-09-055-699-20
; Sequence 20, Application US/09055699
; Patent No. 6005088
;  GENERAL INFORMATION:
;    APPLICANT:  Tsutomu, FUJIWARA
;    APPLICANT:  Takeshi, WATANABE
;    APPLICANT:  Masato, HORIE
;    APPLICANT:  Toyomasa, KATAGIRI
;    TITLE OF INVENTION:  HUMAN GENE
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak & Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/055,699
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/820,170
;      FILING DATE:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1518 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA(genomic)
US-09-055-699-20

Alignment Scores:
Pred. No.:                12.9          Length:          1518
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-055-699-20 (1-1518)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::  :::||||||     ||||||                          |||
Db        451 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 510

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              |||                        ::::::||||||     |||:::|||
Db        511 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 546

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        547 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 585

RESULT 40
US-09-273-565-20
; Sequence 20, Application US/09273565A
; Patent No. 6166190
; GENERAL INFORMATION:
;  APPLICANT: FUJIWARA, TSUTOMU
;  APPLICANT: WATANABE, TAKESHI
;  APPLICANT: HORIE, MASATO
;  TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;  TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;  FILE REFERENCE: Q-53599
;  CURRENT APPLICATION NUMBER: US/09/273,565A
;  CURRENT FILING DATE: 1999-03-22
;  EARLIER APPLICATION NUMBER: 09/055,699
;  EARLIER FILING DATE: 1998-04-07
;  EARLIER APPLICATION NUMBER: 08/820,170
;  EARLIER FILING DATE: 1997-03-19
;  EARLIER APPLICATION NUMBER: JP 63410/1996
;  EARLIER FILING DATE: 1996-03-19
;  EARLIER APPLICATION NUMBER: JP 69163/1997
;  EARLIER FILING DATE: 1997-03-05
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
```

```
;   LENGTH: 1518
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-273-565-20

Alignment Scores:
Pred. No.:                12.9          Length:          1518
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-273-565-20 (1-1518)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                        |||
Db        451 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 510

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              |||                          ::::::||||||      |||:::|||
Db        511 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 546

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        547 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 585

RESULT 41
US-09-565-538-20
; Sequence 20, Application US/09565538
; Patent No. 6333404
; GENERAL INFORMATION:
;  APPLICANT: FUJIWARA, TSUTOMU
;  APPLICANT:  WATANABE, TAKESHI
;  APPLICANT:  HORIE, MASATO
;  TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;  TITLE OF INVENTION:  SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;  FILE REFERENCE: Q-53599
;  CURRENT APPLICATION NUMBER: US/09/565,538
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER: 09/273,565
;  PRIOR FILING DATE: 1999-03-22
;  PRIOR APPLICATION NUMBER: 09/055,699
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: 08/820,170
;  PRIOR FILING DATE: 1997-03-19
;  PRIOR APPLICATION NUMBER: JP 63410/1996
;  PRIOR FILING DATE: 1996-03-19
;  PRIOR APPLICATION NUMBER: JP 69163/1997
;  PRIOR FILING DATE: 1997-03-05
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 1518
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-565-538-20

Alignment Scores:
Pred. No.:                12.9          Length:          1518
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-565-538-20 (1-1518)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                        |||
Db        451 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 510

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              |||                          ::::::||||||      |||:::|||
Db        511 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 546

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        547 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 585

RESULT 42
US-09-661-468-20
; Sequence 20, Application US/09661468
; Patent No. 6376189
; GENERAL INFORMATION:
;  APPLICANT:  FUJIWARA, TSUTOMU
;  APPLICANT:  WATANABE, TAKESHI
;  APPLICANT:  HORIE, MASATO
;  TITLE OF INVENTION:  AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;  TITLE OF INVENTION:  SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;  FILE REFERENCE:  Q-53599
;  CURRENT APPLICATION NUMBER: US/09/661,468
;  CURRENT FILING DATE:  2000-09-13
;  PRIOR APPLICATION NUMBER:  09/055,699
;  PRIOR FILING DATE:  1998-04-07
;  PRIOR APPLICATION NUMBER:  08/820,170
;  PRIOR FILING DATE:  1997-03-19
;  PRIOR APPLICATION NUMBER:  JP 63410/1996
;  PRIOR FILING DATE:  1996-03-19
;  PRIOR APPLICATION NUMBER:  JP 69163/1997
;  PRIOR FILING DATE:  1997-03-05
;  NUMBER OF SEQ ID NOS:  95
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 1518
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-661-468-20

Alignment Scores:
Pred. No.:                12.9          Length:          1518
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-661-468-20 (1-1518)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                        |||
Db        451 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 510

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              |||                          ::::::||||||      |||:::|||
Db        511 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 546

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        547 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 585

RESULT 43
US-09-976-165-20
; Sequence 20, Application US/09976165
; Patent No. 6562947
; GENERAL INFORMATION:
;  APPLICANT: FUJIWARA, TSUTOMU
;  APPLICANT:  WATANABE, TAKESHI
;  APPLICANT:  HORIE, MASATO
;  TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;  TITLE OF INVENTION:  SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;  FILE REFERENCE: Q-53599
;  CURRENT APPLICATION NUMBER: US/09/976,165
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: 09/565,538
;  PRIOR FILING DATE: 2000-05-05
```

```
;  PRIOR APPLICATION NUMBER: 09/055,699
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: 08/820,170
;  PRIOR FILING DATE: 1997-03-19
;  PRIOR APPLICATION NUMBER: JP 63410/1996
;  PRIOR FILING DATE: 1996-03-19
;  PRIOR APPLICATION NUMBER: JP 69163/1997
;  PRIOR FILING DATE: 1997-03-05
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 1518
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-976-165-20

Alignment Scores:
Pred. No.:                12.9          Length:          1518
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-976-165-20 (1-1518)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::  :::||||||     ||||||                 |||
Db        451 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 510

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                |||                  ::::::||||||    |||:::|||
Db        511 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 546

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        547 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 585

RESULT 44
US-08-336-408B-5
; Sequence 5, Application US/08336408B
; Patent No. 5723329
;  GENERAL INFORMATION:
;    APPLICANT:  EVANS, RONALD M.
;    APPLICANT:  MANGELSDORF, DAVID J.
;    TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
;      STREET:  444 SOUTH FLOWER STREET, SUITE 2000
;      CITY:  LOS ANGELES
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/336,408B
;      FILING DATE:  08-NOV-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/933,453
;      FILING DATE:  21-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US91/00399
;      FILING DATE:  22-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/478,071
;      FILING DATE:  09-FEB-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  REITER, STEPHEN E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P41 9851
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-1995
;      TELEFAX:  619-546-9392
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1662 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  123..1514
US-08-336-408B-5

Alignment Scores:
Pred. No.:                14.6          Length:          1662
Score:                    67.00         Matches:         23
Percent Similarity:       50.7%         Conservative:    11
Best Local Similarity:    34.3%         Mismatches:      23
Query Match:              15.0%         Indels:          10
DB:                       2             Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x US-08-336-408B-5 (1-1662)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   |||   :::|||:::   :::   |||||||||         ||||||
Db        795 GAATGTGCCAGTAGTAGCCACGAAGACATGCCCGTGGAGAGG---ATTCTAGAA------ 845

Qy         22 ProGluThrIleSerGluProLysThr------TyrValAspLeuThrAsnGluGluThr 39
                |||          ||||||||||||      |||   |||:::       |||   :::
Db        846 GCCGAACTTGCTGTGGAACCAAAGACAGAATCCTACGGTGACATGAACGTGGAGAACTCA 905

Qy         40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 59
              |||:::         :::   |||       :::   |||   |||
Db        906 ACAAATGACCCTGTTACCAACATATGCCATGCTGCAGATAAGCAA--------------- 950

Qy         60 MetPheSerLeuIleThrTrp 66
              :::|||:::|||:::   |||
Db        951 CTTTTCACCCTCGTTGAGTGG 971

RESULT 45
PCT-US91-00399-5
; Sequence 5, Application PC/TUS9100399
;  GENERAL INFORMATION:
;    APPLICANT:  Mangelsdorf, Dr., David J.
;    APPLICANT:  Evans Dr., Ronald M.
;    TITLE OF INVENTION:  RETINOID RECEPTOR COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FITCH, EVEN, TABIN & FLANNERY
;      STREET:  135 South LaSalle Street, Suite 900
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/00399
;      FILING DATE:  19910122
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/478,071
;      FILING DATE:  09-FEB-1990
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Mr., Stephen E.
;      REGISTRATION NUMBER:  31192
;      REFERENCE/DOCKET NUMBER:  50852
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 552-1311
;      TELEFAX:  (619) 552-0095
;      TELEX:  20 6566 PATLAW CGO
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1662 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    IMMEDIATE SOURCE:
;      CLONE:  RXR MOUSE GAMMA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(123..1514)
PCT-US91-00399-5

Alignment Scores:
Pred. No.:              14.6          Length:        1662
Score:                  67.00         Matches:       23
Percent Similarity:     50.7%         Conservative:  11
Best Local Similarity:  34.3%         Mismatches:    23
Query Match:            15.0%         Indels:        10
DB:                     6             Gaps:          4

US-10-757-745-2_COPY_54_140 (1-87) x PCT-US91-00399-5 (1-1662)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   |||   :::|||:::   :::   |||||||||          ||||||
Db        795 GAATGTGCCAGTAGTAGCCACGAAGACATGCCCGTGGAGAGG---ATTCTAGAA------ 845

Qy         22 ProGluThrIleSerGluProLysThr------TyrValAspLeuThrAsnGluGluThr 39
                 |||          ||||||||||||      |||   |||:::     |||   :::
Db        846 GCCGAACTTGCTGTGGAACCAAAGACAGAATCCTACGGTGACATGAACGTGGAGAACTCA 905

Qy         40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 59
              |||:::         :::   |||       :::   |||   |||
Db        906 ACAAATGACCCTGTTACCAACATATGCCATGCTGCAGATAAGCAA--------------- 950

Qy         60 MetPheSerLeuIleThrTrp 66
              :::|||:::|||:::   |||
Db        951 CTTTTCACCCTCGTTGAGTGG 971

RESULT 46
US-08-820-170A-21
; Sequence 21, Application US/08820170A
; Patent No. 5831058
;  GENERAL INFORMATION:
;    APPLICANT:  Tsutomu, FUJIWARA
;    APPLICANT:  Takeshi, WATANABE
;    APPLICANT:  Masato, HORIE
;    APPLICANT:  Toyomasa, KATAGIRI
;    TITLE OF INVENTION:  HUMAN GENE
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak & Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/820,170A
;      FILING DATE:
;      CLASSIFICATION:  536
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2636 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA(genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    IMMEDIATE SOURCE:
;      LIBRARY:  Human fetal brain cDNA library
;      CLONE:  GEN-078D05
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  266..1783
US-08-820-170A-21

Alignment Scores:
Pred. No.:              27.8          Length:        2636
Score:                  67.00         Matches:       16
Percent Similarity:     49.1%         Conservative:  10
Best Local Similarity:  30.2%         Mismatches:    19
Query Match:            15.0%         Indels:        8
DB:                     2             Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-08-820-170A-21 (1-2636)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                        |||
Db        716 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 775

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                |||                          ::::::||||||      |||:::|||
Db        776 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 811

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        812 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 850

RESULT 47
US-09-055-699-21
; Sequence 21, Application US/09055699
; Patent No. 6005088
;  GENERAL INFORMATION:
;    APPLICANT:  Tsutomu, FUJIWARA
;    APPLICANT:  Takeshi, WATANABE
;    APPLICANT:  Masato, HORIE
;    APPLICANT:  Toyomasa, KATAGIRI
;    TITLE OF INVENTION:  HUMAN GENE
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak & Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/055,699
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  08/820,170
;       FILING DATE:
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 293-7060
;       TELEFAX:  (202) 293-7860
;       TELEX:  6491103
;   INFORMATION FOR SEQ ID NO:  21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2636 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA(genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     IMMEDIATE SOURCE:
;       LIBRARY:  Human fetal brain cDNA library
;       CLONE:  GEN-078D05
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  266..1783
US-09-055-699-21

Alignment Scores:
Pred. No.:                27.8          Length:          2636
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-055-699-21 (1-2636)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                        |||
Db        716 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 775

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
               |||                        ::::::||||||      |||:::|||
Db        776 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 811

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        812 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 850

RESULT 48
US-09-273-565-21
; Sequence 21, Application US/09273565A
; Patent No. 6166190
; GENERAL INFORMATION:
;  APPLICANT: FUJIWARA, TSUTOMU
;  APPLICANT: WATANABE, TAKESHI
;  APPLICANT: HORIE, MASATO
;  TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;  TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;  FILE REFERENCE: Q-53599
;  CURRENT APPLICATION NUMBER: US/09/273,565A
;  CURRENT FILING DATE: 1999-03-22
;  EARLIER APPLICATION NUMBER: 09/055,699
;  EARLIER FILING DATE: 1998-04-07
;  EARLIER APPLICATION NUMBER: 08/820,170
;  EARLIER FILING DATE: 1997-03-19
;  EARLIER APPLICATION NUMBER: JP 63410/1996
;  EARLIER FILING DATE: 1996-03-19
;  EARLIER APPLICATION NUMBER: JP 69163/1997
;  EARLIER FILING DATE: 1997-03-05
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 2636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (266)..(1783)
US-09-273-565-21

Alignment Scores:
Pred. No.:                27.8          Length:          2636
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-273-565-21 (1-2636)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                        |||
Db        716 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 775

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
               |||                        ::::::||||||      |||:::|||
Db        776 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 811

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        812 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 850

RESULT 49
US-09-565-538-21
; Sequence 21, Application US/09565538
; Patent No. 6333404
; GENERAL INFORMATION:
;  APPLICANT: FUJIWARA, TSUTOMU
;  APPLICANT:  WATANABE, TAKESHI
;  APPLICANT:  HORIE, MASATO
;  TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;  TITLE OF INVENTION:  SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;  FILE REFERENCE: Q-53599
;  CURRENT APPLICATION NUMBER: US/09/565,538
;  CURRENT FILING DATE:  2000-05-05
;  PRIOR APPLICATION NUMBER: 09/273,565
;  PRIOR FILING DATE: 1999-03-22
;  PRIOR APPLICATION NUMBER: 09/055,699
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: 08/820,170
;  PRIOR FILING DATE: 1997-03-19
;  PRIOR APPLICATION NUMBER: JP 63410/1996
;  PRIOR FILING DATE: 1996-03-19
;  PRIOR APPLICATION NUMBER: JP 69163/1997
;  PRIOR FILING DATE: 1997-03-05
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 2636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (266)..(1783)
US-09-565-538-21

Alignment Scores:
Pred. No.:                27.8          Length:          2636
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-565-538-21 (1-2636)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                        |||
Db        716 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 775
```

```
Qy          25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
               |||                          ::::::||||||      |||:::|||
Db         776 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 811

Qy          45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
               |||::::::   |||   |||      ::::::||||||
Db         812 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 850

RESULT 50
US-09-661-468-21
; Sequence 21, Application US/09661468
; Patent No. 6376189
; GENERAL INFORMATION:
;  APPLICANT:  FUJIWARA, TSUTOMU
;  APPLICANT:  WATANABE, TAKESHI
;  APPLICANT:  HORIE, MASATO
;  TITLE OF INVENTION:  AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;  TITLE OF INVENTION:  SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;  FILE REFERENCE:  Q-53599
;  CURRENT APPLICATION NUMBER: US/09/661,468
;  CURRENT FILING DATE:  2000-09-13
;  PRIOR APPLICATION NUMBER:  09/055,699
;  PRIOR FILING DATE:  1998-04-07
;  PRIOR APPLICATION NUMBER:  08/820,170
;  PRIOR FILING DATE:  1997-03-19
;  PRIOR APPLICATION NUMBER:  JP 63410/1996
;  PRIOR FILING DATE:  1996-03-19
;  PRIOR APPLICATION NUMBER:  JP 69163/1997
;  PRIOR FILING DATE:  1997-03-05
;  NUMBER OF SEQ ID NOS:  95
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 2636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (266)..(1783)
US-09-661-468-21

Alignment Scores:
Pred. No.:                27.8          Length:          2636
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-661-468-21 (1-2636)

Qy           5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
               :::|||:::   :::||||||      ||||||                        |||
Db         716 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 775

Qy          25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
               |||                          ::::::||||||      |||:::|||
Db         776 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 811

Qy          45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
               |||::::::   |||   |||      ::::::||||||
Db         812 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 850

RESULT 51
US-09-976-165-21
; Sequence 21, Application US/09976165
; Patent No. 6562947
; GENERAL INFORMATION:
;  APPLICANT: FUJIWARA, TSUTOMU
;  APPLICANT:  WATANABE, TAKESHI
;  APPLICANT:  HORIE, MASATO
;  TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;  TITLE OF INVENTION:  SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;  FILE REFERENCE: Q-53599
;  CURRENT APPLICATION NUMBER: US/09/976,165
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: 09/565,538
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: 09/055,699
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: 08/820,170
;  PRIOR FILING DATE: 1997-03-19
;  PRIOR APPLICATION NUMBER: JP 63410/1996
;  PRIOR FILING DATE: 1996-03-19
;  PRIOR APPLICATION NUMBER: JP 69163/1997
;  PRIOR FILING DATE: 1997-03-05
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 2636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (266)..(1783)
US-09-976-165-21

Alignment Scores:
Pred. No.:                27.8          Length:          2636
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-976-165-21 (1-2636)

Qy           5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
               :::|||:::   :::||||||      ||||||                        |||
Db         716 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 775

Qy          25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
               |||                          ::::::||||||      |||:::|||
Db         776 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 811

Qy          45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
               |||::::::   |||   |||      ::::::||||||
Db         812 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 850

RESULT 52
US-08-993-118-1/c
; Sequence 1, Application US/08993118
; Patent No. 5997872
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie;
;    APPLICANT:  DE SMET, Charles;
;    APPLICANT:  BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
;    TITLE OF INVENTION:  REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/993,118
;      FILING DATE:
```

```
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,528
;      FILING DATE:  April 25, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5455
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4031 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double-stranded
;      TOPOLOGY:  linear
US-08-993-118-1

Alignment Scores:
Pred. No.:                50.1          Length:          4031
Score:                    67.00         Matches:         29
Percent Similarity:       46.0%         Conservative:    17
Best Local Similarity:    29.0%         Mismatches:      31
Query Match:              15.0%         Indels:          23
DB:                       2             Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x US-08-993-118-1 (1-4031)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::             :::   ||||||   ||||||   :::   |||
Db       1883 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 1824

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
              ||||||   |||   |||             :::   ::::::|||||||||
Db       1823 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 1764

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                               ::::::   ::::::   :::|||   ||||||
Db       1763 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1704

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   ||||||  :::             ||||||
Db       1703 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1646

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                          |||||||||   ||||||      |||:::      :::
Db       1645 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1610

RESULT 53
US-08-845-528C-1/c
; Sequence 1, Application US/08845528C
; Patent No. 6027924
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie;
;    APPLICANT:  DE SMET, Charles;
;    APPLICANT:  BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
;    TITLE OF INVENTION:  REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,528C
;      FILING DATE:  April 25, 1997
;      CLASSIFICATION:  4335
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5455
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4031 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double-stranded
;      TOPOLOGY:  linear
US-08-845-528C-1

Alignment Scores:
Pred. No.:                50.1          Length:          4031
Score:                    67.00         Matches:         29
Percent Similarity:       46.0%         Conservative:    17
Best Local Similarity:    29.0%         Mismatches:      31
Query Match:              15.0%         Indels:          23
DB:                       3             Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x US-08-845-528C-1 (1-4031)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::             :::   ||||||   ||||||   :::   |||
Db       1883 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 1824

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
              ||||||   |||   |||             :::   ::::::|||||||||
Db       1823 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 1764

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                               ::::::   ::::::   :::|||   ||||||
Db       1763 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1704

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   ||||||  :::             ||||||
Db       1703 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1646

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                          |||||||||   ||||||      |||:::      :::
Db       1645 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1610

RESULT 54
US-09-066-281B-1/c
; Sequence 1, Application US/09066281B
; Patent No. 6475783
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie; DE SMET, Charles; BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING
;    TITLE OF INVENTION:  FOR TUMOR REJECTION ANTIGEN PRECURSOR MAGE-C1 AND MAGE-C2
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski L.L.P.
;      STREET:  666 Fifth Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/066,281B
```

```
;       FILING DATE:  April 24, 1998
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/845,528
;       FILING DATE:  April 25, 1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mary Anne Schofield
;       REGISTRATION NUMBER:  36,669
;       REFERENCE/DOCKET NUMBER:  LUD 5455.2 US - JEL/MAS
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 318-3100
;       TELEFAX:  (212) 752-5958
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4031 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double-stranded
;       TOPOLOGY:  linear
US-09-066-281B-1

Alignment Scores:
Pred. No.:                50.1          Length:          4031
Score:                    67.00         Matches:         29
Percent Similarity:       46.0%         Conservative:    17
Best Local Similarity:    29.0%         Mismatches:      31
Query Match:              15.0%         Indels:          23
DB:                       3             Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-066-281B-1 (1-4031)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::           :::   ||||||   ||||||   :::   |||
Db       1883 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 1824

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
                ||||||   |||   |||           :::   ::::::|||||||||
Db       1823 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 1764

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                              ::::::   ::::::   :::|||   ||||||
Db       1763 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1704

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   ||||||  :::             ||||||
Db       1703 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1646

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                          |||||||||   ||||||      |||:::      :::
Db       1645 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1610

RESULT 55
US-09-468-433C-1/c
; Sequence 1, Application US/09468433C
; Patent No. 6680191
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie; BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULES CODING FOR
;    TITLE OF INVENTION:  TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C AN
;    TITLE OF INVENTION:  MAGE-B FAMILIES AND USES THEREOF
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski L.L.P.
;      STREET:  801 Pennsylvania Avenue, NW
;      CITY:  Washington
;      STATE:  District of Columbia
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/468,433C
;      FILING DATE:  December 17, 1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/066,281
;      FILING DATE:  April 24, 1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/845,528
;      FILING DATE:  April 25, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5611 JEL/MAS
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 662-0200
;      TELEFAX:  (202) 662-4643
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4031 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double-stranded
;      TOPOLOGY:  linear
US-09-468-433C-1

Alignment Scores:
Pred. No.:                50.1          Length:          4031
Score:                    67.00         Matches:         29
Percent Similarity:       46.0%         Conservative:    17
Best Local Similarity:    29.0%         Mismatches:      31
Query Match:              15.0%         Indels:          23
DB:                       3             Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-468-433C-1 (1-4031)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::           :::   ||||||   ||||||   :::   |||
Db       1883 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 1824

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
                ||||||   |||   |||           :::   ::::::|||||||||
Db       1823 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 1764

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                              ::::::   ::::::   :::|||   ||||||
Db       1763 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1704

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   ||||||  :::             ||||||
Db       1703 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1646

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                          |||||||||   ||||||      |||:::      :::
Db       1645 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1610

RESULT 56
US-08-993-118-9/c
; Sequence 9, Application US/08993118
; Patent No. 5997872
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie;
;    APPLICANT:  DE SMET, Charles;
;    APPLICANT:  BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
;    TITLE OF INVENTION:  REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
```

```
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  PC-DOS
;       SOFTWARE:  Wordperfect
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/993,118
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/845,528
;       FILING DATE:  April 25, 1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mary Anne Schofield
;       REGISTRATION NUMBER:  36,669
;       REFERENCE/DOCKET NUMBER:  LUD 5455
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 688-9200
;       TELEFAX:  (212) 838-3884
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4225 base pairs
;       TYPE:  nucleic acids
;       STRANDEDNESS:  double-stranded
;       TOPOLOGY:  linear
US-08-993-118-9

Alignment Scores:
Pred. No.:                53.5          Length:          4225
Score:                    67.00         Matches:         29
Percent Similarity:       46.0%         Conservative:    17
Best Local Similarity:    29.0%         Mismatches:      31
Query Match:              15.0%         Indels:          23
DB:                       2             Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x US-08-993-118-9 (1-4225)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::             :::   ||||||   ||||||   :::    |||
Db       2093 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 2034

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
                ||||||   |||   |||            :::   ::::::|||||||||
Db       2033 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 1974

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                              ::::::   ::::::   :::|||   ||||||
Db       1973 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1914

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   ||||||  :::          ||||||
Db       1913 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1856

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                              |||||||||   ||||||      |||:::      :::
Db       1855 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1820

RESULT 57
US-08-845-528C-9/c
; Sequence 9, Application US/08845528C
; Patent No. 6027924
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie;
;    APPLICANT:  DE SMET, Charles;
;    APPLICANT:  BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
;    TITLE OF INVENTION:  REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,528C
;      FILING DATE:  April 25, 1997
;      CLASSIFICATION:  4335
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5455
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4225 base pairs
;      TYPE:  nucleic acids
;      STRANDEDNESS:  double-stranded
;      TOPOLOGY:  linear
US-08-845-528C-9

Alignment Scores:
Pred. No.:                53.5          Length:          4225
Score:                    67.00         Matches:         29
Percent Similarity:       46.0%         Conservative:    17
Best Local Similarity:    29.0%         Mismatches:      31
Query Match:              15.0%         Indels:          23
DB:                       3             Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x US-08-845-528C-9 (1-4225)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::             :::   ||||||   ||||||   :::    |||
Db       2093 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 2034

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
                ||||||   |||   |||            :::   ::::::|||||||||
Db       2033 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 1974

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                              ::::::   ::::::   :::|||   ||||||
Db       1973 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1914

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   ||||||  :::          ||||||
Db       1913 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1856

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                              |||||||||   ||||||      |||:::      :::
Db       1855 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1820

RESULT 58
US-09-066-281B-9/c
; Sequence 9, Application US/09066281B
; Patent No. 6475783
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie; DE SMET, Charles; BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING
;    TITLE OF INVENTION:  FOR TUMOR REJECTION ANTIGEN PRECURSOR MAGE-C1 AND MAGE-C2
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski L.L.P.
;      STREET:  666 Fifth Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
```

```
;      ZIP:  10103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/066,281B
;      FILING DATE:  April 24, 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/845,528
;      FILING DATE:  April 25, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5455.2 US - JEL/MAS
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 318-3100
;      TELEFAX:  (212) 752-5958
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4225 base pairs
;      TYPE:  nucleic acids
;      STRANDEDNESS:  double-stranded
;      TOPOLOGY:  linear
US-09-066-281B-9

Alignment Scores:
Pred. No.:              53.5          Length:         4225
Score:                  67.00         Matches:        29
Percent Similarity:     46.0%         Conservative:   17
Best Local Similarity:  29.0%         Mismatches:     31
Query Match:            15.0%         Indels:         23
DB:                     3             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-066-281B-9 (1-4225)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::           :::   ||||||   ||||||   :::   |||
Db       2093 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 2034

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
                ||||||   |||   |||           :::   ::::::|||||||||
Db       2033 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 1974

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                              ::::::   ::::::   :::|||   ||||||
Db       1973 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1914

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                    ||||||::::::   :::|||   ||||||  :::           ||||||
Db       1913 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1856

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                         |||||||||   ||||||      |||:::      :::
Db       1855 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1820

RESULT 59
US-09-468-433C-9/c
; Sequence 9, Application US/09468433C
; Patent No. 6680191
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie; BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULES CODING FOR
;    TITLE OF INVENTION:  TUMOR REJECTION ANTIGEN PRECURSORS OF MEMBERS OF THE MAGE-C AN
;    TITLE OF INVENTION:  MAGE-B FAMILIES AND USES THEREOF
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski L.L.P.
;      STREET:  801 Pennsylvania Avenue, NW
;      CITY:  Washington
;      STATE:  District of Columbia
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/468,433C
;      FILING DATE:  December 17, 1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/066,281
;      FILING DATE:  April 24, 1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/845,528
;      FILING DATE:  April 25, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5611 JEL/MAS
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 662-0200
;      TELEFAX:  (202) 662-4643
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4225 base pairs
;      TYPE:  nucleic acids
;      STRANDEDNESS:  double-stranded
;      TOPOLOGY:  linear
US-09-468-433C-9

Alignment Scores:
Pred. No.:              53.5          Length:         4225
Score:                  67.00         Matches:        29
Percent Similarity:     46.0%         Conservative:   17
Best Local Similarity:  29.0%         Mismatches:     31
Query Match:            15.0%         Indels:         23
DB:                     3             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-468-433C-9 (1-4225)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::           :::   ||||||   ||||||   :::   |||
Db       2093 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 2034

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
                ||||||   |||   |||           :::   ::::::|||||||||
Db       2033 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 1974

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                              ::::::   ::::::   :::|||   ||||||
Db       1973 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1914

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                    ||||||::::::   :::|||   ||||||  :::           ||||||
Db       1913 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1856

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                         |||||||||   ||||||      |||:::      :::
Db       1855 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1820

RESULT 60
US-09-061-709-1/c
; Sequence 1, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
;  APPLICANT: Chen, Yao-Tseng
;  APPLICANT: Gure, Ali
;  APPLICANT: Tsang, Solam
;  APPLICANT: Stockert, Elisabeth
;  APPLICANT: Jager, Elke
;  APPLICANT: Knuth, Alexander
```

```
;  APPLICANT: Old, Lloyd J.
;  TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
;  TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
;  FILE REFERENCE: LUD 5538
;  CURRENT APPLICATION NUMBER: US/09/061,709B
;  CURRENT FILING DATE: 1998-04-17
;  NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 1
;   LENGTH: 4265
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-061-709-1

Alignment Scores:
Pred. No.:              54.2          Length:         4265
Score:                  67.00         Matches:        29
Percent Similarity:     46.0%         Conservative:   17
Best Local Similarity:  29.0%         Mismatches:     31
Query Match:            15.0%         Indels:         23
DB:                     3             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-061-709-1 (1-4265)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::           :::   ||||||  ||||||   :::   |||
Db       2123 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 2064

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
              ||||||  |||  |||           :::   ::::::|||||||||
Db       2063 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 2004

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                              ::::::   ::::::   :::|||   ||||||
Db       2003 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1944

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   ||||||:::                ||||||
Db       1943 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1886

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                        |||||||||   ||||||      |||:::     :::
Db       1885 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1850

RESULT 61
US-09-899-651-1/c
; Sequence 1, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
;  APPLICANT: Chen, Yao-Tseng
;  APPLICANT:  Gure, Ali
;  APPLICANT:  Tsang, Solam
;  APPLICANT:  Stockert, Elisabeth
;  APPLICANT:  Jager, Elke
;  APPLICANT:  Knuth, Alexander
;  APPLICANT:  Old, Lloyd J.
;  TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
;  TITLE OF INVENTION:  Associated
;  TITLE OF INVENTION:  Antigen, The Antigens Per Se, And Uses Thereof
;  FILE REFERENCE: LUD 5538
;  CURRENT APPLICATION NUMBER: US/09/899,651
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: US/09/061,709
;  PRIOR FILING DATE: 1998-04-17
;  NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 1
;   LENGTH: 4265
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-899-651-1

Alignment Scores:
Pred. No.:              54.2          Length:         4265
Score:                  67.00         Matches:        29
Percent Similarity:     46.0%         Conservative:   17
Best Local Similarity:  29.0%         Mismatches:     31
Query Match:            15.0%         Indels:         23
DB:                     3             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-899-651-1 (1-4265)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::           :::   ||||||  ||||||   :::   |||
Db       2123 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 2064

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
              ||||||  |||  |||           :::   ::::::|||||||||
Db       2063 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 2004

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                              ::::::   ::::::   :::|||   ||||||
Db       2003 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1944

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   ||||||:::                ||||||
Db       1943 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1886

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                        |||||||||   ||||||      |||:::     :::
Db       1885 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1850

RESULT 62
US-09-392-714-15/c
; Sequence 15, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
;  APPLICANT: Scanlan, Matthew J.
;  APPLICANT: Gure, Ali O.
;  APPLICANT: Williamson, Barbara
;  APPLICANT: Chen, Yao-Tseng
;  APPLICANT: Old, Lloyd J.
;  TITLE OF INVENTION: Cancer Associated Antigens and Uses
;  TITLE OF INVENTION: Therefor
;  FILE REFERENCE: L0461/7062
;  CURRENT APPLICATION NUMBER: US/09/392,714A
;  CURRENT FILING DATE: 1999-09-09
;  EARLIER APPLICATION NUMBER: PCT/US98/14679
;  EARLIER FILING DATE: 1998-07-15
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 15
;   LENGTH: 4265
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-392-714-15

Alignment Scores:
Pred. No.:              54.2          Length:         4265
Score:                  67.00         Matches:        29
Percent Similarity:     46.0%         Conservative:   17
Best Local Similarity:  29.0%         Mismatches:     31
Query Match:            15.0%         Indels:         23
DB:                     3             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-392-714-15 (1-4265)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::           :::   ||||||  ||||||   :::   |||
Db       2123 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 2064

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
              ||||||  |||  |||           :::   ::::::|||||||||
Db       2063 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 2004

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
```

```
                                ::::::   ::::::   :::|||   ||||||
Db       2003 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1944

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   |||||| :::            ||||||
Db       1943 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1886

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                              |||||||||   ||||||      |||:::      :::
Db       1885 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1850

RESULT 63
US-09-270-437D-1/c
; Sequence 1, Application US/09270437D
; Patent No. 6830924
; GENERAL INFORMATION:
;  APPLICANT: Chen, Yao-Tseng
;  APPLICANT:  Gure, Ali
;  APPLICANT:  Tsang, Solam
;  APPLICANT:  Stockert, Elisabeth
;  APPLICANT:  Jager, Elke
;  APPLICANT:  Knuth, Alexander
;  APPLICANT:  Old, Lloyd J.
;  TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antige
;  TITLE OF INVENTION:  Antigens Per Se, And Uses Thereof
;  FILE REFERENCE: LUD 5538.1
;  CURRENT APPLICATION NUMBER: US/09/270,437D
;  CURRENT FILING DATE:  1999-03-16
;  PRIOR APPLICATION NUMBER: 09/061,709
;  PRIOR FILING DATE: 1998-04-17
;  NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 1
;   LENGTH: 4265
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-270-437D-1

Alignment Scores:
Pred. No.:              54.2          Length:         4265
Score:                  67.00         Matches:        29
Percent Similarity:     46.0%         Conservative:   17
Best Local Similarity:  29.0%         Mismatches:     31
Query Match:            15.0%         Indels:         23
DB:                     3             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-270-437D-1 (1-4265)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::           :::   ||||||  ||||||   :::   |||
Db       2123 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 2064

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
              ||||||  |||  |||           :::  ::::::|||||||||
Db       2063 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 2004

Qy         40 -----------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                               ::::::   ::::::   :::|||   ||||||
Db       2003 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1944

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   |||||| :::            ||||||
Db       1943 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1886

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                              |||||||||   ||||||      |||:::      :::
Db       1885 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1850

RESULT 64
US-09-949-016-13215
; Sequence 13215, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13215
;   LENGTH: 10160
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13215

Alignment Scores:
Pred. No.:              181           Length:         10160
Score:                  67.00         Matches:        16
Percent Similarity:     52.2%         Conservative:   8
Best Local Similarity:  34.8%         Mismatches:     18
Query Match:            15.0%         Indels:         4
DB:                     3             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-949-016-13215 (1-10160)

Qy         12 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
              |||   ||||||      |||   :::|||||:::|||   :::            |||
Db       6254 CCTGCTGAGGAGGTAATTCTGCCCCAAAGACCACAGACTCCCACCCACCACAGCCCATAC 6313

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
                                        |||      ::::::|||:::   ||||||:::
Db       6314 ------------ACTGCCCCTCACACTCCCATGTCCTCCTCCAGGTCCAGCCCAGCTGCT 6361

Qy         52 AspThrGlnGlnGluAsn 57
              |||   |||:::||||||
Db       6362 GACGTCCAGGAAGAAAAC 6379

RESULT 65
US-09-248-796A-9100
; Sequence 9100, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 9100
;   LENGTH: 891
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-9100

Alignment Scores:
Pred. No.:              7.24          Length:         891
Score:                  66.50         Matches:        21
Percent Similarity:     49.2%         Conservative:   9
Best Local Similarity:  34.4%         Mismatches:     18
Query Match:            14.8%         Indels:         13
DB:                     3             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-248-796A-9100 (1-891)
```

```
Qy          22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu--------- 38
               |||::::::   |||    |||  |||  |||      |||||||||
Db         451 CCCAAATCAACATCGTCAACAAAAAATTATCATGAC------AATGAGGAAGCCTTTTCT 504

Qy          39 ---ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
                  :::      ||||||||||||   :::      |||      ::::::|||   :::
Db         505 GACTCTGACTCTTCTACAACGTCTACACTACTAGGTAGTCCATTTTCGGAACAAGTAGAT 564

Qy          58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 77
                                           |||||||||   :::|||||||||   :::
Db         565 GATAGA--------------------AATATTGATGCCATAGATTTAAATGAAATTGAA 603

Qy          78 Glu 78
               |||
Db         604 GAA 606

RESULT 66
US-09-248-796A-6805
; Sequence 6805, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6805
;   LENGTH: 3003
;   TYPE: DNA
;   ORGANISM: Candida albicans
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (2931),(2974),(2981)
;   OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno
US-09-248-796A-6805

Alignment Scores:
Pred. No.:                39.2          Length:          3003
Score:                    66.50         Matches:         20
Percent Similarity:       43.8%         Conservative:    8
Best Local Similarity:    31.2%         Mismatches:      17
Query Match:              14.8%         Indels:          19
DB:                       3             Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-248-796A-6805 (1-3003)

Qy          31 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer--------------- 45
               :::      |||||||||      |||   |||:::   |||:::
Db         799 TGGGCAGGGTTAACTAATAATCCAACATTGGATACTGTTACTACCTGGCAACAACAGAAC 858

Qy          46 ---------LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
                        |||:::||||||      |||  |||      :::|||      |||
Db         859 ATGCCCGATAAGGTATCTCCATGCAATGACAAGCAGTTGGTCCACGGAAAACAATTTGGC 918

Qy          63 ---------------------LeuIleThrTrpAsnIleAsp-----------GlyLeu 71
                                   |||:::     :::   |||            ||||||
Db         919 AAAAAAGAGTTTGATAGCCATTTGCTTAGATTGCACTTTGATGGCCAATATACAGGCCTA 978

Qy          72 AspLeuAsnAsn 75
               |||||||||:::
Db         979 GATTTGAATTCA 990

RESULT 67
US-09-112-450-1
; Sequence 1, Application US/09112450
; Patent No. 6120999
; GENERAL INFORMATION:
;  APPLICANT: Abad, Antonio Jose C.
;  APPLICANT: Choi, Gil
;  APPLICANT: Calderone, Richard A.
;  TITLE OF INVENTION: Histidine Kinase Two-component in Candida albicans
;  FILE REFERENCE: PF393
;  CURRENT APPLICATION NUMBER: US/09/112,450
;  CURRENT FILING DATE: 1998-07-09
;  EARLIER APPLICATION NUMBER: 60/052,273
;  EARLIER FILING DATE: 1997-07-10
;  EARLIER APPLICATION NUMBER: 60/074,308
;  EARLIER FILING DATE: 1998-02-11
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3119
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-112-450-1

Alignment Scores:
Pred. No.:                41.4          Length:          3119
Score:                    66.50         Matches:         20
Percent Similarity:       43.8%         Conservative:    8
Best Local Similarity:    31.2%         Mismatches:      17
Query Match:              14.8%         Indels:          19
DB:                       3             Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-112-450-1 (1-3119)

Qy          31 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer--------------- 45
               :::      |||||||||      |||   |||:::   |||:::
Db         934 TGGGCAGGGTTAACTAATAATCCAACATTGGATACTGTTACTACCTGGCAACAACAGAAC 993

Qy          46 ---------LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
                        |||:::||||||      |||  |||      :::|||      |||
Db         994 ATGCCCGATAAGGTATCTCCATGCAATGACAAGCAGTTGGTCCACGGAAAACAATTTGGC 1053

Qy          63 ---------------------LeuIleThrTrpAsnIleAsp-----------GlyLeu 71
                                   |||:::     :::   |||            ||||||
Db        1054 AAAAAAGAGTTTGATAGCCATTTGCTCAGATTGCACTTTGATGGCCAATATACAGGCCTA 1113

Qy          72 AspLeuAsnAsn 75
               |||||||||:::
Db        1114 GATTTGAATTCA 1125

RESULT 68
US-09-419-291A-1
; Sequence 1, Application US/09419291A
; Patent No. 6416989
; GENERAL INFORMATION:
;  APPLICANT: Abad et al.
;  TITLE OF INVENTION: Histidine Kinase Two-component in Candida albicans
;  FILE REFERENCE: PB393D1
;  CURRENT APPLICATION NUMBER: US/09/419,291A
;  CURRENT FILING DATE:  1999-10-15
;  PRIOR APPLICATION NUMBER: US 09/112,450
;  PRIOR FILING DATE: 1998-07-09
;  PRIOR APPLICATION NUMBER: US 60/074,308
;  PRIOR FILING DATE: 1998-02-11
;  PRIOR APPLICATION NUMBER: US 60/052,273
;  PRIOR FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 3119
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-419-291A-1

Alignment Scores:
Pred. No.:                41.4          Length:          3119
Score:                    66.50         Matches:         20
```

```
Percent Similarity:     43.8%          Conservative:  8
Best Local Similarity:  31.2%          Mismatches:    17
Query Match:            14.8%          Indels:        19
DB:                     3              Gaps:          3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-419-291A-1 (1-3119)

Qy         31 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer---------------- 45
              :::     |||||||||       |||   |||:::   |||:::
Db        934 TGGGCAGGGTTAACTAATAATCCAACATTGGATACTGTTACTACCTGGCAACAACAGAAC 993

Qy         46 ---------LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
                      |||:::||||||       |||   |||       :::|||       |||
Db        994 ATGCCCGATAAGGTATCTCCATGCAATGACAAGCAGTTGGTCCACGGAAAACAATTTGGC 1053

Qy         63 --------------------LeuIleThrTrpAsnIleAsp------------GlyLeu 71
                                 |||:::      :::   |||             ||||||
Db       1054 AAAAAAGAGTTTGATAGCCATTTGCTCAGATTGCACTTTGATGGCCAATATACAGGCCTA 1113

Qy         72 AspLeuAsnAsn 75
              |||||||||:::
Db       1114 GATTTGAATTCA 1125

RESULT 69
US-10-116-048-1
; Sequence 1, Application US/10116048
; Patent No. 6824992
; GENERAL INFORMATION:
;  APPLICANT: Abad et al.
;  TITLE OF INVENTION: Histidine Kinase Two-component in Candida albicans
;  FILE REFERENCE: PB393D2
;  CURRENT APPLICATION NUMBER: US/10/116,048
;  CURRENT FILING DATE:  2002-04-05
;  PRIOR APPLICATION NUMBER: US 09/419,291
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: US 09/112,450
;  PRIOR FILING DATE: 1998-07-09
;  PRIOR APPLICATION NUMBER: US 60/074,308
;  PRIOR FILING DATE: 1998-02-11
;  PRIOR APPLICATION NUMBER: US 60/052,273
;  PRIOR FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 3119
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-10-116-048-1

Alignment Scores:
Pred. No.:              41.4           Length:        3119
Score:                  66.50          Matches:       20
Percent Similarity:     43.8%          Conservative:  8
Best Local Similarity:  31.2%          Mismatches:    17
Query Match:            14.8%          Indels:        19
DB:                     3              Gaps:          3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-116-048-1 (1-3119)

Qy         31 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer---------------- 45
              :::     |||||||||       |||   |||:::   |||:::
Db        934 TGGGCAGGGTTAACTAATAATCCAACATTGGATACTGTTACTACCTGGCAACAACAGAAC 993

Qy         46 ---------LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
                      |||:::||||||       |||   |||       :::|||       |||
Db        994 ATGCCCGATAAGGTATCTCCATGCAATGACAAGCAGTTGGTCCACGGAAAACAATTTGGC 1053

Qy         63 --------------------LeuIleThrTrpAsnIleAsp------------GlyLeu 71
                                 |||:::      :::   |||             ||||||
Db       1054 AAAAAAGAGTTTGATAGCCATTTGCTCAGATTGCACTTTGATGGCCAATATACAGGCCTA 1113

Qy         72 AspLeuAsnAsn 75
              |||||||||:::
Db       1114 GATTTGAATTCA 1125

RESULT 70
US-09-112-450-3
; Sequence 3, Application US/09112450
; Patent No. 6120999
; GENERAL INFORMATION:
;  APPLICANT: Abad, Antonio Jose C.
;  APPLICANT: Choi, Gil
;  APPLICANT: Calderone, Richard A.
;  TITLE OF INVENTION: Histidine Kinase Two-component in Candida albicans
;  FILE REFERENCE: PF393
;  CURRENT APPLICATION NUMBER: US/09/112,450
;  CURRENT FILING DATE: 1998-07-09
;  EARLIER APPLICATION NUMBER: 60/052,273
;  EARLIER FILING DATE: 1997-07-10
;  EARLIER APPLICATION NUMBER: 60/074,308
;  EARLIER FILING DATE: 1998-02-11
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 8561
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-112-450-3

Alignment Scores:
Pred. No.:              168            Length:        8561
Score:                  66.50          Matches:       20
Percent Similarity:     43.8%          Conservative:  8
Best Local Similarity:  31.2%          Mismatches:    17
Query Match:            14.8%          Indels:        19
DB:                     3              Gaps:          3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-112-450-3 (1-8561)

Qy         31 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer---------------- 45
              :::     |||||||||       |||   |||:::   |||:::
Db       6370 TGGGCAGGGTTAACTAATAATCCAACATTGGATACTGTTACTACCTGGCAACAACAGAAC 6429

Qy         46 ---------LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
                      |||:::||||||       |||   |||       :::|||       |||
Db       6430 ATGCCCGATAAGGTATCTCCATGCAATGACAAGCAGTTGGTCCACGGAAAACAATTTGGC 6489

Qy         63 --------------------LeuIleThrTrpAsnIleAsp------------GlyLeu 71
                                 |||:::      :::   |||             ||||||
Db       6490 AAAAAAGAGTTTGATAGCCATTTGCTCAGATTGCACTTTGATGGCCAATATACAGGCCTA 6549

Qy         72 AspLeuAsnAsn 75
              |||||||||:::
Db       6550 GATTTGAATTCA 6561

RESULT 71
US-09-419-291A-3
; Sequence 3, Application US/09419291A
; Patent No. 6416989
; GENERAL INFORMATION:
;  APPLICANT: Abad et al.
;  TITLE OF INVENTION: Histidine Kinase Two-component in Candida albicans
;  FILE REFERENCE: PB393D1
;  CURRENT APPLICATION NUMBER: US/09/419,291A
;  CURRENT FILING DATE:  1999-10-15
;  PRIOR APPLICATION NUMBER: US 09/112,450
;  PRIOR FILING DATE: 1998-07-09
;  PRIOR APPLICATION NUMBER: US 60/074,308
;  PRIOR FILING DATE: 1998-02-11
;  PRIOR APPLICATION NUMBER: US 60/052,273
;  PRIOR FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 8561
;   TYPE: DNA
```

```
;   ORGANISM: Candida albicans
US-09-419-291A-3

Alignment Scores:
Pred. No.:                168             Length:          8561
Score:                    66.50           Matches:         20
Percent Similarity:       43.8%           Conservative:    8
Best Local Similarity:    31.2%           Mismatches:      17
Query Match:              14.8%           Indels:          19
DB:                       3               Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-419-291A-3 (1-8561)

Qy         31 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer--------------- 45
              :::     |||||||||      |||   |||:::   |||:::
Db       6370 TGGGCAGGGTTAACTAATAATCCAACATTGGATACTGTTACTACCTGGCAACAACAGAAC 6429

Qy         46 ---------LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
                      |||:::||||||      |||   |||      :::|||      |||
Db       6430 ATGCCCGATAAGGTATCTCCATGCAATGACAAGCAGTTGGTCCACGGAAAACAATTTGGC 6489

Qy         63 ---------------------LeuIleThrTrpAsnIleAsp------------GlyLeu 71
                                  |||:::      :::   |||            ||||||
Db       6490 AAAAAAGAGTTTGATAGCCATTTGCTCAGATTGCACTTTGATGGCCAATATACAGGCCTA 6549

Qy         72 AspLeuAsnAsn 75
              |||||||||:::
Db       6550 GATTTGAATTCA 6561

RESULT 72
US-10-116-048-3
; Sequence 3, Application US/10116048
; Patent No. 6824992
; GENERAL INFORMATION:
;  APPLICANT: Abad et al.
;  TITLE OF INVENTION: Histidine Kinase Two-component in Candida albicans
;  FILE REFERENCE: PB393D2
;  CURRENT APPLICATION NUMBER: US/10/116,048
;  CURRENT FILING DATE:  2002-04-05
;  PRIOR APPLICATION NUMBER: US 09/419,291
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: US 09/112,450
;  PRIOR FILING DATE: 1998-07-09
;  PRIOR APPLICATION NUMBER: US 60/074,308
;  PRIOR FILING DATE: 1998-02-11
;  PRIOR APPLICATION NUMBER: US 60/052,273
;  PRIOR FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 8561
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-10-116-048-3

Alignment Scores:
Pred. No.:                168             Length:          8561
Score:                    66.50           Matches:         20
Percent Similarity:       43.8%           Conservative:    8
Best Local Similarity:    31.2%           Mismatches:      17
Query Match:              14.8%           Indels:          19
DB:                       3               Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-116-048-3 (1-8561)

Qy         31 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer--------------- 45
              :::     |||||||||      |||   |||:::   |||:::
Db       6370 TGGGCAGGGTTAACTAATAATCCAACATTGGATACTGTTACTACCTGGCAACAACAGAAC 6429

Qy         46 ---------LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
                      |||:::||||||      |||   |||      :::|||      |||
Db       6430 ATGCCCGATAAGGTATCTCCATGCAATGACAAGCAGTTGGTCCACGGAAAACAATTTGGC 6489

Qy         63 ---------------------LeuIleThrTrpAsnIleAsp------------GlyLeu 71
                                  |||:::      :::   |||            ||||||
Db       6490 AAAAAAGAGTTTGATAGCCATTTGCTCAGATTGCACTTTGATGGCCAATATACAGGCCTA 6549

Qy         72 AspLeuAsnAsn 75
              |||||||||:::
Db       6550 GATTTGAATTCA 6561

RESULT 73
US-09-199-637A-153/c
; Sequence 153, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
;  APPLICANT:  Ausubel, Frederick
;  APPLICANT:  Goodman, Howard M.
;  APPLICANT:  Rahme, Laurence G.
;  APPLICANT:  Mahajan-Miklos, Shalina
;  APPLICANT:  Tan, Man-Wah
;  APPLICANT:  Cao, Hui
;  APPLICANT:  Drenkard, Eliana
;  APPLICANT:  Tsongalis, John
;  TITLE OF INVENTION:  VIRULENCE-ASSOCIATED NUCLEIC ACID
;  TITLE OF INVENTION:  SEQUENCES AND USES THEREOF
;  FILE REFERENCE:  00786/361002
;  CURRENT APPLICATION NUMBER: US/09/199,637A
;  CURRENT FILING DATE:  1998-11-25
;  PRIOR APPLICATION NUMBER:  60/066,517
;  PRIOR FILING DATE:  1997-11-25
;  NUMBER OF SEQ ID NOS:  437
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 153
;   LENGTH: 762
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-153

Alignment Scores:
Pred. No.:                6.87            Length:          762
Score:                    66.00           Matches:         15
Percent Similarity:       45.5%           Conservative:    5
Best Local Similarity:    34.1%           Mismatches:      20
Query Match:              14.7%           Indels:          4
DB:                       3               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-199-637A-153 (1-762)

Qy         11 ProProValGluGluSer------------AlaLeuGluArgArgProGluThrIleSer 26
              ||||||::::::   |||                ||||||||||
Db       649 CCACCAATAGATAGAAGCCGGCGCCCGAAGACTGCCAGGCGTCGCCCCAGCGCTCCAGGC 590

Qy         27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 46
              ||||||                |||      :::          ||||||||   |||
Db       589 GAGCCAGGCGCGATTGCCAGGCTGCCGACGCAAGCGGCGGACGATAGCACCCGAAGCGCC 530

Qy         47 IleSerProSer 50
              :::||||||:::
Db       529 GTTTCTCCAGCG 518

RESULT 74
US-09-199-637A-154/c
; Sequence 154, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
;  APPLICANT:  Ausubel, Frederick
;  APPLICANT:  Goodman, Howard M.
;  APPLICANT:  Rahme, Laurence G.
;  APPLICANT:  Mahajan-Miklos, Shalina
;  APPLICANT:  Tan, Man-Wah
;  APPLICANT:  Cao, Hui
;  APPLICANT:  Drenkard, Eliana
;  APPLICANT:  Tsongalis, John
;  TITLE OF INVENTION:  VIRULENCE-ASSOCIATED NUCLEIC ACID
;  TITLE OF INVENTION:  SEQUENCES AND USES THEREOF
```

```
;  FILE REFERENCE:  00786/361002
;  CURRENT APPLICATION NUMBER: US/09/199,637A
;  CURRENT FILING DATE:  1998-11-25
;  PRIOR APPLICATION NUMBER:  60/066,517
;  PRIOR FILING DATE:  1997-11-25
;  NUMBER OF SEQ ID NOS:  437
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 154
;   LENGTH: 801
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-154

Alignment Scores:
Pred. No.:              7.36          Length:          801
Score:                  66.00         Matches:         15
Percent Similarity:     45.5%         Conservative:    5
Best Local Similarity:  34.1%         Mismatches:      20
Query Match:            14.7%         Indels:          4
DB:                     3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-199-637A-154 (1-801)

Qy         11 ProProValGluGluSer------------AlaLeuGluArgArgProGluThrIleSer 26
              ||||||::::::   |||                ||||||||
Db        239 CCACCAATAGATAGAAGCCGGCGCCCGAAGACTGCCAGGCGTCGCCCCAGCGCTCCAGGC 180

Qy         27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 46
              ||||||              |||      :::       ||||||||   |||
Db        179 GAGCCAGGCGCGATTGCCAGGCTGCCGACGCAAGCGGCGGACGATAGCACCCGAAGCGCC 120

Qy         47 IleSerProSer 50
              :::||||||:::
Db        119 GTTTCTCCAGCG 108

RESULT 75
US-09-252-991A-13439/c
; Sequence 13439, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 13439
;   LENGTH: 891
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13439

Alignment Scores:
Pred. No.:              8.54          Length:          891
Score:                  66.00         Matches:         15
Percent Similarity:     45.5%         Conservative:    5
Best Local Similarity:  34.1%         Mismatches:      20
Query Match:            14.7%         Indels:          4
DB:                     3             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-252-991A-13439 (1-891)

Qy         11 ProProValGluGluSer------------AlaLeuGluArgArgProGluThrIleSer 26
              ||||||::::::   |||                ||||||||
Db        778 CCACCAATAGATAGAAGCCGGCGCCCGAAGACTGCCAGGCGTCGCCCCAGCGCTCCAGGC 719

Qy         27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 46
              ||||||              |||      :::       ||||||||   |||
Db        718 GAGCCAGGCGCGATTGCCAGGCTGCCGACGCAAGCGGCGGACGATAGCACCCGAAGCGCC 659

Qy         47 IleSerProSer 50
              :::||||||:::
Db        658 GTTTCTCCAGCG 647

Search completed: March 10, 2006, 22:26:26
Job time : 110.939 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 03:44:14 ; Search time 251.78 Seconds
(without alignments)
797.263 Million cell updates/sec

Title: US-10-757-745-2_COPY_54_140
Perfect score: 448
Sequence: 1 MERALNSYFEPPVEESALER.....IDGLDLNNLSERARGVCSYL 87

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7673375 seqs, 1153648444 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170339_4330/app_query.fasta_1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1676_@runat_09032006_170339_4330 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	288	64.3	606	6	US-09-925-065A-783537	Sequence 783537,
	2	288	64.3	634	6	US-09-925-065A-783536	Sequence 783536,
	3	78	17.4	3672	8	US-10-821-234-301	Sequence 301, App
c	4	78	17.4	4140	12	US-11-091-883-189	Sequence 189, App
	5	72	16.1	1806	7	US-10-932-182A-487	Sequence 487, App
	6	72	16.1	1806	7	US-10-932-182A-487	Sequence 487, App
	7	70.5	15.7	184868	12	US-11-121-086-88	Sequence 88, Appl
c	8	70	15.6	496	6	US-09-925-065A-747874	Sequence 747874,
c	9	67.5	15.1	1194	9	US-11-096-568A-7380	Sequence 7380, Ap
	10	67.5	15.1	2361	8	US-10-485-517-118	Sequence 118, App
	11	67	15.0	1410	9	US-11-096-568A-12999	Sequence 12999, A
c	12	67	15.0	4265	12	US-11-044-051-74	Sequence 74, Appl
	13	66.5	14.8	600	12	US-11-136-527-6443	Sequence 6443, Ap
c	14	66.5	14.8	664	6	US-09-925-065A-610817	Sequence 610817,
c	15	66.5	14.8	810	7	US-10-932-182A-3877	Sequence 3877, Ap
c	16	66.5	14.8	810	7	US-10-932-182A-3877	Sequence 3877, Ap
	17	66.5	14.8	2124	7	US-10-932-182A-77405	Sequence 77405, A
	18	66.5	14.8	2124	7	US-10-932-182A-77405	Sequence 77405, A
	19	66.5	14.8	2270	12	US-11-136-527-2347	Sequence 2347, Ap
c	20	66	14.7	499	8	US-10-521-162-33	Sequence 33, Appl
	21	66	14.7	499	8	US-10-521-162-35	Sequence 35, Appl
	22	66	14.7	870	12	US-11-037-243-11	Sequence 11, Appl
c	23	66	14.7	1443	8	US-10-521-162-39	Sequence 39, Appl
	24	66	14.7	1443	8	US-10-521-162-41	Sequence 41, Appl
	25	66	14.7	2121	7	US-10-932-182A-77604	Sequence 77604, A
	26	66	14.7	2121	7	US-10-932-182A-77604	Sequence 77604, A
	27	66	14.7	2352	12	US-11-037-243-8	Sequence 8, Appli
	28	65	14.5	390	8	US-10-873-528-198	Sequence 198, App
	29	65	14.5	585	6	US-09-925-065A-887050	Sequence 887050,
	30	65	14.5	585	6	US-09-925-065A-887051	Sequence 887051,
	31	65	14.5	1077	7	US-10-932-182A-4655	Sequence 4655, Ap
	32	65	14.5	1077	7	US-10-932-182A-4655	Sequence 4655, Ap
	33	65	14.5	3741	12	US-11-136-527-751	Sequence 751, App
	34	64.5	14.4	1914	8	US-10-750-185-42446	Sequence 42446, A
	35	64.5	14.4	1914	8	US-10-750-623-42446	Sequence 42446, A
	36	64.5	14.4	3930	7	US-10-932-182A-76026	Sequence 76026, A
	37	64.5	14.4	3930	7	US-10-932-182A-76026	Sequence 76026, A
c	38	64	14.3	312	6	US-09-925-065A-203999	Sequence 203999,
	39	64	14.3	861	7	US-10-932-182A-5731	Sequence 5731, Ap
	40	64	14.3	861	7	US-10-932-182A-5731	Sequence 5731, Ap
c	41	64	14.3	1098	8	US-10-454-437-141	Sequence 141, App
	42	64	14.3	1623	9	US-11-096-568A-19504	Sequence 19504, A
c	43	64	14.3	1777	6	US-09-925-065A-711158	Sequence 711158,
	44	64	14.3	2430	7	US-10-932-182A-1233	Sequence 1233, Ap
	45	64	14.3	2430	7	US-10-932-182A-1233	Sequence 1233, Ap
	46	64	14.3	2457	7	US-10-932-182A-79549	Sequence 79549, A
	47	64	14.3	2457	7	US-10-932-182A-79549	Sequence 79549, A
c	48	64	14.3	2995	12	US-11-136-527-3388	Sequence 3388, Ap
c	49	63.5	14.2	648	6	US-09-925-065A-750498	Sequence 750498,
	50	63.5	14.2	849	7	US-10-932-182A-1169	Sequence 1169, Ap
	51	63.5	14.2	849	7	US-10-932-182A-1169	Sequence 1169, Ap
	52	63.5	14.2	1673	8	US-10-750-185-41248	Sequence 41248, A
	53	63.5	14.2	1673	8	US-10-750-623-41248	Sequence 41248, A
	54	63.5	14.2	2895	7	US-10-932-182A-1899	Sequence 1899, Ap
	55	63.5	14.2	2895	7	US-10-932-182A-1899	Sequence 1899, Ap
	56	63	14.1	921	12	US-11-052-554A-605	Sequence 605, App
	57	63	14.1	1400	12	US-11-136-527-4231	Sequence 4231, Ap
	58	63	14.1	1604	12	US-11-136-527-135	Sequence 135, App
	59	63	14.1	5256	12	US-11-103-957-46	Sequence 46, Appl
	60	62.5	14.0	594	7	US-10-932-182A-81472	Sequence 81472, A
	61	62.5	14.0	594	7	US-10-932-182A-81472	Sequence 81472, A
	62	62.5	14.0	624	6	US-09-925-065A-180175	Sequence 180175,
	63	62.5	14.0	624	6	US-09-925-065A-180178	Sequence 180178,
	64	62.5	14.0	2526	7	US-10-932-182A-3132	Sequence 3132, Ap
	65	62.5	14.0	2526	7	US-10-932-182A-3132	Sequence 3132, Ap
c	66	62.5	14.0	45764	7	US-10-330-773-251	Sequence 251, App
c	67	62.5	14.0	201990	8	US-10-995-561-13303	Sequence 13303, A
c	68	62.5	14.0	215248	7	US-10-330-773-261	Sequence 261, App
	69	62	13.8	201	8	US-10-995-561-43975	Sequence 43975, A
	70	62	13.8	201	8	US-10-995-561-43976	Sequence 43976, A
	71	62	13.8	201	8	US-10-995-561-43977	Sequence 43977, A
	72	62	13.8	201	8	US-10-995-561-43994	Sequence 43994, A
	73	62	13.8	201	12	US-11-124-367A-29340	Sequence 29340, A
	74	62	13.8	201	12	US-11-124-367A-29341	Sequence 29341, A
	75	62	13.8	201	12	US-11-124-367A-29342	Sequence 29342, A

ALIGNMENTS

```
RESULT 1
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783537

Alignment Scores:
Pred. No.:                4.75e-28        Length:          606
Score:                    288.00          Matches:         56
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              64.3%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-783537 (1-606)

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 473

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 533

Qy         72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 581

RESULT 2
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783536

Alignment Scores:
Pred. No.:                5.01e-28        Length:          634
Score:                    288.00          Matches:         56
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              64.3%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-783536 (1-634)

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 GTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAA 324

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTA 384

Qy         72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        385 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 432

RESULT 3
US-10-821-234-301
; Sequence 301, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
;  APPLICANT: Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Andarmani, Susan
;  APPLICANT:  Tang, Y. Tom
;  TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;  FILE REFERENCE: 821A
;  CURRENT APPLICATION NUMBER: US/10/821,234
;  CURRENT FILING DATE:  2004-04-07
;  PRIOR APPLICATION NUMBER: US 60/462,047
;  PRIOR FILING DATE: 2003-04-07
;  NUMBER OF SEQ ID NOS: 1704
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 301
;   LENGTH: 3672
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-821-234-301

Alignment Scores:
Pred. No.:                13.4            Length:          3672
Score:                    78.00           Matches:         26
Percent Similarity:       48.5%           Conservative:    7
Best Local Similarity:    38.2%           Mismatches:      32
Query Match:              17.4%           Indels:          3
DB:                       8               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-821-234-301 (1-3672)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::     |||     |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
              :::       |||   |||   |||       |||:::|||         |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
```

```
                    |||   ||||||   |||         |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                |||   :::     |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 4
US-11-091-883-189/c
; Sequence 189, Application US/11091883
; Publication No. US20060024693A1
; GENERAL INFORMATION:
;  APPLICANT: CIBELLI, JOSE
;  APPLICANT:  FERNANDEZ, EMILIO O.
;  APPLICANT:  JORDAO DE MEGALHAES, GUILHERME
;  APPLICANT:  KOCABAS, ARIF
;  APPLICANT:  CROSBY, JAVIER A.
;  TITLE OF INVENTION: IDENTIFICATION OF GENES OR POLYPEPTIDES THE EXPRESSION OF WHICH
;  TITLE OF INVENTION:  CORRELATES TO FERTILITY, OVARIAN FUNCTION AND/OR FETAL/NEWBORN
;  TITLE OF INVENTION:  VIABILITY
;  FILE REFERENCE: 53942US
;  CURRENT APPLICATION NUMBER: US/11/091,883
;  CURRENT FILING DATE:  2005-03-29
;  PRIOR APPLICATION NUMBER: 60/556,875
;  PRIOR FILING DATE: 2004-03-29
;  NUMBER OF SEQ ID NOS: 513
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 189
;   LENGTH: 4140
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (319)..(319)
;   OTHER INFORMATION: a, c, g, or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (690)..(690)
;   OTHER INFORMATION: a, c, g, or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1263)..(1263)
;   OTHER INFORMATION: a, c, g, or t
US-11-091-883-189

Alignment Scores:
Pred. No.:              15.4          Length:         4140
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     12            Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-091-883-189 (1-4140)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::    |||      |||
Db       2224 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 2165

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
              :::      |||   |||   |||       |||:::|||          |||||||||
Db       2164 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 2105

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||         |||   ||||||:::|||:::
Db       2104 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 2045

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                |||   :::     |||
Db       2044 TGTCTGGACTTTGAGCAGGCTT 2023

RESULT 5
US-10-932-182A-487
; Sequence 487, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 487
;   LENGTH: 1806
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-487

Alignment Scores:
Pred. No.:              35.4          Length:         1806
Score:                  72.00         Matches:        27
Percent Similarity:     38.8%         Conservative:   11
Best Local Similarity:  27.6%         Mismatches:     34
Query Match:            16.1%         Indels:         26
DB:                     7             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-487 (1-1806)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              :::|||      :::   |||      :::      |||:::            |||
Db        397 TTAGAAAGTGTAATTCAAAGCATGGTTCAAAGCGTGGTCCAAAGCGTGGTCACGGAA--- 453

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
                    ||||||         |||      |||                 |||
Db        454 -----------ATTAGCTATGACACCACCGTACCTGACACGGATGCTACTGCCACTCAA 501

Qy         41 AspSerThrThrSerLysIle-----------------------------------Ser 48
              |||||||||:::|||   :::                                   |||
Db        502 GACAGTACATCTTCTGCTTTACAGACGATTGATTGCACCGAGTACGGTACATTTGATAGC 561

Qy         49 ProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle---------- 64
                 :::      |||   |||   |||   :::   ||||||:::
Db        562 TCCAACTCTTCTACTTTTCAATCCAATAAGACCGATTTTTCGATCACCTATGGTGATTTG 621

Qy         65 ---------------ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 77
                           ||||||      |||   |||:::|||||||||||||:::
Db        622 ACCTTTGCCTCTGGCACGTGGGGCCGTGATCAATTGAACTTGAACAATTTGAAC 675

RESULT 6
US-10-932-182A-487
; Sequence 487, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 487
;   LENGTH: 1806
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-487

Alignment Scores:
```

```
Pred. No.:              35.4          Length:         1806
Score:                  72.00         Matches:        27
Percent Similarity:     38.8%         Conservative:   11
Best Local Similarity:  27.6%         Mismatches:     34
Query Match:            16.1%         Indels:         26
DB:                     7             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-487 (1-1806)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              :::|||     :::   |||     :::     |||:::         |||
Db        397 TTAGAAAGTGTAATTCAAAGCATGGTTCAAAGCGTGGTCCAAAGCGTGGTCACGGAA--- 453

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
                      ||||||       |||     |||                 |||
Db        454 ------------ATTAGCTATGACACCACCGTACCTGACACGGATGCTACTGCCACTCAA 501

Qy         41 AspSerThrThrSerLysIle----------------------------------Ser 48
              |||||||||:::|||   :::                                 |||
Db        502 GACAGTACATCTTCTGCTTTACAGACGATTGATTGCACCGAGTACGGTACATTTGATAGC 561

Qy         49 ProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle------------ 64
                :::     |||   |||   |||   :::   ||||||:::
Db        562 TCCAACTCTTCTACTTTTCAATCCAATAAGACCGATTTTTCGATCACCTATGGTGATTTG 621

Qy         65 --------------ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 77
                           ||||||       |||   |||:::|||||||||||:::
Db        622 ACCTTTGCCTCTGGCACGTGGGGCCGTGATCAATTGAACTTGAACAATTTGAAC 675

RESULT 7
US-11-121-086-88
; Sequence 88, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 88
;   LENGTH: 184868
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-88

Alignment Scores:
Pred. No.:              1.3e+04       Length:         184868
Score:                  70.50         Matches:        21
Percent Similarity:     44.6%         Conservative:   12
Best Local Similarity:  28.4%         Mismatches:     26
Query Match:            15.7%         Indels:         15
DB:                     12            Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-11-121-086-88 (1-184868)

Qy         16 SerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThr 35
              ||||||   |||     ||||||   |||:::  |||           |||
Db      89444 TCTGCAGCTGAAGCAACCCCTGAAAGGATTAATAGCCCAAATCTGTCATCTGACAGCATT 89503

Qy         36 AsnGluGluThrThrAspSerThrThrSerLys--------------------IleSer 48
                    |||:::   ::::::   |||                          :::
Db      89504 CCCAACCGCTGCACTAACAGGTCTTCCTTCAAGGGGTTTTTGAGGGGCACATGCATGCCA 89563

Qy         49 ProSerGluAspThrGlnGlnGluAsnGlySer---MetPheSerLeuIleThrTrp--- 66
              |||   :::   |||       :::|||:::     ::::::   :::||||||
Db      89564 CCAGACAAGACAACTACTCCCTCACATGGAAATGGCTGCTGGGCATTCCTCACTTGGAAA 89623

Qy         67 ------------------AsnIleAspGlyLeuAspLeuAsn 74
                                |||:::|||   ||||||||||||
Db      89624 GAGCTTGCCCTTCTCTGCAACGTCGATCATTTAGACCTCAAT 89665

RESULT 8
US-09-925-065A-747874/c
; Sequence 747874, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 747874
;   LENGTH: 496
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-747874

Alignment Scores:
Pred. No.:              14.1          Length:         496
Score:                  70.00         Matches:        29
Percent Similarity:     48.7%         Conservative:   9
Best Local Similarity:  37.2%         Mismatches:     18
Query Match:            15.6%         Indels:         22
DB:                     6             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-747874 (1-496)

Qy         16 SerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThr 35
              ||||||:::||||||           |||:::       |||:::::::::|||
Db        324 TCAGCCATAGAAAGGAAT---GYGTTAATAGCATTCACAGCAACCTGGATGGAACTG--- 271

Qy         36 AsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGln 55
                    ||||||                    ||||||||||||||   ||||||:::
Db        270 ------GAGACT--------------------ATTAGTCCAAGTGAAGTAACCCAGGAA 238

Qy         56 ------GluAsnGlySerMetPheSerLeuIleThrTrpAsn----------------- 67
                         |||||||||||||||||:::   :::
Db        237 TGGAAAACCAAACATTCTATGTTCTCACTCATAAGTGGGAGCTAAGCTATGAGGATGCAA 178

Qy         68 ---------IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGly 82
                       |||   :::|||       |||       ||||||   ||||||
Db        177 AGGCATAAGAATGATACAATGGACTTTGCCAACTCAGGAGAAAGGGTGAGAGGG 124

RESULT 9
US-11-096-568A-7380/c
; Sequence 7380, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7380
;   LENGTH: 1194
;   TYPE: DNA
```

```
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1194)
;   OTHER INFORMATION: Ceres Seq. ID no. 15171516
US-11-096-568A-7380

Alignment Scores:
Pred. No.:                84.5          Length:         1194
Score:                    67.50         Matches:        21
Percent Similarity:       45.3%         Conservative:   18
Best Local Similarity:    24.4%         Mismatches:     32
Query Match:              15.1%         Indels:         15
DB:                       9             Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-096-568A-7380 (1-1194)

Qy          10 GluProProValGluGluSer-AlaLeuGluArgArgProGluThrIleSer-GluProL 29
               :::||||||   :::   |||   |||   |||:::      |||:::       |||:
Db         763 CAACCACCCTAGCAAGTGTCCCTGCCTCCGTAGACAAAAGTCCACCGTTGTAGGTTCCAC 704

Qy          29 ysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerP 49
               ::      :::|||              |||::::::|||            |||::::::|
Db         703 GAGAAAGCCTCGATGACATTACTCGCGAAAGCAGTGAC------------AAGTTGACAC 656

Qy          49 roSerGluAspThrGlnGlnGluAsn--------------------------GlySerM 60
               |||||      :::   :::   |||                          ||||||
Db         655 CTTCAAGTACTTCGGCAGAAACAAATAAAAGGAGACCCGAACAGATGTATTGTGGCTCAG 596

Qy          60 etPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgA 80
                 :::   :::||||||||   :::          :::      :::          |
Db         595 AATATGGAATGATTACTTGGAAACTCAGTAGAACGCCTATGAGAACCATAATTTCTGATG 536

Qy          80 laArgGlyValCys 84
               ||    :::|||
Db         535 CCAACAGAATTTGC 522

RESULT 10
US-10-485-517-118
; Sequence 118, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
;  APPLICANT: University of Sheffield
;  APPLICANT:  Biosynexus Incorporated
;  APPLICANT:  Foster, Simon
;  APPLICANT:  Mond, James
;  TITLE OF INVENTION: Antigenic Polypeptides
;  FILE REFERENCE: P100629WO
;  CURRENT APPLICATION NUMBER: US/10/485,517
;  CURRENT FILING DATE:  2004-02-02
;  PRIOR APPLICATION NUMBER: GB 0118825.9
;  PRIOR FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: GB 0200349.9
;  PRIOR FILING DATE: 2002-01-09
;  NUMBER OF SEQ ID NOS: 424
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 118
;   LENGTH: 2361
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-10-485-517-118

Alignment Scores:
Pred. No.:                189           Length:         2361
Score:                    67.50         Matches:        19
Percent Similarity:       47.5%         Conservative:   9
Best Local Similarity:    32.2%         Mismatches:     22
Query Match:              15.1%         Indels:         9
DB:                       8             Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-485-517-118 (1-2361)

Qy          24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
               |||:::      |||      |||   |||          ::::::||||||      |||
Db         200 ACTTTAGTTTCACCTACCGCTTATGCTGATACACCTCAAAAAGATACTACAGCTAAGACA 259

Qy          44 Thr---------SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               |||         ||||||  |||       :::::::|||   ::::::      |||
Db         260 ACATCTCATGATTCAAAAAAATCTAATGACGATGAAACTTCTAAGGATACTACAAGTAAA 319

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                              :::||||||      |||   ||||||   |||   :::
Db         320 -----------------GATATTGATAAAGCAGACAAAAATAATACAAGTAACCAA 358

RESULT 11
US-11-096-568A-12999
; Sequence 12999, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12999
;   LENGTH: 1410
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1410)
;   OTHER INFORMATION: Ceres Seq. ID no. 14316032
US-11-096-568A-12999

Alignment Scores:
Pred. No.:                120           Length:         1410
Score:                    67.00         Matches:        19
Percent Similarity:       57.4%         Conservative:   16
Best Local Similarity:    31.1%         Mismatches:     20
Query Match:              15.0%         Indels:         6
DB:                       9             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-11-096-568A-12999 (1-1410)

Qy           2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
               |||          :::|||      :::   |||   :::|||:::|||   |||::::::
Db         424 GAGCCGCAGGCCGACAGCTCAGACAAGAAGCCGGAGCAGGAGACGGCCTCGGAGAAGAAG 483

Qy          22 Pro------GluThrIleSerGlu------ProLysThrTyrValAspLeuThrAsnGlu 37
               |||      ||||||   ||||||      |||:::|||      :::      :::|||
Db         484 CCGGAGCAGGAGACGGCCTCGGAGAAGAAGCCTGAGACGGAGAAGGAAGCAGAGTCTGAG 543

Qy          38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
               :::      |||         :::|||      ||||||::::::   :::::::::
Db         544 AAGAAGGCTGACAAGGAGGAGGCGAAA------CCAAGCGACGAGGCAAAGAAAGACGGC 597

Qy          58 Gly 58
               |||
Db         598 GGC 600

RESULT 12
US-11-044-051-74/c
; Sequence 74, Application US/11044051
; Publication No. US20050255553A1
; GENERAL INFORMATION:
;  APPLICANT: VAN PEL, Aline
;  APPLICANT:  GODELAINE, DaniSle
;  APPLICANT:  CARRASCO, Javier
;  APPLICANT:  BRASSEUR, Francis
;  APPLICANT:  BOON-FALLEUR, Thierry
;  TITLE OF INVENTION: MAGE C2 ANTIGENIC PEPTIDES AND USES THEREOF
;  FILE REFERENCE: LUD 5888 US
;  CURRENT APPLICATION NUMBER: US/11/044,051
;  CURRENT FILING DATE:  2005-01-28
```

```
;  PRIOR APPLICATION NUMBER: US 60/459,263
;  PRIOR FILING DATE: 2004-02-09
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 74
;   LENGTH: 4265
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-044-051-74

Alignment Scores:
Pred. No.:                441             Length:         4265
Score:                    67.00           Matches:        29
Percent Similarity:       46.0%           Conservative:   17
Best Local Similarity:    29.0%           Mismatches:     31
Query Match:              15.0%           Indels:         23
DB:                       12              Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-11-044-051-74 (1-4265)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   :::             :::   ||||||  ||||||  :::   |||
Db       2123 GAGGAAAGTAGAGAGGAGACATGGAGTCCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGT 2064

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr------ 39
              ||||||   |||   |||           :::   ::::::|||||||||
Db       2063 GAGGAGACAGGGAGTCCTCCCCCTGAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGT 2004

Qy         40 ------------------ThrAspSerThrThrSerLysIle---SerPro--------- 49
                              ::::::  ::::::   :::|||  ||||||
Db       2003 CCTCCCCCTGAGGAGGGCTCTGAGGAAAGTAGTGAGGAGATAGGGAGTCCTCCCCCTGAG 1944

Qy         50 ------SerGluAspThrGlnGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
                  ||||||::::::   :::|||   |||||| :::             ||||||
Db       1943 GAGGGCTCTGAGGAAAGTAGTGAGGAGACAGGGAGTCCTCCCACTCAGGAAGACTTGG-- 1886

Qy         67 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                          |||||||||   ||||||      |||:::      :::
Db       1885 ------------------AACAATCTG---GAGAGGAGAATGGGTATTCTCCCCTTC 1850

RESULT 13
US-11-136-527-6443
; Sequence 6443, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6443
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-6443

Alignment Scores:
Pred. No.:                50.8            Length:         600
Score:                    66.50           Matches:        16
Percent Similarity:       57.1%           Conservative:   16
Best Local Similarity:    28.6%           Mismatches:     17
Query Match:              14.8%           Indels:         7
DB:                       12              Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-136-527-6443 (1-600)

Qy          4 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg------ 21
              ::|||:::|||:::      |||||||||:::      :::      ::::::

Db          6 AGTCTCAGTTCACATACCCGGCCTCCTGTGCAGGGACGCAGCACCTCCCAGCAGGGTGAA 65

Qy         22 ---ProGluThrIleSerGluProLysThrTyrValAsp-----------LeuThrAsn 36
                 |||:::      :::   |||     :::::::::           |||||||||
Db         66 GAACCCCAAAAGACAGCATCACCTCACCCCCACCTCAACAAATCTCAGTCCCTGACTAAC 125

Qy         37 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 52
                 :::|||      |||:::   :::   :::|||||||||||||
Db        126 AGCCTCAGCACGTCAGATACCTCCCATCGAGGAACCCCAAGTGAAGAT 173

RESULT 14
US-09-925-065A-610817/c
; Sequence 610817, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 610817
;   LENGTH: 664
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-610817

Alignment Scores:
Pred. No.:                57.3            Length:         664
Score:                    66.50           Matches:        19
Percent Similarity:       51.3%           Conservative:   20
Best Local Similarity:    25.0%           Mismatches:     24
Query Match:              14.8%           Indels:         13
DB:                       6               Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-610817 (1-664)

Qy         10 GluProProValGluGlu----------SerAlaLeuGluArgArgProGluThr 24
              :::|||||:::::|||                :::|||||     |||
Db        521 AAGCCACCACTACAAGAACTGCTAAAAGGAGCTTTAAGTCTTGAAACAAATCCTGGAAAC 462

Qy         25 IleSerGluProLysThrTyr----------ValAspLeuThrAsnGluGluThr 39
                 |||:::      :::              ||||||   |||:::   |||
Db        461 ACATCAAAACAGAACSTCTTTAAAGCATAAATCACACAGGACCTATGAAACAAAAATACC 402

Qy         40 Thr---------AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 56
              :::         :::      |||   :::   |||::::::      :::
Db        401 AGTTAAAAAGCAAAAACAATAAAACCAAAGTACACAGGCAGCAAAGAGCACCATGAATGC 342

Qy         57 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 72
              ||||||:::   :::   ::::::|||   |||::::::||||||:::
Db        341 AATGGTACCTCTTATCCAATACTAACATTGAATGTAAATGGCCTAAAT 294

RESULT 15
US-10-932-182A-3877/c
; Sequence 3877, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
```

```
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3877
;   LENGTH: 810
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-3877

Alignment Scores:
Pred. No.:              72.4          Length:         810
Score:                  66.50         Matches:        29
Percent Similarity:     39.4%         Conservative:   10
Best Local Similarity:  29.3%         Mismatches:     31
Query Match:            14.8%         Indels:         29
DB:                     7             Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-3877 (1-810)

Qy          3 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 22
              ||||||:::|||:::     |||||||||          |||                |||
Db        494 AGGGCGATCAATACA------GAACCACCGTACTTCTTTAGC--------------CCC 456

Qy         23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
                 ||||||  ||||||   :::      ||||||                  :::
Db        455 TCCTGAATATCCGCTCCCAAATGGCATTGCAGTTTGACGCGCCTTTTGGAAGTTGAACGA 396

Qy         43 ThrThr--------------------------------SerLysIleSerProSerGlu 51
                |||                                        ||||||   |||
Db        395 AAAACATAAGTAGCATCCTTCGATTGCAGACGACCTTCTTTGTCCATTTCAAAGTCCAAC 336

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr----------------- 65
                |||      :::         ||||||:::   |||
Db        335 GGCAACCAAACAATTTCTTCGTTGTTTTTTAGTATATCAACTACATTTTTCCAGTGTTCA 276

Qy         66 --------------TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                           ::::::   |||||||||:::|||   |||:::|||   |||
Db        275 TTAAATTCTTTCTGATTTTCCCATGACGGTCTTGAATTAGAAAATATATCATACAGG 219

RESULT 16
US-10-932-182A-3877/c
; Sequence 3877, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3877
;   LENGTH: 810
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-3877

Alignment Scores:
Pred. No.:              72.4          Length:         810
Score:                  66.50         Matches:        29
Percent Similarity:     39.4%         Conservative:   10
Best Local Similarity:  29.3%         Mismatches:     31
Query Match:            14.8%         Indels:         29
DB:                     7             Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-3877 (1-810)

Qy          3 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 22
              ||||||:::|||:::     |||||||||          |||                |||
Db        494 AGGGCGATCAATACA------GAACCACCGTACTTCTTTAGC--------------CCC 456

Qy         23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
                 ||||||  ||||||   :::      ||||||                  :::
Db        455 TCCTGAATATCCGCTCCCAAATGGCATTGCAGTTTGACGCGCCTTTTGGAAGTTGAACGA 396

Qy         43 ThrThr--------------------------------SerLysIleSerProSerGlu 51
                |||                                        ||||||   |||
Db        395 AAAACATAAGTAGCATCCTTCGATTGCAGACGACCTTCTTTGTCCATTTCAAAGTCCAAC 336

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr----------------- 65
                |||      :::         ||||||:::   |||
Db        335 GGCAACCAAACAATTTCTTCGTTGTTTTTTAGTATATCAACTACATTTTTCCAGTGTTCA 276

Qy         66 --------------TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                           ::::::   |||||||||:::|||   |||:::|||   |||
Db        275 TTAAATTCTTTCTGATTTTCCCATGACGGTCTTGAATTAGAAAATATATCATACAGG 219

RESULT 17
US-10-932-182A-77405
; Sequence 77405, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77405
;   LENGTH: 2124
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-77405

Alignment Scores:
Pred. No.:              225           Length:         2124
Score:                  66.50         Matches:        21
Percent Similarity:     47.2%         Conservative:   4
Best Local Similarity:  39.6%         Mismatches:     21
Query Match:            14.8%         Indels:         7
DB:                     7             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-77405 (1-2124)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 26
              :::|||   |||      |||||||||          |||      ||||||   :::
Db        346 AATTACACTGAATTTACTGTTGAAGAATTAGTCAAAGAAATTAGTCCTGAATATTTAAGA 405

Qy         27 GluProLysThrTyrValAspLeuThrAsnGlu--------------------GluThr 39
                |||      ||||||||||||||||                        ||||||
Db        406 TTAAACAAACAGATTGATGACCTAACTAACGAACTAAGCAGAAAATCACAAATCGAAACC 465

Qy         40 ThrAspSerThrThrSerLysIleSerProSerGluAsp 52
              |||||||||          |||:::         :::|||
Db        466 ACGGATTCTATTTTTTCAAATTGATTAAAGAGAAGGAC 504

RESULT 18
US-10-932-182A-77405
; Sequence 77405, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
```

```
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77405
;   LENGTH: 2124
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-77405

Alignment Scores:
Pred. No.:                225             Length:         2124
Score:                    66.50           Matches:        21
Percent Similarity:       47.2%           Conservative:   4
Best Local Similarity:    39.6%           Mismatches:     21
Query Match:              14.8%           Indels:         7
DB:                       7               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-77405 (1-2124)

Qy           7 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 26
               :::|||   |||      |||||||||        |||     ||||||   :::
Db         346 AATTACACTGAATTTACTGTTGAAGAATTAGTCAAAGAAATTAGTCCTGAATATTTAAGA 405

Qy          27 GluProLysThrTyrValAspLeuThrAsnGlu--------------------GluThr 39
                |||          ||||||||||||||||                      ||||||
Db         406 TTAAACAAACAGATTGATGACCTAACTAACGAACTAAGCAGAAAATCACAAATCGAAACC 465

Qy          40 ThrAspSerThrThrSerLysIleSerProSerGluAsp 52
               |||||||||          |||:::        :::|||
Db         466 ACGGATTCTATTTTTTTCAAATTGATTAAAGAGAAGGAC 504

RESULT 19
US-11-136-527-2347
; Sequence 2347, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2347
;   LENGTH: 2270
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-2347

Alignment Scores:
Pred. No.:                244             Length:         2270
Score:                    66.50           Matches:        16
Percent Similarity:       57.1%           Conservative:   16
Best Local Similarity:    28.6%           Mismatches:     17
Query Match:              14.8%           Indels:         7
DB:                       12              Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-136-527-2347 (1-2270)

Qy           4 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg------ 21
               :::|||:::|||:::   |||||||||:::      :::      ::::::
Db        1676 AGTCTCAGTTCACATACCCGGCCTCCTGTGCAGGGACGCAGCACCTCCCAGCAGGGTGAA 1735

Qy          22 ---ProGluThrIleSerGluProLysThrTyrValAsp-----------LeuThrAsn 36
                  |||:::     :::   |||     :::::::::           ||||||||||
Db        1736 GAACCCCAAAAGACAGCATCACCTCACCCCCACCTCAACAAATCTCAGTCCCTGACTAAC 1795

Qy          37 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 52
                  :::|||      |||:::   :::   :::||||||||||||
Db        1796 AGCCTCAGCACGTCAGATACCTCCCATCGAGGAACCCCAAGTGAAGAT 1843

RESULT 20
US-10-521-162-33/c
; Sequence 33, Application US/10521162
; Publication No. US20050287529A1
; GENERAL INFORMATION:
;  APPLICANT: Brandt, Kevin S.
;  TITLE OF INVENTION: FLEA AND TICK OCTOPAMINE RECEPTOR NUCLEIC ACID MOLECULES,
;  TITLE OF INVENTION:  PROTEINS AND USES THEREOF
;  FILE REFERENCE: FC-11-PCT
;  CURRENT APPLICATION NUMBER: US/10/521,162
;  CURRENT FILING DATE:  2005-01-13
;  PRIOR APPLICATION NUMBER: 60/319,402
;  PRIOR FILING DATE: 2003-07-17
;  PRIOR APPLICATION NUMBER: 60/426,601
;  PRIOR FILING DATE: 2003-11-15
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
;   LENGTH: 499
;   TYPE: DNA
;   ORGANISM: Rhipicephalus saguineus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(278)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: n = unknown
;   OTHER INFORMATION: Xaa at location 2 = Tyr
US-10-521-162-33

Alignment Scores:
Pred. No.:                47.6            Length:         499
Score:                    66.00           Matches:        16
Percent Similarity:       44.2%           Conservative:   7
Best Local Similarity:    30.8%           Mismatches:     23
Query Match:              14.7%           Indels:         6
DB:                       8               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-521-162-33 (1-499)

Qy           7 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 26
               |||:::||||||              :::   |||||||||
Db         200 TCCTTCTTTGAGGCGGCAGCGACACAGGATGCGCTTGAA-----------------GGC 159

Qy          27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 46
               :::|||||||          |||          :::|||      |||          :::
Db         158 CAGCCGAAAGTCCTTGCTGACAAGCACGTAGATGAGCGGGTTGATGGCCGAGTTGCAGTA 99

Qy          47 IleSerProSerGluAspThrGlnGlnGluAsnGly 58
                  |||   ||||||   :::|||:::   |||
Db          98 TCCGAGCCAGAAGAAGACCGAGAACAGCAGGTTGGG 63

RESULT 21
US-10-521-162-35
; Sequence 35, Application US/10521162
; Publication No. US20050287529A1
; GENERAL INFORMATION:
;  APPLICANT: Brandt, Kevin S.
;  TITLE OF INVENTION: FLEA AND TICK OCTOPAMINE RECEPTOR NUCLEIC ACID MOLECULES,
;  TITLE OF INVENTION:  PROTEINS AND USES THEREOF
;  FILE REFERENCE: FC-11-PCT
;  CURRENT APPLICATION NUMBER: US/10/521,162
;  CURRENT FILING DATE:  2005-01-13
;  PRIOR APPLICATION NUMBER: 60/319,402
```

```
;  PRIOR FILING DATE: 2003-07-17
;  PRIOR APPLICATION NUMBER: 60/426,601
;  PRIOR FILING DATE: 2003-11-15
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35
;   LENGTH: 499
;   TYPE: DNA
;   ORGANISM: Rhipicephalus saguineus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (492)..(492)
;   OTHER INFORMATION: n = unknown
US-10-521-162-35

Alignment Scores:
Pred. No.:                47.6          Length:          499
Score:                    66.00         Matches:         16
Percent Similarity:       44.2%         Conservative:    7
Best Local Similarity:    30.8%         Mismatches:      23
Query Match:              14.7%         Indels:          6
DB:                       8             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-521-162-35 (1-499)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 26
              |||:::||||||           :::   |||||||||
Db        300 TCCTTCTTTGAGGCGGCAGCGACACAGGATGCGCTTGAA-----------------GGC 341

Qy         27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 46
              :::||||||        |||         :::|||      |||         :::
Db        342 CAGCCGAAAGTCCTTGCTGACAAGCACGTAGATGAGCGGGTTGATGGCCGAGTTGCAGTA 401

Qy         47 IleSerProSerGluAspThrGlnGlnGluAsnGly 58
                 |||   ||||||   :::|||:::   |||
Db        402 TCCGAGCCAGAAGAAGACCGAGAACAGCAGGTTGGG 437

RESULT 22
US-11-037-243-11
; Sequence 11, Application US/11037243
; Publication No. US20050287546A1
; GENERAL INFORMATION:
;  APPLICANT: PLOWMAN, GREGORY
;  APPLICANT:  WHYTE, DAVID
;  APPLICANT:  CAENEPEEL, SEAN
;  APPLICANT:  CHARYDCZAK, GLEN
;  APPLICANT:  MANNING, GERARD
;  APPLICANT:  SUDARSANAM, SUCHA
;  TITLE OF INVENTION: NOVEL PROTEASES
;  FILE REFERENCE: 038602/1214
;  CURRENT APPLICATION NUMBER: US/11/037,243
;  CURRENT FILING DATE:  2005-05-26
;  PRIOR APPLICATION NUMBER: US/09/888,615
;  PRIOR FILING DATE: 2001-06-26
;  PRIOR APPLICATION NUMBER: 60/214,047
;  PRIOR FILING DATE: 2000-06-26
;  NUMBER OF SEQ ID NOS: 150
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 870
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-037-243-11

Alignment Scores:
Pred. No.:                91.6          Length:          870
Score:                    66.00         Matches:         18
Percent Similarity:       49.0%         Conservative:    7
Best Local Similarity:    35.3%         Mismatches:      19
Query Match:              14.7%         Indels:          7
DB:                       12            Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-037-243-11 (1-870)

Qy         23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
              :::   :::::::::       |||       ||||||   ||||||          |||
Db        616 CAGAACTTGGCACAGACTTATACTCTTACTGATCTGATGAATGAGATCAAAGAAAGTAGT 675

Qy         43 ThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
              |||       ||||||   ||||||    |||:::|||   :::
Db        676 ACAAAACTCAAGATTTTTCCTTCCTCAGACTCTCAGCTGGACCCA--------------- 720

Qy         63 LeuIleThr-TrpAsnIleAspGlyLeuAsp 72
                 :::    ||||||      |||||||||
Db        721 ---TTGGTGGTGGAACTTTCAAGGCCTGGAC 748

RESULT 23
US-10-521-162-39/c
; Sequence 39, Application US/10521162
; Publication No. US20050287529A1
; GENERAL INFORMATION:
;  APPLICANT: Brandt, Kevin S.
;  TITLE OF INVENTION: FLEA AND TICK OCTOPAMINE RECEPTOR NUCLEIC ACID MOLECULES,
;  TITLE OF INVENTION:  PROTEINS AND USES THEREOF
;  FILE REFERENCE: FC-11-PCT
;  CURRENT APPLICATION NUMBER: US/10/521,162
;  CURRENT FILING DATE:  2005-01-13
;  PRIOR APPLICATION NUMBER: 60/319,402
;  PRIOR FILING DATE: 2003-07-17
;  PRIOR APPLICATION NUMBER: 60/426,601
;  PRIOR FILING DATE: 2003-11-15
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39
;   LENGTH: 1443
;   TYPE: DNA
;   ORGANISM: Rhipicephalus sanguineus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1440)
US-10-521-162-39

Alignment Scores:
Pred. No.:                166           Length:          1443
Score:                    66.00         Matches:         16
Percent Similarity:       44.2%         Conservative:    7
Best Local Similarity:    30.8%         Mismatches:      23
Query Match:              14.7%         Indels:          6
DB:                       8             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-521-162-39 (1-1443)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 26
              |||:::||||||           :::   |||||||||
Db       1362 TCCTTCTTTGAGGCGGCAGCGACACAGGATGCGCTTGAA-----------------GGC 1321

Qy         27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 46
              :::||||||        |||         :::|||      |||         :::
Db       1320 CAGCCGAAAGTCCTTGCTGACAAGCACGTAGATGAGCGGGTTGATGGCCGAGTTGCAGTA 1261

Qy         47 IleSerProSerGluAspThrGlnGlnGluAsnGly 58
                 |||   ||||||   :::|||:::   |||
Db       1260 TCCGAGCCAGAAGAAGACCGAGAACAGCAGGTTGGG 1225

RESULT 24
US-10-521-162-41
; Sequence 41, Application US/10521162
; Publication No. US20050287529A1
; GENERAL INFORMATION:
;  APPLICANT: Brandt, Kevin S.
;  TITLE OF INVENTION: FLEA AND TICK OCTOPAMINE RECEPTOR NUCLEIC ACID MOLECULES,
;  TITLE OF INVENTION:  PROTEINS AND USES THEREOF
;  FILE REFERENCE: FC-11-PCT
;  CURRENT APPLICATION NUMBER: US/10/521,162
;  CURRENT FILING DATE:  2005-01-13
;  PRIOR APPLICATION NUMBER: 60/319,402
```

```
;  PRIOR FILING DATE: 2003-07-17
;  PRIOR APPLICATION NUMBER: 60/426,601
;  PRIOR FILING DATE: 2003-11-15
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41
;   LENGTH: 1443
;   TYPE: DNA
;   ORGANISM: Rhipicephalus saguineus
US-10-521-162-41

Alignment Scores:
Pred. No.:                166             Length:         1443
Score:                    66.00           Matches:        16
Percent Similarity:       44.2%           Conservative:   7
Best Local Similarity:    30.8%           Mismatches:     23
Query Match:              14.7%           Indels:         6
DB:                       8               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-521-162-41 (1-1443)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 26
              |||:::|||||||             :::   |||||||||
Db         82 TCCTTCTTTGAGGCGGCAGCGACACAGGATGCGCTTGAA-----------------GGC 123

Qy         27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 46
              :::||||||          |||         :::|||      |||          :::
Db        124 CAGCCGAAAGTCCTTGCTGACAAGCACGTAGATGAGCGGGTTGATGGCCGAGTTGCAGTA 183

Qy         47 IleSerProSerGluAspThrGlnGlnGluAsnGly 58
                 |||   ||||||   :::|||:::   |||
Db        184 TCCGAGCCAGAAGAAGACCGAGAACAGCAGGTTGGG 219

RESULT 25
US-10-932-182A-77604
; Sequence 77604, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77604
;   LENGTH: 2121
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-77604

Alignment Scores:
Pred. No.:                262             Length:         2121
Score:                    66.00           Matches:        27
Percent Similarity:       42.6%           Conservative:   16
Best Local Similarity:    26.7%           Mismatches:     34
Query Match:              14.7%           Indels:         24
DB:                       7               Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-77604 (1-2121)

Qy          9 PheGluPro---ProValGluGluSerAlaLeuGluArgArgPro--------------- 22
              |||||||||   |||      :::      ||||||      |||
Db        769 TTTGAACCAGATCCTAAACTACAGGTGAATCTTGAAAATGAACCACTTCAAGATGATTTT 828

Qy         23 -----------------GluThrIleSerGluProLys---ThrTyrValAspLeuThr 35
                              |||      :::|||||||||         ::::::::::|||
Db        829 CCAGAAGCAAAACAAGAGGAAAAGAATACAGAACCTAAAATTCCGGAAATTAATGTTACG 888

Qy         36 AsnGluGluThrThrAspSerThrThr---------------SerLysIleSerProSer 50
              |||       |||   |||   |||                 |||||||||   |||
Db        889 AGGGAAAGTAATACACCATCTTTAACCATGAACGCACTAGACTCTAAAATATATCCTGAT 948

Qy         51 -----------------GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                              |||   ::::::::::::                  |||
Db        949 GATAACTTTTCTGGCCTGGAATCTTCAAAGGAACAGAAATCACCTGGTGTGTCATCATCG 1008

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCys 84
              :::     ::::::   |||   |||:::   |||:::|||:::         :::
Db       1009 TCAACAAAAGTTGAAGATTTAAGTCTTGATGGGTTAAATGAGAAACGTCTTTCAATAACT 1068

Qy         85 Ser 85
              |||
Db       1069 TCT 1071

RESULT 26
US-10-932-182A-77604
; Sequence 77604, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77604
;   LENGTH: 2121
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-77604

Alignment Scores:
Pred. No.:                262             Length:         2121
Score:                    66.00           Matches:        27
Percent Similarity:       42.6%           Conservative:   16
Best Local Similarity:    26.7%           Mismatches:     34
Query Match:              14.7%           Indels:         24
DB:                       7               Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-77604 (1-2121)

Qy          9 PheGluPro---ProValGluGluSerAlaLeuGluArgArgPro--------------- 22
              |||||||||   |||      :::      ||||||      |||
Db        769 TTTGAACCAGATCCTAAACTACAGGTGAATCTTGAAAATGAACCACTTCAAGATGATTTT 828

Qy         23 -----------------GluThrIleSerGluProLys---ThrTyrValAspLeuThr 35
                              |||      :::|||||||||         ::::::::::|||
Db        829 CCAGAAGCAAAACAAGAGGAAAAGAATACAGAACCTAAAATTCCGGAAATTAATGTTACG 888

Qy         36 AsnGluGluThrThrAspSerThrThr---------------SerLysIleSerProSer 50
              |||       |||   |||   |||                 |||||||||   |||
Db        889 AGGGAAAGTAATACACCATCTTTAACCATGAACGCACTAGACTCTAAAATATATCCTGAT 948

Qy         51 -----------------GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                              |||   ::::::::::::                  |||
Db        949 GATAACTTTTCTGGCCTGGAATCTTCAAAGGAACAGAAATCACCTGGTGTGTCATCATCG 1008

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCys 84
              :::     ::::::   |||   |||:::   |||:::|||:::         :::
Db       1009 TCAACAAAAGTTGAAGATTTAAGTCTTGATGGGTTAAATGAGAAACGTCTTTCAATAACT 1068

Qy         85 Ser 85
              |||
Db       1069 TCT 1071

RESULT 27
```

```
US-11-037-243-8
; Sequence 8, Application US/11037243
; Publication No. US20050287546A1
; GENERAL INFORMATION:
;  APPLICANT: PLOWMAN, GREGORY
;  APPLICANT:  WHYTE, DAVID
;  APPLICANT:  CAENEPEEL, SEAN
;  APPLICANT:  CHARYDCZAK, GLEN
;  APPLICANT:  MANNING, GERARD
;  APPLICANT:  SUDARSANAM, SUCHA
;  TITLE OF INVENTION: NOVEL PROTEASES
;  FILE REFERENCE: 038602/1214
;  CURRENT APPLICATION NUMBER: US/11/037,243
;  CURRENT FILING DATE:  2005-05-26
;  PRIOR APPLICATION NUMBER: US/09/888,615
;  PRIOR FILING DATE: 2001-06-26
;  PRIOR APPLICATION NUMBER: 60/214,047
;  PRIOR FILING DATE: 2000-06-26
;  NUMBER OF SEQ ID NOS: 150
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 2352
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-037-243-8

Alignment Scores:
Pred. No.:              295          Length:         2352
Score:                  66.00        Matches:        18
Percent Similarity:     49.0%        Conservative:   7
Best Local Similarity:  35.3%        Mismatches:     19
Query Match:            14.7%        Indels:         7
DB:                     12           Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-037-243-8 (1-2352)

Qy         23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
              :::   :::::::::      |||      ||||||   ||||||           |||
Db        616 CAGAACTTGGCACAGACTTATACTCTTACTGATCTGATGAATGAGATCAAAGAAAGTAGT 675

Qy         43 ThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
              |||      ||||||   ||||||   |||:::|||   :::
Db        676 ACAAAACTCAAGATTTTTCCTTCCTCAGACTCTCAGCTGGACCCA--------------- 720

Qy         63 LeuIleThr-TrpAsnIleAspGlyLeuAsp 72
                 :::    ||||||      |||||||
Db        721 ---TTGGTGGTGGAACTTTCAAGGCCTGGAC 748

RESULT 28
US-10-873-528-198
; Sequence 198, Application US/10873528
; Publication No. US20050276814A1
; GENERAL INFORMATION:
;  APPLICANT: Microbial Technics Limited
;  APPLICANT:  Gilbert, Christophe FG
;  APPLICANT:  Hansbro, Philip M
;  TITLE OF INVENTION: Proteins
;  FILE REFERENCE: PWC/P21129WO
;  CURRENT APPLICATION NUMBER: US/10/873,528
;  CURRENT FILING DATE:  2004-06-23
;  PRIOR APPLICATION NUMBER: US/09/769,787
;  PRIOR FILING DATE: 2001-01-26
;  PRIOR APPLICATION NUMBER: GB 9816337.1
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: US 60/125164
;  PRIOR FILING DATE: 1999-03-19
;  NUMBER OF SEQ ID NOS: 388
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 198
;   LENGTH: 390
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
US-10-873-528-198

Alignment Scores:
Pred. No.:              48.1         Length:         390
Score:                  65.00        Matches:        17
Percent Similarity:     50.0%        Conservative:   9
Best Local Similarity:  32.7%        Mismatches:     20
Query Match:            14.5%        Indels:         6
DB:                     8            Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-873-528-198 (1-390)

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              ||||||      |||   |||          |||   :::|||:::   :::   |||||
Db         55 CCTGAGCAATCATCTATTCCTTCGTATCCTGTATCTGTTACATCTAACCAAGGAACAGAT 114

Qy         42 ---------SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 58
                                  ::::::|||:::   |||   :::|||||||||||
Db        115 GTAGCAGTAGAACCAGCTAAAGCAGTTGCTCCAACAACAGACTGGAAACAAGAAAATGGT 174

Qy         59 SerMetPheSerLeuIleThrTrpAsnIleAspGly 70
                    :::            :::|||   ||||||
Db        175 ATGTGGTAT---------TTTTATAATACTGATGGT 201

RESULT 29
US-09-925-065A-887050
; Sequence 887050, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 887050
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-887050

Alignment Scores:
Pred. No.:              77.6         Length:         585
Score:                  65.00        Matches:        23
Percent Similarity:     41.2%        Conservative:   5
Best Local Similarity:  33.8%        Mismatches:     22
Query Match:            14.5%        Indels:         18
DB:                     6            Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-887050 (1-585)

Qy         10 GluProProValGlu--------------------GluSerAlaLeuGluArgArgPro 22
              |||||||||                          ||||||   |||||||||
Db         53 GAGCCCCCACACAGAGTTCCTACTGGGGCATGACCTAGTGGAGCTTTGAGAAGAAGGCCA 112

Qy         23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
              :::   |||   |||:::                  |||         |||||||||
Db        113 CCATCC---TCCAGGCCTCAG-----------------AACGGTAGATTGACTGATAGC 151

Qy         43 ThrThrSerLysIleSerProSerGluAspThrGln------------GlnGluAsnGly 58
              |||                 :::   |||||||||             |||      |||
Db        152 TTGACCCATGCRCCTGGAATAGCAGCAGACACTCAAAGCCTGTGAAAGCAGCCAGGAGGG 211
```

```
Qy          59 SerMetPheSerLeuIleThrTrp 66
               ::::::|||   |||   |||
Db         212 AGGCTCTACTCTGCAATCCCATGG 235
```

RESULT 30
US-09-925-065A-887051
; Sequence 887051, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 887051
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-887051

Alignment Scores:

Pred. No.:	77.6	Length:	585
Score:	65.00	Matches:	23
Percent Similarity:	41.2%	Conservative:	5
Best Local Similarity:	33.8%	Mismatches:	22
Query Match:	14.5%	Indels:	18
DB:	6	Gaps:	4

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-887051 (1-585)

```
Qy          10 GluProProValGlu--------------------GluSerAlaLeuGluArgArgPro 22
               |||||||||                          ||||||   ||||||||||
Db          53 GAGCCCCCACACAGAGTTCCTACTGGGGCATGACCTAGTGGAGCTTTGAGAAGAAGGCCA 112

Qy          23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
               :::   |||   |||:::                 |||      |||||||||
Db         113 CCATCC---TCCAGGCCTCAG-----------------AACGGTAGATTGACTGATAGC 151

Qy          43 ThrThrSerLysIleSerProSerGluAspThrGln------------GlnGluAsnGly 58
               |||                 :::   |||||||||            |||      |||
Db         152 TTGACCCATGCGCCTGGAATAGCAGCAGACACTCAAAGCCTGTGAAAGCAGCCAGGAGGG 211

Qy          59 SerMetPheSerLeuIleThrTrp 66
               :::::::|||   |||   |||
Db         212 AGGCTCTACTCTGCAATCCCATGG 235
```

RESULT 31
US-10-932-182A-4655
; Sequence 4655, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4655
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-4655

Alignment Scores:

Pred. No.:	159	Length:	1077
Score:	65.00	Matches:	21
Percent Similarity:	42.9%	Conservative:	12
Best Local Similarity:	27.3%	Mismatches:	32
Query Match:	14.5%	Indels:	12
DB:	7	Gaps:	3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-4655 (1-1077)

```
Qy           4 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 23
               ||||||      |||     |||   :::     ||||||   :::   |||
Db         802 GCACTAGAGGATGTATTTGCGACGCCC---AAAAGTGGAGCTTTGTCACAATTACCAAGC 858

Qy          24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
               |||                          |||      |||||||||:::   |||
Db         859 ACGTTC--------------------GACAAAAATAATGAAGAGTCAGATGATTGTCGT 897

Qy          44 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 63
               :::                          |||   ::::::||||||   ::::::   |||
Db         898 TCAGTTAATGACGACAAGGTAGTCACAACGGACAAAGACAATGGGGACTTGTATAAACTG 957

Qy          64 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                        :::     |||::::::|||:::   |||   |||
Db         958 ------------CAGAATAATTACGACATGAGCAACATTGGGGAAATTGCT 996
```

RESULT 32
US-10-932-182A-4655
; Sequence 4655, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4655
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-4655

Alignment Scores:

Pred. No.:	159	Length:	1077
Score:	65.00	Matches:	21
Percent Similarity:	42.9%	Conservative:	12
Best Local Similarity:	27.3%	Mismatches:	32
Query Match:	14.5%	Indels:	12
DB:	7	Gaps:	3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-4655 (1-1077)

```
Qy           4 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 23
               ||||||      |||     |||   :::     ||||||   :::   |||
Db         802 GCACTAGAGGATGTATTTGCGACGCCC---AAAAGTGGAGCTTTGTCACAATTACCAAGC 858

Qy          24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
               |||                          |||      |||||||||:::   |||
```

```
Db        859 ACGTTC--------------------GACAAAAATAATGAAGAGTCAGATGATTGTCGT 897

Qy         44 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 63
              :::                          |||   ::::::||||||   ::::::   |||
Db        898 TCAGTTAATGACGACAAGGTAGTCACAACGGACAAAGACAATGGGGACTTGTATAAACTG 957

Qy         64 IleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                        :::      |||::::::|||:::   |||   |||
Db        958 ------------CAGAATAATTACGACATGAGCAACATTGGGGAAATTGCT 996
```

RESULT 33
US-11-136-527-751
; Sequence 751, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 751
; LENGTH: 3741
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-751

Alignment Scores:

Pred. No.:	690	Length:	3741
Score:	65.00	Matches:	26
Percent Similarity:	42.7%	Conservative:	12
Best Local Similarity:	29.2%	Mismatches:	33
Query Match:	14.5%	Indels:	18
DB:	12	Gaps:	5

US-10-757-745-2_COPY_54_140 (1-87) x US-11-136-527-751 (1-3741)

```
Qy         10 GluProProValGluGluSerAlaLeuGlu------ArgArgProGluThrIleSerGlu 27
              :::|||   ::::::|||   :::|||        ||||||            ||||||
Db       1975 CAGCCCAAAATGGATGAGCAGAGCCTGARTACTTGYCGGCGGATCAGTGAGACATCAGAG 2034

Qy         28 ProLysThrTyrValAspLeu---ThrAsnGluGluThrThrAspSerThrThrSerLys 46
              :::|||   ||||||     :::         ||||||||||||   ||||||
Db       2035 GAYCAGACCTTGGTGGATCRACAGTCCTYCTCCCGGACCACAGACTCAGACACMAGC--- 2091

Qy         47 IleSerProSerGluAsp--------------ThrGlnGlnGluAsnGlySerMetPhe 61
              |||   |||                     :::       :::|||      :::
Db       2092 ATAGAGCCAGGGCYGGGCARAGTCAAAKCGAGAAGCRGGTCTAACAGYGGCCTGGAYTAC 2151

Qy         62 SerLeuIleThrTrpAsn--------------------------IleAspGlyLeuAsp 72
                 ||||||                               :::   ||||||
Db       2152 ATCAAARTCACCTGGAAGAGGCTCCGCTCACAYTCCAGACAGTATGTGTCCGGGYTGCAC 2211

Qy         73 LeuAsnAsnLeuSerGluArgAlaArg 81
              ||||||        :::   |||:::
Db       2212 TTGAACCGAGAGCGGAAGGCAGCCAAG 2238
```

RESULT 34
US-10-750-185-42446
; Sequence 42446, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 42446
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Bovine 19866881384575
US-10-750-185-42446

Alignment Scores:

Pred. No.:	365	Length:	1914
Score:	64.50	Matches:	17
Percent Similarity:	47.1%	Conservative:	7
Best Local Similarity:	33.3%	Mismatches:	16
Query Match:	14.4%	Indels:	11
DB:	8	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-42446 (1-1914)

```
Qy         48 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer--------------- 62
              ||||||   ::::::|||:::             :::|||
Db        949 TCTCCAGGCAAAAATACTAGAGTGGGTCGCCACTTCCTCTTCCAGGGTATCTTCCCCCTC 1008

Qy         63 -----------------LeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 76
                              |||:::   |||      |||   |||   ||||||:::|||
Db       1009 CAGAAATTGAACCTGCTTCTCCTGTATTGGCAGGCAGATTCTTTACCACTGAACCACCTT 1068

Qy         77 SerGluArgAlaArgGlyValCysSerTyrLeu 87
                    :::|||||||||||||||   |||
Db       1069 TGGCTCAGCTCAAGGGGAGTCTGTTCTGACCTC 1101
```

RESULT 35
US-10-750-623-42446
; Sequence 42446, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 42446
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Bovine 19866881384575
US-10-750-623-42446

Alignment Scores:

Pred. No.:	365	Length:	1914
Score:	64.50	Matches:	17
Percent Similarity:	47.1%	Conservative:	7
Best Local Similarity:	33.3%	Mismatches:	16
Query Match:	14.4%	Indels:	11
DB:	8	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-623-42446 (1-1914)

```
Qy         48 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer--------------- 62
```

```
                ||||||   ::::::|||:::                :::|||
Db        949 TCTCCAGGCAAAAATACTAGAGTGGGTCGCCACTTCCTCTTCCAGGGTATCTTCCCCCTC 1008

Qy         63 ------------------LeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 76
                               |||:::   |||      |||   |||   ||||||:::|||
Db       1009 CAGAAATTGAACCTGCTTCTCCTGTATTGGCAGGCAGATTCTTTACCACTGAACCACCTT 1068

Qy         77 SerGluArgAlaArgGlyValCysSerTyrLeu 87
                :::|||||||||||||||   |||
Db       1069 TGGCTCAGCTCAAGGGGAGTCTGTTCTGACCTC 1101

RESULT 36
US-10-932-182A-76026
; Sequence 76026, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76026
;   LENGTH: 3930
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-76026

Alignment Scores:
Pred. No.:                851             Length:         3930
Score:                    64.50           Matches:        24
Percent Similarity:       42.4%           Conservative:   12
Best Local Similarity:    28.2%           Mismatches:     34
Query Match:              14.4%           Indels:         15
DB:                       7               Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-76026 (1-3930)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              :::   |||||||||:::   ||||||   ||||||      |||   :::
Db       1393 CAAAATGCTTTGAATGCAACCTTTGAAACTCCCGTTACTCTTTCTCGGATTAATTTTGAA 1452

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||      :::      |||:::|||                 |||      :::
Db       1453 CCC---ATATTGGAAGTTCCTGAGACTAGTTCACCATCTAAGAATACGATGTCAAAACCC 1509

Qy         42 SerThrThrSerLysIleSerProSerGluAspThr-----------------GlnGln 55
              |||   :::|||   |||         :::||||||                  :::
Db       1510 TCGAATTCTTCACCTATTCCGAAGGAAAAAGATACATTTAATATACACGAGAGAGAAGTA 1569

Qy         56 GluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsn 75
              |||       ::::::||||||                          |||:::   |||
Db       1570 GAGACAAACAATGTTTTTTCA----------------------AACGATATACAAAAT 1605

Qy         76 LeuSerGluArgAla 80
                 |||      |||
Db       1606 TCTTCAAATGCAGCT 1620

RESULT 37
US-10-932-182A-76026
; Sequence 76026, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76026
;   LENGTH: 3930
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-76026

Alignment Scores:
Pred. No.:                851             Length:         3930
Score:                    64.50           Matches:        24
Percent Similarity:       42.4%           Conservative:   12
Best Local Similarity:    28.2%           Mismatches:     34
Query Match:              14.4%           Indels:         15
DB:                       7               Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-76026 (1-3930)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              :::   |||||||||:::   ||||||   ||||||      |||   :::
Db       1393 CAAAATGCTTTGAATGCAACCTTTGAAACTCCCGTTACTCTTTCTCGGATTAATTTTGAA 1452

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||      :::      |||:::|||                 |||      :::
Db       1453 CCC---ATATTGGAAGTTCCTGAGACTAGTTCACCATCTAAGAATACGATGTCAAAACCC 1509

Qy         42 SerThrThrSerLysIleSerProSerGluAspThr-----------------GlnGln 55
              |||   :::|||   |||         :::||||||                  :::
Db       1510 TCGAATTCTTCACCTATTCCGAAGGAAAAAGATACATTTAATATACACGAGAGAGAAGTA 1569

Qy         56 GluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsn 75
              |||       ::::::||||||                          |||:::   |||
Db       1570 GAGACAAACAATGTTTTTTCA----------------------AACGATATACAAAAT 1605

Qy         76 LeuSerGluArgAla 80
                 |||      |||
Db       1606 TCTTCAAATGCAGCT 1620

RESULT 38
US-09-925-065A-203999/c
; Sequence 203999, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 203999
;   LENGTH: 312
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-203999

Alignment Scores:
Pred. No.:                50              Length:         312
```

```
Score:                 64.00          Matches:       20
Percent Similarity:    48.5%          Conservative:  12
Best Local Similarity: 30.3%          Mismatches:    22
Query Match:           14.3%          Indels:        12
DB:                    6              Gaps:          3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-203999 (1-312)

Qy         18 LeuGluArgArgProGluThrIleSerGluProLysThrTyr--------------Val 32
              ||||||     |||       |||:::     :::
Db        238 CTTGAAACAAATCCTGGAAACACATCAAAACACAACCTCTTTAAAGCATAAATCTCACAG 179

Qy         33 AspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 52
              ||||||   |||:::   |||           ||||||||||||:::   ::::::
Db        178 GACCTATACAACAAAAATACAATK---AAAAAAACCAGCAAAATCACAAAAACCCAAGGG 122

Qy         53 ThrGlnGlnGluAsnGlySerMet----------------PheSerLeuIleThrTrp 66
                 |||       |||   :::|||                 |||::::::|||
Db        121 GTACAGGCAACAAATAGCACAATGAATGGAATGATACCTCACATCTCAGTACTAACATTG 62

Qy         67 AsnIleAspGlyLeuAsp 72
              |||::::::||||||:::
Db         61 AATGTAAATGGTCTAAAT 44

RESULT 39
US-10-932-182A-5731
; Sequence 5731, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5731
;   LENGTH: 861
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-5731

Alignment Scores:
Pred. No.:             165            Length:        861
Score:                 64.00          Matches:       25
Percent Similarity:    47.5%          Conservative:  13
Best Local Similarity: 31.2%          Mismatches:    32
Query Match:           14.3%          Indels:        10
DB:                    7              Gaps:          3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-5731 (1-861)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu---Pro 28
              ||||||            :::|||   |||   |||   ||||||   |||
Db         85 GAACCTAAGGAGCCGTTCACCGCCAAGGAAGAAAGAAGCGAGACTACAAGTGTATTTAGC 144

Qy         29 LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr--------------- 43
                 |||   |||:::         :::||||||      |||
Db        145 CCCGCTTACTCTGATATCGCCGCTACTGATACGACCAGAAAAACCGATGATAGCGAATAC 204

Qy         44 ---------ThrSerLysIleSerProSer---GluAspThrGlnGlnGluAsnGlySer 59
                      ||||||         ||||||   ::::::|||::::::      |||:::
Db        205 TACAATTTTACTAGTCATTTCATGCCTTCCCTAAAAAACACTAGAGAGCTGGAAGGCACA 264

Qy         60 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
              :::   :::||||||            :::|||   |||          :::||||||:::
Db        265 ATACTTAATCTTATACAGAAGATTAAGGAAGGGGATGACGAAACTTTGGTTAGTGAGAAA 324

RESULT 40
```

```
US-10-932-182A-5731
; Sequence 5731, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5731
;   LENGTH: 861
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-5731

Alignment Scores:
Pred. No.:             165            Length:        861
Score:                 64.00          Matches:       25
Percent Similarity:    47.5%          Conservative:  13
Best Local Similarity: 31.2%          Mismatches:    32
Query Match:           14.3%          Indels:        10
DB:                    7              Gaps:          3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-5731 (1-861)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu---Pro 28
              ||||||            :::|||   |||   |||   ||||||   |||
Db         85 GAACCTAAGGAGCCGTTCACCGCCAAGGAAGAAAGAAGCGAGACTACAAGTGTATTTAGC 144

Qy         29 LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr--------------- 43
                 |||   |||:::         :::||||||      |||
Db        145 CCCGCTTACTCTGATATCGCCGCTACTGATACGACCAGAAAAACCGATGATAGCGAATAC 204

Qy         44 ---------ThrSerLysIleSerProSer---GluAspThrGlnGlnGluAsnGlySer 59
                      ||||||         ||||||   ::::::|||::::::      |||:::
Db        205 TACAATTTTACTAGTCATTTCATGCCTTCCCTAAAAAACACTAGAGAGCTGGAAGGCACA 264

Qy         60 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
              :::   :::||||||            :::|||   |||          :::||||||:::
Db        265 ATACTTAATCTTATACAGAAGATTAAGGAAGGGGATGACGAAACTTTGGTTAGTGAGAAA 324

RESULT 41
US-10-454-437-141/c
; Sequence 141, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
;  APPLICANT: Pompejus, Markus
;  APPLICANT:  Kroger, Burkhard
;  APPLICANT:  Schroder, Hartwig
;  APPLICANT:  Zelder, Oskar
;  APPLICANT:  Haberhauer, Gregor
;  TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
;  TITLE OF INVENTION:  INVOLVED IN HOMEOSTASIS AND ADAPTATION
;  FILE REFERENCE: BGI-128CPCN
;  CURRENT APPLICATION NUMBER: US/10/454,437
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: US 60/141031
;  PRIOR FILING DATE: 1999-06-25
;  PRIOR APPLICATION NUMBER: DE 19931636.8
;  PRIOR FILING DATE: 1999-07-08
;  PRIOR APPLICATION NUMBER: DE 19932125.6
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932126.4
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932127.2
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932128.0
```

```
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932129.9
;  PRIOR FILING DATE: 1999-07-19
;  PRIOR APPLICATION NUMBER: DE 19932226.0
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: DE 19932920.6
;  PRIOR FILING DATE: 1999-07-14
;  PRIOR APPLICATION NUMBER: DE 19932922.2
;  PRIOR FILING DATE: 1999-07-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 141
;   LENGTH: 1098
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (101)..(1075)
;   OTHER INFORMATION: RXA02558
US-10-454-437-141

Alignment Scores:
Pred. No.:                220              Length:          1098
Score:                    64.00            Matches:         16
Percent Similarity:       46.8%            Conservative:    6
Best Local Similarity:    34.0%            Mismatches:      25
Query Match:              14.3%            Indels:          1
DB:                       8                Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-454-437-141 (1-1098)

Qy          12 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
               ||||||      |||   |||  |||     |||||||||   |||   |||
Db         842 CCGGTGCGTAGCTCAAGGAA-GAGATGCGGGCTTCGACGATTTCAAGACCTGCAACAGCT 784

Qy          32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
                                   :::   |||   ||||||   :::::::|||   :::
Db         783 GCTCGTTGTGCCACCTCATCTGCGAGTTCTTCACTGACCTCATCGGTAGCGCCACGCAAG 724

Qy          52 AspThrGlnGlnGluAsnGly 58
               :::|||:::       |||
Db         723 GAAACACGACCATCAACTGGG 703

RESULT 42
US-11-096-568A-19504
; Sequence 19504, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 19504
;   LENGTH: 1623
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1623)
;   OTHER INFORMATION: Ceres Seq. ID no. 12373419
US-11-096-568A-19504

Alignment Scores:
Pred. No.:                349              Length:          1623
Score:                    64.00            Matches:         22
Percent Similarity:       44.0%            Conservative:    11
Best Local Similarity:    29.3%            Mismatches:      12
Query Match:              14.3%            Indels:          30
DB:                       9                Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x US-11-096-568A-19504 (1-1623)

Qy          10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
               :::|||||||||   |||   |||:::::::::      ||||||      |||||||||
Db         706 GATCCTCCAGTTTTTGAGCAGGCCGTGCAAAAA------GAGACTGCTGATGAACCAAAG 759

Qy          30 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 49
                                     |||:::||||||                      |||
Db         760 ----------------------GAGTCCACTGAT------------------AAACCA 777

Qy          50 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 69
               |||   :::::::::                                        |||||||||
Db         778 AAGGAGTCCAGTGAAAAA------------------------------AATATTGAC 804

Qy          70 GlyLeu---------AspLeuAsnAsnLeuSerGluArgAlaArg 81
               |||         ||||||      :::   :::|||   :::
Db         805 GATTTAATTGATGAGGACCTGAAAGAAATAGGGGACAGGAAAAAG 849

RESULT 43
US-09-925-065A-711158/c
; Sequence 711158, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 711158
;   LENGTH: 1777
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-711158

Alignment Scores:
Pred. No.:                388              Length:          1777
Score:                    64.00            Matches:         15
Percent Similarity:       54.2%            Conservative:    11
Best Local Similarity:    31.2%            Mismatches:      22
Query Match:              14.3%            Indels:          0
DB:                       6                Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-711158 (1-1777)

Qy          12 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
               |||   ||||||:::|||   |||:::   ||||||      :::|||   |||
Db         989 CCCGCGGAGGAGGCCGCAGATGAGAAACCCCCAGAGGAGGCCGCAGAGGAGAAGCTCACA 930

Qy          32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
                  :::   |||   |||          :::      ||||||:::   :::   :::
Db         929 GAGGAGGCCACAGAGGAACCGGCCGCAGAGGAACCGACCTCAGAGGAGGCCGTGGCCCCC 870

Qy          52 AspThrGlnGlnGluAsnGlySer 59
               :::   :::       :::||||||
Db         869 GAGGGTAAGGGGGCTAGCGGCAGC 846

RESULT 44
US-10-932-182A-1233
; Sequence 1233, Application US/10932182A
; Publication No. US20060046253A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1233
;   LENGTH: 2430
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-1233

Alignment Scores:
Pred. No.:              562          Length:         2430
Score:                  64.00        Matches:        20
Percent Similarity:     43.4%        Conservative:   13
Best Local Similarity:  26.3%        Mismatches:     39
Query Match:            14.3%        Indels:         4
DB:                     7            Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-1233 (1-2430)

Qy         11 ProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThr 30
              ||||||:::      :::      :::   |||      ||||||:::   |||
Db       1330 CCCCCAATCAATAGCAACAATTATGAT---AGAGTTTCTACAATAAATCGTCCAAATAAC 1386

Qy         31 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSer 50
              |||||||||         :::                  |||      :::   ||||||
Db       1387 TATGTCGACTCAGAGACGCAAACTTTAGATAGTATGGCTACTCCCAACTTGAAACCATCC 1446

Qy         51 GluAspThrGlnGlnGluAsnGlySerMetPhe------SerLeuIleThrTrpAsnIle 68
                  |||      :::      |||             |||   :::      |||:::
Db       1447 CCTATTACTACTAATAAAACAACAAGCTTCTCGAAAGAGTCTACTGTTGCAGAAAATGTA 1506

Qy         69 AspGlyLeuAspLeu---AsnAsnLeuSerGluArgAlaArgGlyVal 83
                |||      :::      :::||||||   :::   :::   |||
Db       1507 ACTGGAGGAAGTATTGTACGTTCGCTAAGTTCTAAACTAAAAAGGGTT 1554

RESULT 45
US-10-932-182A-1233
; Sequence 1233, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1233
;   LENGTH: 2430
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-1233

Alignment Scores:
Pred. No.:              562          Length:         2430
Score:                  64.00        Matches:        20
Percent Similarity:     43.4%        Conservative:   13
Best Local Similarity:  26.3%        Mismatches:     39
Query Match:            14.3%        Indels:         4
DB:                     7            Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-1233 (1-2430)

Qy         11 ProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThr 30
              ||||||:::      :::      :::   |||      ||||||:::   |||
Db       1330 CCCCCAATCAATAGCAACAATTATGAT---AGAGTTTCTACAATAAATCGTCCAAATAAC 1386

Qy         31 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSer 50
              |||||||||         :::                  |||      :::   ||||||
Db       1387 TATGTCGACTCAGAGACGCAAACTTTAGATAGTATGGCTACTCCCAACTTGAAACCATCC 1446

Qy         51 GluAspThrGlnGlnGluAsnGlySerMetPhe------SerLeuIleThrTrpAsnIle 68
                  |||      :::      |||             |||   :::      |||:::
Db       1447 CCTATTACTACTAATAAAACAACAAGCTTCTCGAAAGAGTCTACTGTTGCAGAAAATGTA 1506

Qy         69 AspGlyLeuAspLeu---AsnAsnLeuSerGluArgAlaArgGlyVal 83
                |||      :::      :::||||||   :::   :::   |||
Db       1507 ACTGGAGGAAGTATTGTACGTTCGCTAAGTTCTAAACTAAAAAGGGTT 1554

RESULT 46
US-10-932-182A-79549
; Sequence 79549, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79549
;   LENGTH: 2457
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-79549

Alignment Scores:
Pred. No.:              569          Length:         2457
Score:                  64.00        Matches:        24
Percent Similarity:     43.8%        Conservative:   15
Best Local Similarity:  27.0%        Mismatches:     30
Query Match:            14.3%        Indels:         20
DB:                     7            Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-79549 (1-2457)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              :::|||      |||           ||||||  ||||||:::   :::|||   |||
Db        439 TTGGAAAATAAACTTGAAGATGCAACTGAGCCAGCAGTGGAGCAAAGATCCCTAAACCGC 498

Qy         21 ArgProGluThrIle----------------------------------------Ser 26
                 |||:::                                            :::
Db        499 TCCGATGAATCTGCCCATACAAGTTCAAGTCCAGAAGCCAACAATTCTTCCTTGCCAGCA 558

Qy         27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 46
                |||::::::      |||:::|||::::::|||:::          |||
Db        559 TCTCCAAGGTCAATTCCGCTACTTTCAAACAAGAAAACTTCAAAAATGATTGATTCTTCG 618

Qy         47 IleSerProSerGluAspThr-----------GlnGlnGluAsnGlySerMetPheSer 62
              :::            ||||||          |||  ||||||          |||
Db        619 CTA------GAATTCGATACTGCCCTCATAGATCAAGCAGAGAATCTAGGAAGGTTTTTA 672

Qy         63 LeuIleThrTrpAsnIleAspGlyLeu 71
              |||   :::   :::|||:::||||||
Db        673 TTATCGTCCGAGGACATAGAGGGACTG 699

RESULT 47
US-10-932-182A-79549
```

```
; Sequence 79549, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79549
;   LENGTH: 2457
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-79549

Alignment Scores:
Pred. No.:              569          Length:          2457
Score:                  64.00        Matches:           24
Percent Similarity:     43.8%        Conservative:      15
Best Local Similarity:  27.0%        Mismatches:        30
Query Match:            14.3%        Indels:            20
DB:                     7            Gaps:               3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-79549 (1-2457)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              :::|||      |||           ||||||  ||||||:::  :::|||  |||
Db        439 TTGGAAAATAAACTTGAAGATGCAACTGAGCCAGCAGTGGAGCAAAGATCCCTAAACCGC 498

Qy         21 ArgProGluThrIle-----------------------------------------Ser 26
                 |||:::                                             :::
Db        499 TCCGATGAATCTGCCCATACAAGTTCAAGTCCAGAAGCCAACAATTCTTCCTTGCCAGCA 558

Qy         27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 46
               |||::::::      |||:::|||:::::::|||:::          |||
Db        559 TCTCCAAGGTCAATTCCGCTACTTTCAAACAAGAAAACTTCAAAAATGATTGATTCTTCG 618

Qy         47 IleSerProSerGluAspThr----------GlnGlnGluAsnGlySerMetPheSer 62
              :::           ||||||          |||   ||||||          |||
Db        619 CTA------GAATTCGATACTGCCCTCATAGATCAAGCAGAGAATCTAGGAAGGTTTTTA 672

Qy         63 LeuIleThrTrpAsnIleAspGlyLeu 71
              |||  :::  :::|||:::||||||
Db        673 TTATCGTCCGAGGACATAGAGGGACTG 699

RESULT 48
US-11-136-527-3388/c
; Sequence 3388, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3388
;   LENGTH: 2995
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-3388

Alignment Scores:
Pred. No.:              718          Length:          2995
Score:                  64.00        Matches:           16
Percent Similarity:     39.3%        Conservative:       8
Best Local Similarity:  26.2%        Mismatches:        23
Query Match:            14.3%        Indels:            14
DB:                     12           Gaps:               2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-136-527-3388 (1-2995)

Qy          11 ProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThr 30
               ||||||  ||||||        |||||||||  ||||||   :::   |||
Db        2265 CCTCCTACGGAAGAGGATGTGCTTGAAAGAGCGCCTGAAGTCCTGATTGAAGACAGTATA 2206

Qy          31 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSer 50
                                  ::::::      :::   |||   :::          |||
Db        2205 -----------------GATGACCGGGTTGAGCGAACTGTTGCAGTAGCCGATCCAGAA 2164

Qy          51 GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGly 70
               |||       ::::::                                |||   :::   |||
Db        2163 GAAAAACTTAAAGAGCGG----------------------TTCTGGCAGCTGGCAGGC 2128

Qy          71 Leu 71
               |||
Db        2127 CTC 2125

RESULT 49
US-09-925-065A-750498/c
; Sequence 750498, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 750498
;   LENGTH: 648
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-750498

Alignment Scores:
Pred. No.:              138          Length:          648
Score:                  63.50        Matches:          24
Percent Similarity:     43.6%        Conservative:     10
Best Local Similarity:  30.8%        Mismatches:       21
Query Match:            14.2%        Indels:           23
DB:                     6            Gaps:              4

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-750498 (1-648)

Qy          28 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp------SerThrThrSer 45
               |||   |||:::|||   |||         |||||||||          |||
Db        586 CCTTGGACTTTTGTTTCCCTCCACACTCTAGAAACTACACCCATTTGATCCCTTAAGGTA 527

Qy          46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet--------------- 60
               |||::::::              ||||||:::      |||:::
Db        526 AAGCTAACCAGGGAAGTTTCTCTCCAGCAGAAGACGGCATCCTTGATGTGAACAGCTTTT 467

Qy          61 ---------------------PheSerLeuIleThr--------------------Trp 66
                                    ||||||  ||||||                     :::
```

```
Db        466 TCCAAGATCATGGATCAAGACTTCTCTACTATTACGAGACTGTTATCTTTGAAYATTTTT 407

Qy         67 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCys 84
              :::::|||         :::||||||   ||||||      :::||||||||||
Db        406 TCCCTGGATTATGCCTCTATAAACAATAGAAGTGAA------AAGGGGGTCTGT 359
```

RESULT 50
US-10-932-182A-1169
; Sequence 1169, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1169
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-1169

Alignment Scores:

Pred. No.:	189	Length:	849
Score:	63.50	Matches:	21
Percent Similarity:	42.2%	Conservative:	14
Best Local Similarity:	25.3%	Mismatches:	31
Query Match:	14.2%	Indels:	17
DB:	7	Gaps:	2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-1169 (1-849)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              :::|||      :::||||||   |||:::      :::             |||:::
Db        466 CTGGAACTTCCTATAAATAGTTGTTTTGACTCTTTAATCAACGCAGAGTTCTTGAAACTA 525

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
                        :::   |||
Db        526 GAAGGTAACGATTTACACGAA------------------------------------ 546

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              :::|||   |||   :::||||||||||:::         ::::::|||      :::|||
Db        547 GAATCCAATACAGGCGAAATAAGTCCAAACGCGGGGTTGGAAAAGGAAGAAGCTGCCATG 606

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn----------AsnLeu 76
                 |||   :::      :::|||      |||   |||             ||||||
Db        607 AAAGGGTTGGCTTCCGATATAGTGGATTTTTGTGATAGCAATAGCACGGATGATAACCTA 666

Qy         77 SerGluArg 79
                 ||||||
Db        667 AAAGAAAGA 675
```

RESULT 51
US-10-932-182A-1169
; Sequence 1169, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1169
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-1169

Alignment Scores:

Pred. No.:	189	Length:	849
Score:	63.50	Matches:	21
Percent Similarity:	42.2%	Conservative:	14
Best Local Similarity:	25.3%	Mismatches:	31
Query Match:	14.2%	Indels:	17
DB:	7	Gaps:	2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-1169 (1-849)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              :::|||      :::||||||   |||:::      :::             |||:::
Db        466 CTGGAACTTCCTATAAATAGTTGTTTTGACTCTTTAATCAACGCAGAGTTCTTGAAACTA 525

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
                        :::   |||
Db        526 GAAGGTAACGATTTACACGAA------------------------------------ 546

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              :::|||   |||   :::||||||||||:::         ::::::|||      :::|||
Db        547 GAATCCAATACAGGCGAAATAAGTCCAAACGCGGGGTTGGAAAAGGAAGAAGCTGCCATG 606

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn----------AsnLeu 76
                 |||   :::      :::|||      |||   |||             ||||||
Db        607 AAAGGGTTGGCTTCCGATATAGTGGATTTTTGTGATAGCAATAGCACGGATGATAACCTA 666

Qy         77 SerGluArg 79
                 ||||||
Db        667 AAAGAAAGA 675
```

RESULT 52
US-10-750-185-41248
; Sequence 41248, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 41248
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Bovine 19866881067831
US-10-750-185-41248

Alignment Scores:

Pred. No.:	421	Length:	1673
Score:	63.50	Matches:	15
Percent Similarity:	48.1%	Conservative:	11
Best Local Similarity:	27.8%	Mismatches:	27
Query Match:	14.2%	Indels:	1
DB:	8	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-41248 (1-1673)

```
Qy          6 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle 25
              :::|||          |||          ||||||      :::   :::|||:::
Db       1117 TCCTCCCTCCTCCCTCCCCATACCCCAGAGAGTCCCAAACAGGATAAGCCCAAGGCG--- 1173

Qy         26 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 45
              |||||||||:::::::::      |||::::::   |||       |||
Db       1174 AAACACCCCAAGACACATATTAATCAAATTAACAAAGATCAAACACAAAGAACAAATATT 1233

Qy         46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 59
              |||   |||      :::   |||            |||:::
Db       1234 AAAAGCAGCAAGGGAAAAACAACAAATACCACACAAGGGAAT 1275

RESULT 53
US-10-750-623-41248
; Sequence 41248, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 41248
;   LENGTH: 1673
;   TYPE: DNA
;   ORGANISM: Bovine  19866881067831
US-10-750-623-41248

Alignment Scores:
Pred. No.:                421             Length:         1673
Score:                    63.50           Matches:        15
Percent Similarity:       48.1%           Conservative:   11
Best Local Similarity:    27.8%           Mismatches:     27
Query Match:              14.2%           Indels:         1
DB:                       8               Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-623-41248 (1-1673)

Qy          6 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle 25
              :::|||          |||          ||||||      :::   :::|||:::
Db       1117 TCCTCCCTCCTCCCTCCCCATACCCCAGAGAGTCCCAAACAGGATAAGCCCAAGGCG--- 1173

Qy         26 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 45
              |||||||||:::::::::      |||::::::   |||       |||
Db       1174 AAACACCCCAAGACACATATTAATCAAATTAACAAAGATCAAACACAAAGAACAAATATT 1233

Qy         46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 59
              |||   |||      :::   |||            |||:::
Db       1234 AAAAGCAGCAAGGGAAAAACAACAAATACCACACAAGGGAAT 1275

RESULT 54
US-10-932-182A-1899
; Sequence 1899, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1899
;   LENGTH: 2895
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-1899

Alignment Scores:
Pred. No.:                803             Length:         2895
Score:                    63.50           Matches:        24
Percent Similarity:       43.0%           Conservative:   13
Best Local Similarity:    27.9%           Mismatches:     26
Query Match:              14.2%           Indels:         23
DB:                       7               Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-1899 (1-2895)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGlu---------------------- 19
              ||||||   :::|||||||   :::::::::   |||
Db       1342 TCCTACATACAACCGCCTCAAGATCAAGCTCCATTGAACGATTTCCAAGAGATTTTCGAA 1401

Qy         20 ------ArgArgProGluThrIle------SerGluProLysThrTyrValAspLeuThr 35
                    ||||||      |||          :::|||          |||||||||   |||
Db       1402 CATTACAGAAGAACAAGCACTGATTGGGGTACTGAAAATACAAAATACGTCGATAGCACC 1461

Qy         36 Asn---------------GluGluThrThrAspSerThrThrSerLysIleSerProSer 50
              |||               |||:::  |||   ::::::  |||:::   |||
Db       1462 AACTACTATTCAGACGAAGAAGATTTGACACATGCATCCATTTCCCAAGATAGTAGTTTC 1521

Qy         51 GluAspThrGlnGlnGluAsnGlySerMetPhe--------------------SerLeu 63
                 |||          |||   |||:::                          |||:::
Db       1522 TTGTCCACAGATAGCTCTAACAATAGTGTTTTGGTGAAACCAAATAATGCAGGCTCAATG 1581

Qy         64 IleThrTrpAsnIleAsp 69
              |||:::      :::|||
Db       1582 ATTTCAGAAAGATTAGAT 1599

RESULT 55
US-10-932-182A-1899
; Sequence 1899, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1899
;   LENGTH: 2895
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-1899

Alignment Scores:
Pred. No.:                803             Length:         2895
Score:                    63.50           Matches:        24
Percent Similarity:       43.0%           Conservative:   13
Best Local Similarity:    27.9%           Mismatches:     26
Query Match:              14.2%           Indels:         23
DB:                       7               Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-1899 (1-2895)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGlu---------------------- 19
              ||||||   :::|||||||   :::::::::   |||
```

```
Db       1342 TCCTACATACAACCGCCTCAAGATCAAGCTCCATTGAACGATTTCCAAGAGATTTTCGAA 1401

Qy         20 ------ArgArgProGluThrIle------SerGluProLysThrTyrValAspLeuThr 35
                  ||||||      |||          :::|||          |||||||||   |||
Db       1402 CATTACAGAAGAACAAGCACTGATTGGGGTACTGAAAATACAAAATACGTCGATAGCACC 1461

Qy         36 Asn--------------GluGluThrThrAspSerThrThrSerLysIleSerProSer 50
              |||              |||:::   |||   ::::::   |||:::   |||
Db       1462 AACTACTATTCAGACGAAGAAGATTTGACACATGCATCCATTTCCCAAGATAGTAGTTTC 1521

Qy         51 GluAspThrGlnGlnGluAsnGlySerMetPhe--------------------SerLeu 63
                 |||          |||   |||:::                        |||:::
Db       1522 TTGTCCACAGATAGCTCTAACAATAGTGTTTTGGTGAAACCAAATAATGCAGGCTCAATG 1581

Qy         64 IleThrTrpAsnIleAsp 69
              |||:::      :::|||
Db       1582 ATTTCAGAAAGATTAGAT 1599
```

RESULT 56
US-11-052-554A-605
; Sequence 605, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 605
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-605

Alignment Scores:

Pred. No.:	242	Length:	921
Score:	63.00	Matches:	18
Percent Similarity:	46.9%	Conservative:	12
Best Local Similarity:	28.1%	Mismatches:	26
Query Match:	14.1%	Indels:	8
DB:	12	Gaps:	2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-052-554A-605 (1-921)

```
Qy         15 GluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeu 34
              ||||||:::::::::   ||||||          |||   :::   :::
Db        751 GAAAGCAGTGTTGATGATCGTCCAATGAAAACTGTTTCAAAAGATACAGGTATCCCAATT 810

Qy         35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
                 :::      |||||||||          |||
Db        811 GCCGCTAAAATTTTTACAGATTCAGTTGCTAAAAAA--------------------GGA 849

Qy         55 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn 74
              |||          |||   ::::::::::::   ||||||||||||   :::      |||
Db        850 CAGGCTGGAGATAGTTACTATGCGATGATGAAGTGGAATATAGATAAAATT---GCAAAT 906

Qy         75 AsnLeuSerGlu 78
                 ||||||:::
Db        907 GGTCTGTCACAA 918
```

RESULT 57
US-11-136-527-4231
; Sequence 4231, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4231
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4231

Alignment Scores:

Pred. No.:	397	Length:	1400
Score:	63.00	Matches:	22
Percent Similarity:	49.3%	Conservative:	11
Best Local Similarity:	32.8%	Mismatches:	24
Query Match:	14.1%	Indels:	10
DB:	12	Gaps:	4

US-10-757-745-2_COPY_54_140 (1-87) x US-11-136-527-4231 (1-1400)

```
Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||   |||          :::   :::   |||||||||          ||||||
Db        315 GAATGTGCCAGTACTGGCCATGAAGACATGCCGGTGGAGAGG---ATTCTAGAA------ 365

Qy         22 ProGluThrIleSerGluProLysThr------TyrValAspLeuThrAsnGluGluThr 39
                 |||          ||||||||||||          |||   |||::::::   |||   :::
Db        366 GCTGAGCTTGCTGTGGAACCAAAGACCGAATCCTACGGCGACATGAGTGTGGAGAGCTCG 425

Qy         40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 59
              |||:::          :::   |||       :::   |||   |||
Db        426 ACAAATGACCCTGTCACCAATATATGCCACGCTGCGGATAAGCAG--------------- 470

Qy         60 MetPheSerLeuIleThrTrp 66
              :::|||:::|||:::   |||
Db        471 CTYTTCACCCTTGTTGAGTGG 491
```

RESULT 58
US-11-136-527-135
; Sequence 135, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 135
; LENGTH: 1604
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-135

Alignment Scores:

Pred. No.:	466	Length:	1604
Score:	63.00	Matches:	22
Percent Similarity:	49.3%	Conservative:	11
Best Local Similarity:	32.8%	Mismatches:	24
Query Match:	14.1%	Indels:	10
DB:	12	Gaps:	4

US-10-757-745-2_COPY_54_140 (1-87) x US-11-136-527-135 (1-1604)

```
Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
```

```
Db        519 GAATGTGCCAGTACTGGCCATGAAGACATGCCGGTGGAGAGG---ATTCTAGAA------ 569

Qy         22 ProGluThrIleSerGluProLysThr------TyrValAspLeuThrAsnGluGluThr 39

Db        570 GCTGAGCTTGCTGTGGAACCAAAGACCGAATCCTACGGCGACATGAGTGTGGAGAGCTCG 629

Qy         40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 59

Db        630 ACAAATGACCCTGTCACCAATATATGCCACGCTGCGGATAAGCAG-------------- 674

Qy         60 MetPheSerLeuIleThrTrp 66

Db        675 CTYTTCACCCTTGTTGAGTGG 695

RESULT 59
US-11-103-957-46
; Sequence 46, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
;  APPLICANT: Berthet, Francois-Xavier Jacques
;  APPLICANT:  Lobet, Yves
;  APPLICANT:  Poolman, Jan
;  APPLICANT:  Verlant, Vincent Georges Christian Louis
;  TITLE OF INVENTION: Vaccine Composition
;  FILE REFERENCE: B45261
;  CURRENT APPLICATION NUMBER: US/11/103,957
;  CURRENT FILING DATE:  2005-04-12
;  PRIOR APPLICATION NUMBER: US/10/467,534
;  PRIOR FILING DATE: 2004-02-03
;  PRIOR APPLICATION NUMBER: PCT/EP02/01356
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: GB 0103169.9
;  PRIOR FILING DATE: 2001-02-08
;  NUMBER OF SEQ ID NOS: 108
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46
;   LENGTH: 5256
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-11-103-957-46

Alignment Scores:
Pred. No.:                1.88e+03        Length:          5256
Score:                    63.00           Matches:         21
Percent Similarity:       41.5%           Conservative:    13
Best Local Similarity:    25.6%           Mismatches:      40
Query Match:              14.1%           Indels:          8
DB:                       12              Gaps:            2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-103-957-46 (1-5256)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29

Db       1915 GAACAGCCCGCAGCAGCTTCTGCCGCGACGTCTACTCCCGAGTCTGCCCCAGTGGTCTCA 1974

Qy         30 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr---------------- 44

Db       1975 ACTGCTCTAAGCACACCTTCATCTTCTACCGTCTCTTCATTAACCTTACTAGCAGCCTCT 2034

Qy         45 SerLysIleSerPro---------SerGluAspThrGlnGlnGluAsnGlySerMetPhe 61

Db       2035 TCACAAGCCTCTCCTGCAACCTCTAATAAGGAAACTCAAGATCCTAATGCTGATACAGAC 2094

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81

Db       2095 TTATTGATCGATTATGTAGTTGATACGACTATCAGCAAAAACACTGCTAAGAAAGGCGGT 2154

Qy         82 GlyVal 83

Db       2155 GGAATC 2160

RESULT 60
US-10-932-182A-81472
; Sequence 81472, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81472
;   LENGTH: 594
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-81472

Alignment Scores:
Pred. No.:                168             Length:          594
Score:                    62.50           Matches:         20
Percent Similarity:       42.9%           Conservative:    13
Best Local Similarity:    26.0%           Mismatches:      41
Query Match:              14.0%           Indels:          3
DB:                       7               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-81472 (1-594)

Qy          6 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle 25

Db        130 TCCAGCTACGGGATGCCACCCACCAACCGATCTAATTTGGATAGAATAAGAAGAAGGACT 189

Qy         26 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 45

Db        190 AGCTCACCAGTAAGAACTGATAAATTTGCATCACAAAATGTGATAGATGACCAAAGATTA 249

Qy         46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 65

Db        250 GAAATAAAATATTTAGAACGAATCGTATATGACCAAGGCACTGTCATTGATAATCTAACA 309

Qy         66 TrpAsnIleAspGlyLeuAsp---------LeuAsnAsnLeuSerGluArg 79

Db        310 TCAAGAATAACAAGACTGGAATCATTCATATTGAATTCGATCTCAGACAGA 360

RESULT 61
US-10-932-182A-81472
; Sequence 81472, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81472
;   LENGTH: 594
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-81472

Alignment Scores:
Pred. No.:                168             Length:          594
Score:                    62.50           Matches:         20
Percent Similarity:       42.9%           Conservative:    13
```

```
Best Local Similarity:  26.0%          Mismatches:    41
Query Match:            14.0%          Indels:        3
DB:                     7              Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-81472 (1-594)

Qy          6 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle 25
              :::||||||      ||||||          |||   |||:::|||
Db        130 TCCAGCTACGGGATGCCACCCACCAACCGATCTAATTTGGATAGAATAAGAAGAAGGACT 189

Qy         26 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 45
              |||   |||                 ::::::         |||
Db        190 AGCTCACCAGTAAGAACTGATAAATTTGCATCACAAAATGTGATAGATGACCAAAGATTA 249

Qy         46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 65
              :::|||         |||             :::   |||::::::         :::|||
Db        250 GAAATAAAATATTTAGAACGAATCGTATATGACCAAGGCACTGTCATTGATAATCTAACA 309

Qy         66 TrpAsnIleAspGlyLeuAsp---------LeuAsnAsnLeuSerGluArg 79
                 |||      |||:::          ||||||::::::|||:::|||
Db        310 TCAAGAATAACAAGACTGGAATCATTCATATTGAATTCGATCTCAGACAGA 360

RESULT 62
US-09-925-065A-180175
; Sequence 180175, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 180175
;   LENGTH: 624
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-180175

Alignment Scores:
Pred. No.:              178            Length:        624
Score:                  62.50          Matches:       21
Percent Similarity:     59.5%          Conservative:  4
Best Local Similarity:  50.0%          Mismatches:    8
Query Match:            14.0%          Indels:        9
DB:                     6              Gaps:          3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-180175 (1-624)

Qy         46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet--------------- 60
              ||||||||||||   |||:::||||||:::   |||      |||
Db        176 AAGATAAGCCCA---GAAAACACCCAAAAAACAAATAACAAAATGGCTGGAATATGTCCT 232

Qy         61 ---PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                 ||||||||||||   ||||||:::|||      ||||||   |||   :::|||
Db        233 TACTTATCACTAATAACATTAAATATAAATGGA------CTAAACTGTTTAATCAAAAGA 286

Qy         80 AlaArg 81
              |||
Db        287 CATAGA 292

RESULT 63
US-09-925-065A-180178
; Sequence 180178, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 180178
;   LENGTH: 624
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-180178

Alignment Scores:
Pred. No.:              178            Length:        624
Score:                  62.50          Matches:       21
Percent Similarity:     59.5%          Conservative:  4
Best Local Similarity:  50.0%          Mismatches:    8
Query Match:            14.0%          Indels:        9
DB:                     6              Gaps:          3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-180178 (1-624)

Qy         46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet--------------- 60
              ||||||||||||   |||:::||||||:::   |||      |||
Db        176 AAGATAAGCCCA---GAAAACACCCAAAAAACAAATAACAAAATGGCTGGAATATGTCCT 232

Qy         61 ---PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                 ||||||||||||   ||||||:::|||      ||||||   |||   :::|||
Db        233 TACTTATCACTAATAACATTAAATATAAATGGA------CTAAACTGTTTAATCAAAAGA 286

Qy         80 AlaArg 81
              |||
Db        287 CATAGA 292

RESULT 64
US-10-932-182A-3132
; Sequence 3132, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3132
;   LENGTH: 2526
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-3132
```

```
Alignment Scores:
Pred. No.:                924             Length:          2526
Score:                    62.50           Matches:         14
Percent Similarity:       48.0%           Conservative:    10
Best Local Similarity:    28.0%           Mismatches:      21
Query Match:              14.0%           Indels:          5
DB:                       7               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-3132 (1-2526)

Qy          10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
               ||||||      |||         |||::::::   |||   :::      :::   |||
Db        1900 GAGCCTGACACCGAAACTGGCAAACTTGATAAGCTGCCTGCGAGCACAGAAGACGGTAAG 1959

Qy          30 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 49
               |||   :::::::::     ::::::                  |||   |||:::
Db        1960 ACGGTAATTGAAATAGCCAAAAAGAAA--------------GATAGCGCCATTAACGAG 2004

Qy          50 SerGluAspThrGlnGlnGluAsnGlySer 59
                  |||       ||||||   ||||||
Db        2005 GAGCATGATAACCCTCAAGAAAAGGGTTCA 2034

RESULT 65
US-10-932-182A-3132
; Sequence 3132, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3132
;   LENGTH: 2526
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-3132

Alignment Scores:
Pred. No.:                924             Length:          2526
Score:                    62.50           Matches:         14
Percent Similarity:       48.0%           Conservative:    10
Best Local Similarity:    28.0%           Mismatches:      21
Query Match:              14.0%           Indels:          5
DB:                       7               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-932-182A-3132 (1-2526)

Qy          10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
               ||||||      |||         |||::::::   |||   :::      :::   |||
Db        1900 GAGCCTGACACCGAAACTGGCAAACTTGATAAGCTGCCTGCGAGCACAGAAGACGGTAAG 1959

Qy          30 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 49
               |||   :::::::::     ::::::                  |||   |||:::
Db        1960 ACGGTAATTGAAATAGCCAAAAAGAAA--------------GATAGCGCCATTAACGAG 2004

Qy          50 SerGluAspThrGlnGlnGluAsnGlySer 59
                  |||       ||||||   ||||||
Db        2005 GAGCATGATAACCCTCAAGAAAAGGGTTCA 2034

RESULT 66
US-10-330-773-251/c
; Sequence 251, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 251
;   LENGTH: 45764
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(45764)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-251

Alignment Scores:
Pred. No.:                2.8e+04         Length:          45764
Score:                    62.50           Matches:         16
Percent Similarity:       43.2%           Conservative:    16
Best Local Similarity:    21.6%           Mismatches:      25
Query Match:              14.0%           Indels:          17
DB:                       7               Gaps:            2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-330-773-251 (1-45764)

Qy          14 GluGluSerAlaLeuGluArgArgPro---------------------------------- 22
               |||     :::|||          ||||||
Db        44997 GAGGCAGGATCTCTGAGTTCAAGGCCAGGCTGGTGTACAGATTGAGTTCCAGGTCAGCCA 44938

Qy          23 -----------GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 38
                          ||||||:::|||:::   :::                  |||:::::::
Db        44937 GGGTTACACATAGAAACCCTGTCTCAAAAACAACAACAACAGCAACAACAAAACAAACAA 44878

Qy          39 ThrThr------AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 56
               |||           :::|||   :::|||   :::|||::::::   |||::::::::::
Db        44877 ACAAACAAAAAACCCACCACCAACAACAAAACAAATCCAAACAAACAAACTAAAGAACAG 44818

Qy          57 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGly 70
                           |||   :::   |||  :::   |||
Db        44817 CAGCAACAAAAAGCGTCCTTTCTCATGTGGACTTTGTTTGGC 44776

RESULT 67
US-10-995-561-13303/c
; Sequence 13303, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13303
;   LENGTH: 201990
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(201990)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13303

Alignment Scores:
Pred. No.:                1.6e+05         Length:          201990
Score:                    62.50           Matches:         19
Percent Similarity:       43.5%           Conservative:    11
Best Local Similarity:    27.5%           Mismatches:      26
Query Match:              14.0%           Indels:          13
DB:                       8               Gaps:            1
```

```
US-10-757-745-2_COPY_54_140 (1-87) x US-10-995-561-13303 (1-201990)

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr--- 40
              |||      :::|||   |||                 ||||||::::::   :::|||
Db      59123 CCATCACAACTATCCAGGCCAGATCTCAAGTACAGGCTCACCGATCAATGATCGACTGTC 59064

Qy         41 ----------------------------------AspSerThrThrSerLysIleSer 48
                                                |||     |||      ::::::
Db      59063 TTCCTCTCCACTGCTCCCTGTCTCATCATTCTGAGGGATTTCCACACCCACACAGTTAAT 59004

Qy         49 ProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle 68
              ||||||     |||:::   |||      |||      ||||||:::|||
Db      59003 CCATCCCATAGCACACGTTTGGAGCAGCTCAGTGGCTTATCCTTGGTCACTATGCGCAGA 58944

Qy         69 AspGlyLeuAspLeuAsnAsnLeuSer 77
              ||||||:::|||   ::::::
Db      58943 TACAGATTAGATGTCAACAAGATGAAC 58917

RESULT 68
US-10-330-773-261/c
; Sequence 261, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 261
;   LENGTH: 215248
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(215248)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-261

Alignment Scores:
Pred. No.:              1.73e+05        Length:         215248
Score:                  62.50           Matches:        16
Percent Similarity:     43.2%           Conservative:   16
Best Local Similarity:  21.6%           Mismatches:     25
Query Match:            14.0%           Indels:         17
DB:                     7               Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-330-773-261 (1-215248)

Qy         14 GluGluSerAlaLeuGluArgArgPro-------------------------------- 22
              |||      :::|||      ||||||
Db       1463 GAGGCAGGATCTCTGAGTTCAAGGCCAGGCTGGTGTACAGATTGAGTTCCAGGTCAGCCA 1404

Qy         23 -----------GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 38
                        ||||||:::|||:::   :::                  |||:::::
Db       1403 GGGTTACACATAGAAACCCTGTCTCAAAAACAACAACAACAGCAACAACAAAACAAACAA 1344

Qy         39 ThrThr------AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 56
              |||           :::|||   :::|||   :::|||::::::   |||:::::::::
Db       1343 ACAAACAAAAAACCCACCACCAACAACAAAACAAATCCAAACAAACAAACTAAAGAACAG 1284

Qy         57 AsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGly 70
                        |||   :::   |||   :::   |||
Db       1283 CAGCAACAAAAAGCGTCCTTTCTCATGTGGACTTTGTTTGGC 1242

RESULT 69
US-10-995-561-43975
; Sequence 43975, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43975
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-43975

Alignment Scores:
Pred. No.:              54.5            Length:         201
Score:                  62.00           Matches:        10
Percent Similarity:     63.0%           Conservative:   7
Best Local Similarity:  37.0%           Mismatches:     10
Query Match:            13.8%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-995-561-43975 (1-201)

Qy         35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
              ||||||::::::|||         |||         :::||||||:::|||   |||
Db         69 ACAAACAAACAAACAAACAAACAAACAAACAAAAACTTGTCTCCAAATGAGCAGACACCA 128

Qy         55 GlnGluAsnGlySerMetPhe 61
              :::   :::|||   :::|||
Db        129 GAATCTCATGGAGGGATTTTT 149

RESULT 70
US-10-995-561-43976
; Sequence 43976, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43976
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-43976

Alignment Scores:
Pred. No.:              54.5            Length:         201
Score:                  62.00           Matches:        10
Percent Similarity:     63.0%           Conservative:   7
Best Local Similarity:  37.0%           Mismatches:     10
Query Match:            13.8%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-995-561-43976 (1-201)

Qy         35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
              ||||||::::::|||         |||         :::||||||:::|||   |||
Db         70 ACAAACAAACAAACAAACAAACAAACAAACAAAAACTTGTCTCCAAATGAGCAGACACCA 129

Qy         55 GlnGluAsnGlySerMetPhe 61
              :::   :::|||   :::|||
Db        130 GAATCTCATGGAGGGATTTTT 150

RESULT 71
US-10-995-561-43977
```

```
; Sequence 43977, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43977
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-43977

Alignment Scores:
Pred. No.:                54.5          Length:          201
Score:                    62.00         Matches:         10
Percent Similarity:       63.0%         Conservative:    7
Best Local Similarity:    37.0%         Mismatches:      10
Query Match:              13.8%         Indels:          0
DB:                       8             Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-995-561-43977 (1-201)

Qy          35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
               ||||||:::::|||        |||         :::||||||:::|||   |||
Db          71 ACAAACAAACAAACAAACAAACAAACAAACAAAAACTTGTCTCCAAATGAGCAGACACCA 130

Qy          55 GlnGluAsnGlySerMetPhe 61
               :::   :::|||   :::|||
Db         131 GAATCTCATGGAGGGATTTTT 151

RESULT 72
US-10-995-561-43994
; Sequence 43994, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43994
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-43994

Alignment Scores:
Pred. No.:                54.5          Length:          201
Score:                    62.00         Matches:         10
Percent Similarity:       63.0%         Conservative:    7
Best Local Similarity:    37.0%         Mismatches:      10
Query Match:              13.8%         Indels:          0
DB:                       8             Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-995-561-43994 (1-201)

Qy          35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
               ||||||:::::|||        |||         :::||||||:::|||   |||
Db          72 ACAAACAAACAAACAAACAAACAAACAAACAAAAACTTGTCTCCAAATGAGCAGACACCA 131

Qy          55 GlnGluAsnGlySerMetPhe 61
               :::   :::|||   :::|||
Db         132 GAATCTCATGGAGGGATTTTT 152

RESULT 73
US-11-124-367A-29340
; Sequence 29340, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  Hongjin Huang
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Fibrosis Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001519.ORD
;  CURRENT APPLICATION NUMBER: US/11/124,367A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,846
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/582,609
;  PRIOR FILING DATE: 2004-06-25
;  PRIOR APPLICATION NUMBER: US 60/599,554
;  PRIOR FILING DATE: 2004-08-09
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29340
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-367A-29340

Alignment Scores:
Pred. No.:                54.5          Length:          201
Score:                    62.00         Matches:         10
Percent Similarity:       63.0%         Conservative:    7
Best Local Similarity:    37.0%         Mismatches:      10
Query Match:              13.8%         Indels:          0
DB:                       12            Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-11-124-367A-29340 (1-201)

Qy          35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
               ||||||:::::|||        |||         :::||||||:::|||   |||
Db          69 ACAAACAAACAAACAAACAAACAAACAAACAAAAACTTGTCTCCAAATGAGCAGACACCA 128

Qy          55 GlnGluAsnGlySerMetPhe 61
               :::   :::|||   :::|||
Db         129 GAATCTCATGGAGGGATTTTT 149

RESULT 74
US-11-124-367A-29341
; Sequence 29341, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  Hongjin Huang
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Fibrosis Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001519.ORD
;  CURRENT APPLICATION NUMBER: US/11/124,367A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,846
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/582,609
;  PRIOR FILING DATE: 2004-06-25
;  PRIOR APPLICATION NUMBER: US 60/599,554
;  PRIOR FILING DATE: 2004-08-09
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29341
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-367A-29341

Alignment Scores:
Pred. No.:                54.5          Length:          201
```

```
Score:                 62.00          Matches:       10
Percent Similarity:    63.0%          Conservative:  7
Best Local Similarity: 37.0%          Mismatches:    10
Query Match:           13.8%          Indels:        0
DB:                    12             Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x US-11-124-367A-29341 (1-201)

Qy         35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
              ||||||::::::|||          |||         :::||||||:::|||   |||
Db         70 ACAAACAAACAAACAAACAAACAAACAAACAAAAACTTGTCTCCAAATGAGCAGACACCA 129

Qy         55 GlnGluAsnGlySerMetPhe 61
              :::   :::|||   :::|||
Db        130 GAATCTCATGGAGGGATTTTT 150

RESULT 75
US-11-124-367A-29342
; Sequence 29342, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
;  APPLICANT: Michele Cargill
;  APPLICANT:  Hongjin Huang
;  TITLE OF INVENTION: Genetic Polymorphisms Associated with
;  TITLE OF INVENTION:  Fibrosis Methods of Detection and Uses Thereof
;  FILE REFERENCE: CL001519.ORD
;  CURRENT APPLICATION NUMBER: US/11/124,367A
;  CURRENT FILING DATE:  2005-05-09
;  PRIOR APPLICATION NUMBER: US 60/568,846
;  PRIOR FILING DATE: 2004-05-07
;  PRIOR APPLICATION NUMBER: US 60/582,609
;  PRIOR FILING DATE: 2004-06-25
;  PRIOR APPLICATION NUMBER: US 60/599,554
;  PRIOR FILING DATE: 2004-08-09
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29342
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-367A-29342

Alignment Scores:
Pred. No.:             54.5           Length:        201
Score:                 62.00          Matches:       10
Percent Similarity:    63.0%          Conservative:  7
Best Local Similarity: 37.0%          Mismatches:    10
Query Match:           13.8%          Indels:        0
DB:                    12             Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x US-11-124-367A-29342 (1-201)

Qy         35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
              ||||||::::::|||          |||         :::||||||:::|||   |||
Db         71 ACAAACAAACAAACAAACAAACAAACAAACAAAAACTTGTCTCCAAATGAGCAGACACCA 130

Qy         55 GlnGluAsnGlySerMetPhe 61
              :::   :::|||   :::|||
Db        131 GAATCTCATGGAGGGATTTTT 151


Search completed: March 11, 2006, 13:51:17
Job time : 338.78 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:38:30 ; Search time 336.652 Seconds
(without alignments)
1722.336 Million cell updates/sec

Title: US-10-757-745-2_COPY_54_140
Perfect score: 448
Sequence: 1 MERALNSYFEPPVEESALER.....IDGLDLNNLSERARGVCSYL 87

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170325_3993/app_query.fasta_1
-DB=N_Geneseq -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1589_@runat_09032006_170325_3993 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	448	100.0	483	2	AAX40590	Aax40590 Human sec
2	448	100.0	1296	3	AAC98160	Aac98160 Human col
3	448	100.0	1920	3	AAZ47118	Aaz47118 Human CD4
4	448	100.0	1936	10	ADD19013	Add19013 Human dis
5	448	100.0	1936	13	ADP25361	Adp25361 PRO polyp
6	448	100.0	1940	14	ADX06356	Adx06356 Cyclin-de
7	448	100.0	1948	4	AAI58997	Aai58997 Human pol
8	448	100.0	1948	5	ADQ99219	Adq99219 DNA encod
9	448	100.0	1948	9	ADB48979	Adb48979 Novel hum
10	448	100.0	1958	4	AAI60783	Aai60783 Human pol
11	448	100.0	2499	2	AAX28153	Aax28153 Topoisome
12	448	100.0	3152	9	ADA10970	Ada10970 Human cDN
13	439	98.0	752	4	AAH08073	Aah08073 Human cDN
14	439	98.0	1898	4	AAH15146	Aah15146 Human cDN
c 15	414	92.4	391	11	ADT95307	Adt95307 Colon can
c 16	414	92.4	391	11	ADX41789	Adx41789 Human cDN
17	332.5	74.2	1227	5	AAS86255	Aas86255 DNA encod
18	211.5	47.2	1312	3	AAZ47119	Aaz47119 Mouse CD4
19	178.5	39.8	602	10	ADD34376	Add34376 Mouse mit
20	114	25.4	725	6	ABQ34465	Abq34465 Oligonucl
c 21	114	25.4	725	6	ABQ34464	Abq34464 Oligonucl
22	113	25.2	258	2	AAX41008	Aax41008 Human sec
23	101	22.5	60	6	ABN41860	Abn41860 Human spl
24	84.5	18.9	1589	5	AAS87500	Aas87500 DNA encod
25	84.5	18.9	2228	5	AAS87499	Aas87499 DNA encod
c 26	84.5	18.9	2717	5	AAS82908	Aas82908 DNA encod
27	84.5	18.9	4002	5	AAS75995	Aas75995 DNA encod
c 28	84.5	18.9	4890	5	AAS87489	Aas87489 DNA encod
29	84.5	18.9	6448	5	AAS82911	Aas82911 DNA encod
30	84.5	18.9	7197	5	AAS87485	Aas87485 DNA encod
c 31	84.5	18.9	7556	5	AAS82893	Aas82893 DNA encod
c 32	84.5	18.9	7731	5	AAS87486	Aas87486 DNA encod
33	84.5	18.9	11742	5	AAS87514	Aas87514 DNA encod
34	84.5	18.9	12036	5	AAS75996	Aas75996 DNA encod
35	84.5	18.9	23157	5	AAS75998	Aas75998 DNA encod
36	84.5	18.9	32313	5	AAS82912	Aas82912 DNA encod
37	83	18.5	6115	10	ADF90711	Adf90711 Human hep
38	83	18.5	37741	14	AEB47090	Aeb47090 Chimpanze
c 39	81.5	18.2	826	10	ADD34445	Add34445 Mouse mit
40	81.5	18.2	3321	3	AAC81290	Aac81290 Mouse apo
41	81.5	18.2	6763	3	AAC81309	Aac81309 Mouse FLA
c 42	79	17.6	1473	6	AAD41820	Aad41820 Human tra
c 43	79	17.6	1703	8	ABZ80231	Abz80231 Human tra
c 44	79	17.6	2456	13	ADR07759	Adr07759 Full leng
c 45	79	17.6	3558	12	ADQ67469	Adq67469 Novel hum
46	78	17.4	480	6	ABQ55431	Abq55431 Human ova
47	78	17.4	1142	9	ADA24517	Ada24517 Human cDN
48	78	17.4	3672	10	ABZ77286	Abz77286 Nucleotid
49	78	17.4	3672	13	ADR24529	Adr24529 Breast ca
50	78	17.4	3672	14	ADY61760	Ady61760 Human gen
51	78	17.4	4180	13	ACN37824	Acn37824 Tumour-as
52	78	17.4	4194	10	AAL60051	Aal60051 Human Pc0
53	78	17.4	4750	2	AAZ41297	Aaz41297 Human nor
54	78	17.4	5511	8	ABX77514	Abx77514 Different
55	78	17.4	5511	8	ABX63285	Abx63285 Human cDN
56	78	17.4	5511	9	ACH04033	Ach04033 Human cDN
57	78	17.4	5511	10	ADJ56346	Adj56346 Human cDN
c 58	78	17.4	6115	10	ADF90711	Adf90711 Human hep
59	78	17.4	6316	12	ADP07274	Adp07274 Human IGF
60	78	17.4	6316	12	ADP07323	Adp07323 Human IGF
61	77	17.2	894	6	ABQ34415	Abq34415 Oligonucl
c 62	77	17.2	894	6	ABQ34414	Abq34414 Oligonucl
63	75.5	16.9	10506	8	AAD52973	Aad52973 Bacteriop
64	75.5	16.9	10506	12	ADP09724	Adp09724 Bacteriop
65	75.5	16.9	10506	13	ADQ96719	Adq96719 Bacteriop
66	75.5	16.9	10506	13	ADT91343	Adt91343 N4 virion
67	75.5	16.9	10617	8	AAD52982	Aad52982 Bacteriop
68	75.5	16.9	10617	12	ADP09737	Adp09737 Bacteriop
69	75.5	16.9	10617	13	ADQ96732	Adq96732 Bacteriop
70	75.5	16.9	10617	13	ADT91356	Adt91356 N4 virion
71	75	16.7	2601	12	ADQ08623	Adq08623 Ciona int
c 72	74	16.5	2148	4	AAF57111	Aaf57111 A. thalia
c 73	74	16.5	2490	5	AAS73107	Aas73107 DNA encod
c 74	74	16.5	2547	4	AAF58469	Aaf58469 AtCNGC1/D
c 75	74	16.5	5996	4	AAF58468	Aaf58468 AtCNGC1/D

ALIGNMENTS

```
RESULT 1
AAX40590
ID   AAX40590 standard; cDNA; 483 BP.
XX
AC   AAX40590;
XX
DT   18-JUN-1999  (first entry)
XX
DE   Human secreted protein 5' EST SEQ ID No: 190.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS   Homo sapiens.
XX
PN   WO9906550-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB001232.
XX
PR   01-AUG-1997;   97US-00905144.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Lacroix B;
XX
DR   WPI; 1999-153780/13.
DR   P-PSDB; AAY11868.
XX
PT   New isolated prostate-derived nucleic acids - used to develop products
PT   which may have cytokine, immune regulatory, haematopoiesis regulating,
PT   anti-inflammatory or tumour inhibition activity.
XX
PS   Claim 1; Page 298; 675pp; English.
XX
CC   AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
CC   human secreted proteins expressed in prostate, and encode the proteins
CC   given in AAY11716 to AAY11993 respectively. The proteins given represent
CC   the signal peptide and an N-terminal fragment of a secreted protein. The
CC   nucleic acid sequences can be used for producing secreted human gene
CC   products. They can also be used to develop products for diagnosis and
CC   therapy. The proteins obtained may have cytokine activity, cell
CC   proliferation and differentiation activity, haematopoiesis regulating
CC   activity, tissue growth regulating activity, reproductive hormone
CC   regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC   thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, tumour inhibition activity or other activities. The products
CC   can be used in forensic, gene therapy and chromosome mapping procedures.
CC   The sequences can also be used for obtaining corresponding promoter
CC   sequences. The nucleic acids encoding the signal peptides can be used for
CC   directing extracellular secretion of a polypeptide or the insertion of a
CC   polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ   Sequence 483 BP; 123 A; 111 C; 139 G; 110 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.22e-51        Length:          483
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       2               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x AAX40590 (1-483)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
                |||||||||||||||||||||
Db        400 CGAGGGGTGTGTTCCTACTTA 420

RESULT 2
AAC98160
ID   AAC98160 standard; cDNA; 1296 BP.
XX
AC   AAC98160;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human colon cancer antigen nucleotide sequence SEQ ID NO:170.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW   immunomodulatory; muscular; gynaecological; gastrointestinal;
KW   nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW   neural disorder; immune system disorder; muscular disorder;
KW   reproductive disorder; gastrointestinal disorder; renal disorder;
KW   infectious disease; cardiovascular disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055351-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005883.
XX
PR   12-MAR-1999;   99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587534/55.
DR   P-PSDB; AAB53403.
XX
PT   Colon cancer associated gene sequences, referred to as colon cancer
PT   antigens, useful for the treatment, prevention, and diagnosis of colon
PT   disorders such as colon cancer.
XX
PS   Claim 1; Page 597; 2104pp; English.
XX
CC   AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC   called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC   human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC   neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC   vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC   can be used in gene therapy. The colon cancer antigen polynucleotides,
CC   proteins and antibodies to the proteins are useful for the prevention,
CC   treatment and diagnosis of colon disorders, such as colon cancer. The
CC   polynucleotides may be used in diagnostics and research, such as for
CC   chromosome identification, and as hybridisation probes. The proteins may
CC   also be used to prevent diseases such as neural disorders, immune system
CC   disorders, muscular disorders, reproductive disorders, gastrointestinal
CC   disorders, wounds, renal disorders, infectious diseases, and
```

```
CC   cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC   sequences used in the exemplification of the present invention
XX
SQ   Sequence 1296 BP; 376 A; 250 C; 326 G; 333 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.:              1.57e-50      Length:         1296
Score:                  448.00        Matches:        87
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AAC98160 (1-1296)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 242

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 302

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 362

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 422

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        423 CGAGGGGTGTGTTCCTACTTA 443

RESULT 3
AAZ47118
ID   AAZ47118 standard; cDNA; 1920 BP.
XX
AC   AAZ47118;
XX
DT   15-MAR-2000  (first entry)
XX
DE   Human CD40 receptor associated protein gene.
XX
KW   Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW   immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW   human; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW   tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW   TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW   diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW   arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW   graft versus host disease; autoimmune disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO9955859-A2.
XX
PD   04-NOV-1999.
XX
PF   28-APR-1999;   99WO-EP003025.
XX
PR   29-APR-1998;   98EP-00201392.
XX
PA   (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI   Pype SMC,  Remacle JEFJG,  Huylebroeck DFE;
XX
DR   WPI; 2000-062029/05.
DR   P-PSDB; AAY56019.
XX
PT   Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT   diseases and for improvement of anti-tumor treatments.
XX
PS   Claim 9; Page 37-39; 48pp; English.
XX
CC   This sequence represents the gene encoding human CD40 receptor associated
CC   protein (CRAP). CRAP is a functional protein capable of interacting with
CC   the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC   necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC   I, where the protein has no homology to TNF receptor associated factor
CC   (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC   the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC   related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC   diseases, and for the improvement of anti-tumour diseases. Diseases which
CC   may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC   systemic lupus erythematosus, graft rejection, graft versus host disease,
CC   allergy, and autoimmune disease. The proteins can be used to sensitize
CC   tumour cells to anti-tumour treatments and to screen for compounds which
CC   interfere with the interaction of the proteins with other protein
CC   components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ   Sequence 1920 BP; 599 A; 327 C; 435 G; 557 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              2.66e-50      Length:         1920
Score:                  448.00        Matches:        87
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AAZ47118 (1-1920)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 238

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 298

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 358

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 418

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        419 CGAGGGGTGTGTTCCTACTTA 439

RESULT 4
ADD19013
ID   ADD19013 standard; DNA; 1936 BP.
XX
AC   ADD19013;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human disease related protein DNA sequence SeqID502.
XX
KW   human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW   antiarteriosclerotic; vulnerary; gene therapy;
KW   hypoxia-regulated condition; tumourigenesis; angiogenesis; apoptosis;
KW   inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW   glucose transportation; catecholamine synthesis; iron transport;
KW   nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW   retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW   inflammatory condition; wound healing; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003018621-A2.
```

XX
PD 06-MAR-2003.
XX
PF 23-AUG-2002; 2002WO-GB003892.
XX
PR 23-AUG-2001; 2001GB-00020558.
PR 05-OCT-2001; 2001GB-00024037.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
DR WPI; 2003-290046/28.
DR P-PSDB; ADD19012.
XX
PT New substantially purified polypeptide, useful for diagnosing or treating
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT wound healing.
XX
PS Claim 27; SEQ ID NO 502; 424pp; English.
XX
CC This invention relates to novel human genes and gene product which are
CC implicated in certain disease states. Compounds which modulate the
CC proteins of the invention may have cytostatic, antiinflammatory,
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The
CC sequences of the invention may be useful for gene therapy. The invention
CC may be useful for diagnosing or treating a hypoxia-regulated condition,
CC such as tumourigenesis, angiogenesis, apoptosis, inflammation,
CC erythropoiesis, or the biological response to hypoxia conditions
CC including processes such as glycolysis, gluconeogenesis, glucose
CC transportation, catecholamine synthesis, iron transport or nitric oxide
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC inflammatory conditions or wound healing. The present sequence is that of
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.69e-50	Length:	1936
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x ADD19013 (1-1936)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 235

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 295

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 355

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 415

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        416 CGAGGGGTGTGTTCCTACTTA 436
```

RESULT 5
ADP25361
ID ADP25361 standard; cDNA; 1936 BP.
XX
AC ADP25361;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide encoding cDNA SEQ ID NO:475.
XX
KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-419628/39.
DR P-PSDB; ADP25362.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 1; SEQ ID NO 475; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence encodes a PRO protein of
CC the invention.
XX
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.69e-50	Length:	1936
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	13	Gaps:	0

```
US-10-757-745-2_COPY_54_140 (1-87) x ADP25361 (1-1936)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 235

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 295

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 355

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 415

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        416 CGAGGGGTGTGTTCCTACTTA 436

RESULT 6
ADX06356
ID   ADX06356 standard; DNA; 1940 BP.
XX
AC   ADX06356;
XX
DT   21-APR-2005  (first entry)
XX
DE   Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 921.
XX
KW   cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2005012875-A2.
XX
PD   10-FEB-2005.
XX
PF   29-JUL-2004; 2004WO-US024424.
XX
PR   29-JUL-2003; 2003US-0490890P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Li M,  Rupnow BA,  Webster KR,  Jackson DG,  Wong TW;
XX
DR   WPI; 2005-163068/17.
DR   P-PSDB; ADX06357.
XX
PT   Biomarkers useful for predicting or determining the response of a mammal
PT   to a cancer treatment comprising administration of a modulator of cyclin-
PT   dependent kinase activity.
XX
PS   Claim 5; SEQ ID NO 921; 141pp; English.
XX
CC   This invention describes a novel method of predicting or determining
CC   whether a mammal will respond or is responding to an anti-cancer agent
CC   that modulates cyclin-dependent kinase (cdk) activity. The method
CC   comprises measuring the level of one or more biomarkers selected from
CC   2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC   NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC   invention is utilized in a kit for determining or predicting whether
CC   patient would be susceptible or resistant to treatment by an agent
CC   modulating cdk activity. The invention also describes a method for
CC   utilizing individualized genetic profiles for treating diseases and
CC   disorders based on patient's response and molecular level, specialized
CC   microarrays comprising the biomarkers described, antibodies directed
CC   against the biomarkers and a cell culture model to identify biomarkers.
CC   The cdk modulator is preferably N-5-[[5-(1,1-Dimethylethyl)-2-
CC   oxazolyl]methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC   tartaric acid salt. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC   sequence encodes a biomarker used in the method of the invention.
XX
SQ   Sequence 1940 BP; 613 A; 330 C; 439 G; 558 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.69e-50      Length:          1940
Score:                  448.00        Matches:         87
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     14            Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x ADX06356 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 7
AAI58997
ID   AAI58997 standard; cDNA; 1948 BP.
XX
AC   AAI58997;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 1200.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
PA   (HYSE-) HYSEQ INC.
```

XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM39841.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 1200; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.71e-50	Length:	1948
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x AAI58997 (1-1948)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        438 CGAGGGGTGTGTTCCTACTTA 458
```

RESULT 8
ADQ99219
ID ADQ99219 standard; cDNA; 1948 BP.
XX
AC ADQ99219;
XX
DT 23-SEP-2004 (first entry)
XX
DE DNA encoding human GPCR-like protein seqid 889.
XX
KW ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
KW antidiabetic; GPCR-like protein; ophthalmic disorder;
KW neurological disorder; immunological disorder; nephritic disorder;
KW hormonal dysfunction; cancer; atherosclerosis; diabetes;
KW molecular weight marker; food supplement; human; ss.
XX
OS Homo sapiens.
XX
PN US6569662-B1.
XX
PD 27-MAY-2003.
XX
PF 19-JUL-2000; 2000US-00620312.
XX
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Drmanac RT;
XX
DR WPI; 2001-442255/47.
XX
PT New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT useful for treating diseases of ophthalmic, neurological, immunological
PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT and diabetes.
XX
PS Example 2; SEQ ID NO 889; 92pp; English.
XX
CC The invention describes an isolated polynucleotide (I) comprising a fully
CC defined (SI) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC given in the specification, its translated or protein coding portion, its
CC extracellular portion or its active domain. The GPCR-like polypeptides
CC and polynucleotides are useful for the treatment of diseases of
CC ophthalmic, neurological, immunological and nephritic systems. They may
CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC diabetes. The antibodies are useful for detecting or quantitating the
CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement. This sequence represents a human
CC polynucleotide of the invention.
XX
SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.71e-50	Length:	1948
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x ADQ99219 (1-1948)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        438 CGAGGGGTGTGTTCCTACTTA 458
```

```
RESULT 9
ADB48979
ID   ADB48979 standard; cDNA; 1948 BP.
XX
AC   ADB48979;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Novel human cDNA SEQ ID NO 889.
XX
KW   ss; cancer; neurodegenerative disease; human.
XX
OS   Homo sapiens.
XX
PN   US2003104529-A1.
XX
PD   05-JUN-2003.
XX
PF   04-JAN-2002; 2002US-00037270.
XX
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   19-JUL-2000; 2000US-00620312.
XX
PA   (ZHOU/) ZHOU P.
PA   (TANG/) TANG Y T.
PA   (LIUC/) LIU C.
PA   (ASUN/) ASUNDI V.
PA   (DRMA/) DRMANAC R T.
XX
PI   Zhou P,  Tang YT,  Liu C,  Asundi V,  Drmanac RT;
XX
DR   WPI; 2003-678194/64.
XX
PT   New polynucleotide, useful for treating diseases e.g., cancer or
PT   neurodegenerative diseases.
XX
PS   Claim 1; SEQ ID NO 889; 99pp; English.
XX
CC   The invention relates to a polynucleotide comprising a sequence given in
CC   the specification, or its mature protein-coding portion, or its
CC   complement. The polynucleotide is useful for treating diseases e.g.,
CC   cancer or neurodegenerative diseases and many others listed in the
CC   specification. The present sequence represents a novel human cDNA. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=20030104529.
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.71e-50        Length:         1948
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x ADB48979 (1-1948)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 10
AAI60783
ID   AAI60783 standard; cDNA; 1958 BP.
XX
AC   AAI60783;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4772.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J,  Zhao QA;
PI   Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41627.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders such
PT   as central nervous system injuries.
XX
PS   Claim 1; SEQ ID NO 4772; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC   encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders. Note: The sequence data for this patent did not form
CC   part of the printed specification
XX
SQ   Sequence 1958 BP; 613 A; 339 C; 442 G; 564 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:              2.73e-50        Length:         1958
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     4               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AAI60783 (1-1958)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 269

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 329

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 389

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 449

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        450 CGAGGGGTGTGTTCCTACTTA 470

RESULT 11
AAX28153
ID   AAX28153 standard; DNA; 2499 BP.
XX
AC   AAX28153;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Topoisomerase II binding protein 1 coding sequence.
XX
KW   Topoisomerase II binding protein; TopBP; anticancer agent; ds.
XX
OS   Homo sapiens.
XX
PN   JP11075856-A.
XX
PD   23-MAR-1999.
XX
PF   17-SEP-1997;   97JP-00251544.
XX
PR   17-SEP-1997;   97JP-00251544.
XX
PA   (TSUR/) TSURUO T.
PA   (CHUS ) CHUGAI PHARM CO LTD.
XX
DR   WPI; 1999-257704/22.
DR   P-PSDB; AAY03182.
XX
PT   New Topoisomerase II- binding protein - useful as an anticancer agent.
XX
PS   Disclosure; Page 18-19; 28pp; Japanese.
XX
CC   This sequence encodes the topoisomerase II binding protein (TopBP) of the
CC   invention. The TopBP protein is useful as an anticancer agent. TopBP can
CC   be used as the target molecule for anticancer agent
XX
SQ   Sequence 2499 BP; 720 A; 505 C; 587 G; 687 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.77e-50        Length:         2499
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     2               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AAX28153 (1-2499)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 700

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 760

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 820

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 880

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        881 CGAGGGGTGTGTTCCTACTTA 901

RESULT 12
ADA10970
ID   ADA10970 standard; cDNA; 3152 BP.
XX
AC   ADA10970;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human cDNA differentially expressed in colon cancer #60.
XX
KW   ss; differential expression; colon cancer; cancer; human.
XX
OS   Homo sapiens.
XX
PN   US2002160382-A1.
XX
PD   31-OCT-2002.
XX
PF   11-OCT-2001; 2001US-00981353.
XX
PR   11-OCT-2000; 2000US-0239841P.
XX
PA   (LASE/) LASEK A W.
PA   (JONE/) JONES D A.
XX
PI   Lasek AW,  Jones DA;
XX
DR   WPI; 2003-265756/26.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   colon disorder, useful for diagnosing, treating, staging or monitoring
PT   treatment for colon cancers.
XX
PS   Claim 1; SEQ ID NO 88; 231pp; English.
XX
CC   The invention relates to a combination comprising cDNAs that are
CC   differentially expressed in colon disorder. The methods and compositions
CC   of the present invention are useful for diagnosing, treating, staging or
CC   monitoring treatment for colon cancer. They are also useful in high
CC   throughput methods for using cDNAs to detect differential expression of
CC   nucleic acids in a sample, screening molecules or compounds to identify a
CC   ligand which specifically binds a cDNA and using a protein to screen
CC   molecules or compounds to identify at least one ligand which specifically
CC   binds the protein. The present sequence represents a human cDNA
CC   differentially expressed in colon cancer.
XX
SQ   Sequence 3152 BP; 875 A; 643 C; 707 G; 926 T; 0 U; 1 Other;
```

```
Alignment Scores:
Pred. No.:              5.14e-50        Length:         3152
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x ADA10970 (1-3152)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1411 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 1470

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1471 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 1530

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1531 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 1590

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1591 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 1650

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db       1651 CGAGGGGTGTGTTCCTACTTA 1671

RESULT 13
AAH08073
ID   AAH08073 standard; cDNA; 752 BP.
XX
AC   AAH08073;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:4908.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 1; SEQ ID NO 4908; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 752 BP; 225 A; 148 C; 200 G; 176 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:              1.27e-49        Length:         752
Score:                  439.00          Matches:        86
Percent Similarity:     98.9%           Conservative:   0
Best Local Similarity:  98.9%           Mismatches:     1
Query Match:            98.0%           Indels:         0
DB:                     4               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x AAH08073 (1-752)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGCTCCTACTTA 420

RESULT 14
AAH15146
ID   AAH15146 standard; cDNA; 1898 BP.
XX
AC   AAH15146;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:13209.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
```

PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 13209; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 1898 BP; 590 A; 327 C; 426 G; 555 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.36e-49	Length:	1898
Score:	439.00	Matches:	86
Percent Similarity:	98.9%	Conservative:	0
Best Local Similarity:	98.9%	Mismatches:	1
Query Match:	98.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x AAH15146 (1-1898)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
```

Db 400 CGAGGGGTGTGCTCCTACTTA 420

RESULT 15
ADT95307/c
ID ADT95307 standard; cDNA; 391 BP.
XX
AC ADT95307;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #826.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 826; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 391 BP; 83 A; 100 C; 103 G; 104 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	1.31e-46	Length:	391
Score:	414.00	Matches:	81

```
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            92.4%           Indels:        0
DB:                     11              Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x ADT95307 (1-391)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 227

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 167

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 107

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 47

Qy         81 Arg 81
              |||
Db         46 CGA 44
```

```
RESULT 16
ADX41789/c
ID   ADX41789 standard; cDNA; 391 BP.
XX
AC   ADX41789;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 826.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 826; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 391 BP; 83 A; 100 C; 103 G; 104 T; 0 U; 1 Other;
```

```
Alignment Scores:
Pred. No.:              1.31e-46        Length:        391
Score:                  414.00          Matches:       81
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            92.4%           Indels:        0
DB:                     11              Gaps:          0

US-10-757-745-2_COPY_54_140 (1-87) x ADX41789 (1-391)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 227

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 167

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 107

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 47

Qy         81 Arg 81
              |||
Db         46 CGA 44
```

```
RESULT 17
AAS86255
ID   AAS86255 standard; cDNA; 1227 BP.
XX
AC   AAS86255;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #22059.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG22068.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
```

```
PS   Claim 1; SEQ ID NO 22059; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1227 BP; 311 A; 239 C; 360 G; 317 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.94e-35        Length:          1227
Score:                    332.50          Matches:         84
Percent Similarity:       89.5%           Conservative:    1
Best Local Similarity:    88.4%           Mismatches:      2
Query Match:              74.2%           Indels:          8
DB:                       5               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x AAS86255 (1-1227)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSer-AlaLeuGluAr 20
              ||||||||||||||||||||||||||||||||||||||||||||    |||||||||| ||||||||||
Db        235 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGCGGAGGAGAGCAGCCTTGGAACG 294

Qy         20 gArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrTh 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 CCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 354

Qy         40 rAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn-GlySerM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        355 TGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGGCAGCA 414

Qy         60 etPheSerLeu-IleThrTrp-AsnIle-AspGlyLeuAsp---LeuAsnAsnLeu-Ser 77
              |||||||||| |||||||| |||||| |||||||:::     |||||||||||| |||
Db        415 TGTTCTCTCTTCATTACCTGGGAATATTTGATGGGATTAGGATCTTAAACAATCTGTTCA 474

Qy         78 Glu-ArgAlaArgGlyValCysSerTyrLeu 87
              ||| ||||||||||||||||||||||||||||
Db        475 GAGGAGGGCTCGAGGGGTGTGTTCCTACTTA 505

RESULT 18
AAZ47119
ID   AAZ47119 standard; cDNA; 1312 BP.
XX
AC   AAZ47119;
XX
DT   15-MAR-2000  (first entry)
XX
DE   Mouse CD40 receptor associated protein gene.
XX
KW   Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW   immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW   mouse; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW   tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW   TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW   diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW   arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW   graft versus host disease; autoimmune disease; ds.
XX
OS   Mus musculus.
XX
PN   WO9955859-A2.
XX
PD   04-NOV-1999.
XX
PF   28-APR-1999;   99WO-EP003025.
XX
PR   29-APR-1998;   98EP-00201392.
XX
PA   (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI   Pype SMC,  Remacle JEFJG,  Huylebroeck DFE;
XX
DR   WPI; 2000-062029/05.
DR   P-PSDB; AAY56020.
XX
PT   Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT   diseases and for improvement of anti-tumor treatments.
XX
PS   Claim 10; Page 41-43; 48pp; English.
XX
CC   This sequence represents the gene encoding mouse CD40 receptor associated
CC   protein (CRAP). CRAP is a functional protein capable of interacting with
CC   the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC   necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC   I, where the protein has no homology to TNF receptor associated factor
CC   (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC   the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC   related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC   diseases, and for the improvement of anti-tumour diseases. Diseases which
CC   may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC   systemic lupus erythematosus, graft rejection, graft versus host disease,
CC   allergy, and autoimmune disease. The proteins can be used to sensitize
CC   tumour cells to anti-tumour treatments and to screen for compounds which
CC   interfere with the interaction of the proteins with other protein
CC   components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ   Sequence 1312 BP; 359 A; 279 C; 352 G; 322 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.01e-18        Length:          1312
Score:                    211.50          Matches:         45
Percent Similarity:       66.3%           Conservative:    12
Best Local Similarity:    52.3%           Mismatches:      28
Query Match:              47.2%           Indels:          1
DB:                       3               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x AAZ47119 (1-1312)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::||||||::::::||||||||    |||       :::             |||:::
Db        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||  |||         :::   :::   |||||||||||||||||||||:::       |||
Db        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436

Qy         42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
              :::|||       :::   |||||||||       |||       |||:::   |||
Db        437 ACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGCAGCACTATT 493

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
              |||   ||||||||||||||||||||||||||       ||||||   ||||||||||||
Db        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553

Qy         82 GlyValCysSerTyrLeu 87
              ||||||||||||   |||
Db        554 GGGGTGTGTTCCTGCCTA 571

RESULT 19
```

ADD34376
ID ADD34376 standard; DNA; 602 BP.
XX
AC ADD34376;
XX
DT 15-JAN-2004 (first entry)
XX
DE Mouse mitochondrial DNA sequence SEQ ID NO:2154.
XX
KW ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW mitochondrial disease; oxidative phosphorylation dysfunction;
KW oxidative stress; apoptosis; aging.
XX
OS Mus musculus.
XX
PN WO2003020220-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027886.
XX
PR 30-AUG-2001; 2001US-0316323P.
PR 31-AUG-2001; 2001CA-02356540.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Wallace DC, Levy S, Kerstann K, Procaccio V;
XX
DR WPI; 2003-300821/29.
XX
PT Array containing probes for genes involved in mitochondrial biology,
PT useful for determining mitochondrial biology gene expression profiles for
PT use in diagnosing pathologies and identifying biochemical pathways.
XX
PS Claim 2; SEQ ID NO 2154; 201pp; English.
XX
CC The invention relates to a novel array comprising at least two isolated
CC nucleotide molecules, each molecule having a sequence capable of uniquely
CC hybridising to a nucleic acid molecule which is an expression product of
CC a gene involved in mitochondrial biology. The array comprises two or more
CC isolated nucleic acid molecules or spots, each molecule having a sequence
CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC of the invention is useful for determining an expression profile of a
CC mouse or human sample containing nucleic acid, by contacting the array
CC with the sample under conditions allowing selective hybridisation, and
CC measuring hybridisation of nucleic acid in the sample to the array to
CC produce an expression profile. The array is also useful for determining
CC an expression profile of a first labelled sample containing nucleic acid
CC relative to a second, differently labelled sample containing nucleic
CC acid. The second sample is a reference or a standard. An array is useful
CC for determining an expression profile diagnostic of an energy-metabolism-
CC related physiological condition. An array of the invention is useful for
CC determining mitochondrial biology gene expression profiles of organisms,
CC such as human, mice and closely related species, tissue and organs of
CC such organisms, which are useful for determining expression profiles
CC diagnostic of energy metabolism-related physiological conditions,
CC diagnosing such physiological conditions, identifying biochemical
CC pathways, genes, and mutations involved in such physiological conditions,
CC identifying therapeutic agents useful for preventing and/or treating such
CC physiological conditions, evaluating and/or monitoring the efficacy of
CC such therapies, and creating and identifying animal models of human
CC energy metabolism-related physiological conditions. An array is also
CC useful for defining expression signatures or profiles for mitochondrial
CC diseases, as well as distinguishing clinical disorders that result from
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC apoptosis and aging. An array of the invention contains probes of genes
CC not previously recognised to participate in mitochondrial biology. The
CC sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC clones used to make the probes of the invention. Some sequences are not
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC 1906, 2408 and 2643.
XX
SQ Sequence 602 BP; 194 A; 119 C; 134 G; 155 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.14e-14	Length:	602
Score:	178.50	Matches:	39
Percent Similarity:	67.6%	Conservative:	7
Best Local Similarity:	57.4%	Mismatches:	21
Query Match:	39.8%	Indels:	1
DB:	10	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x ADD34376 (1-602)

```
Qy         20 ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 39
              |||:::|||   |||      :::   :::   ||||||||||||||||||||:::
Db         15 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 74

Qy         40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 59
                 |||:::|||      :::   |||||||||      |||      |||:::   |||
Db         75 AATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGCAGC 131

Qy         60 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                 |||   ||||||||||||||||||||||||||      ||||||   ||||||
Db        132 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 191

Qy         80 AlaArgGlyValCysSerTyrLeu 87
              |||||||||||||||||   |||
Db        192 GCTCGAGGGGTGTGTTCCTGCCTA 215
```

RESULT 20
ABQ34465
ID ABQ34465 standard; DNA; 725 BP.
XX
AC ABQ34465;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 21056.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on

CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 725 BP; 318 A; 213 C; 64 G; 130 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.56e-05	Length:	725
Score:	114.00	Matches:	19
Percent Similarity:	96.6%	Conservative:	9
Best Local Similarity:	65.5%	Mismatches:	1
Query Match:	25.4%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x ABQ34465 (1-725)

```
Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::  |||||||||||||||||||||||||:::::::::::::|||||||:::||||||

Db        571 CAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAAACGCCTTAAAACGCCGA 630

Qy         22 ProGluThrIleSerGluProLysThr 30
                 |||:::|||||||||:::|||||||||

Db        631 CCTAAAACCATCTCTAAACCCAAAACC 657
```

RESULT 21
ABQ34464/c
ID ABQ34464 standard; DNA; 725 BP.
XX
AC ABQ34464;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 21055.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 725 BP; 130 A; 64 C; 213 G; 318 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.56e-05	Length:	725
Score:	114.00	Matches:	19
Percent Similarity:	96.6%	Conservative:	9
Best Local Similarity:	65.5%	Mismatches:	1
Query Match:	25.4%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x ABQ34464 (1-725)

```
Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              ::::::  |||||||||||||||||||||||||:::::::::::::|||||||:::||||||

Db        155 CAAAAAACTCTAAACTCCTACTTCGAACCTCCGATAAAAAAAAAACGCCTTAAAACGCCGA 96

Qy         22 ProGluThrIleSerGluProLysThr 30
                 |||:::|||||||||:::|||||||||

Db         95 CCTAAAACCATCTCTAAACCCAAAACC 69
```

RESULT 22
AAX41008
ID AAX41008 standard; cDNA; 258 BP.
XX
AC AAX41008;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human secreted protein 5' EST SEQ ID NO: 220.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS Homo sapiens.
XX
PN WO9906554-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB001238.
XX
PR 01-AUG-1997; 97US-00905134.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
DR WPI; 1999-153784/13.
DR P-PSDB; AAY12175.
XX

PT New nucleic acids encoding human secreted proteins - obtained from cDNA
PT libraries prepared from kidney, fetal kidney, dystrophic muscle, muscle
PT and heart tissue.
XX
PS Claim 1; Page 319; 622pp; English.
XX
CC AAX40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY01602 and
CC AAY11994 to AAY12260, respectively. The proteins given represent the
CC signal peptide and an N-terminal fragment of a secreted protein. The
CC nucleic acid sequences can be used for producing secreted human gene
CC products. They can also be used to develop products for diagnosis and
CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 258 BP; 54 A; 54 C; 105 G; 40 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	5.38e-06	Length:	258
Score:	113.00	Matches:	22
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	25.2%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x AAX41008 (1-258)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 249

Qy         21 ArgPro 22
              ||||||
Db        250 CGACCA 255
```

RESULT 23
ABN41860
ID AC ABN41860 standard; DNA; 60 BP.
XX
AC ABN41860;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:14608.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB001903.
XX
PR 28-JUL-2000; 2000US-0221607P.
PR 02-MAY-2001; 2001US-0287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of a
PT genome, useful for detecting tissue-, pathology-, and developmental-
PT specific genes.
XX
PS Example 1; SEQ ID NO 14608; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises several
CC oligonucleotides, each capable of hybridising selectively to a set of
CC messenger RNAs transcribed from a given transcription unit of the genome,
CC which encodes one or more messenger RNA splice variants. The
CC oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a particular
CC biological or pathological state, and so allowing the detection of tissue
CC - and pathology-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rats, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 60 BP; 19 A; 16 C; 11 G; 14 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.27e-05	Length:	60
Score:	101.00	Matches:	19
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	22.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x ABN41860 (1-60)

```
Qy         48 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGG 58
```

RESULT 24
AAS87500
ID AAS87500 standard; cDNA; 1589 BP.
XX
AC AAS87500;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #23304.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.

```
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG23313.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 23304; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1589 BP; 475 A; 331 C; 359 G; 424 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.447          Length:          1589
Score:                    84.50          Matches:           26
Percent Similarity:       47.2%          Conservative:      16
Best Local Similarity:    29.2%          Mismatches:        26
Query Match:              18.9%          Indels:            21
DB:                       5              Gaps:               5

US-10-757-745-2_COPY_54_140 (1-87) x AAS87500 (1-1589)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||            |||||||:::   |||   ||||||
Db        638 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 685

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||     |||:::|||||||||||||:::|||             |||
Db        686 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 745

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                  :::      :::|||   :::   :::|||                  |||   :::
Db        746 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 790

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db        791 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 850

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db        851 TCTAAAGCAGAGGGTATTAAGAACTAT 877

RESULT 25
AAS87499
ID   AAS87499 standard; cDNA; 2228 BP.
XX
AC   AAS87499;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #23303.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG23312.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 23303; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2228 BP; 657 A; 468 C; 501 G; 602 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.702          Length:          2228
Score:                    84.50          Matches:           26
Percent Similarity:       47.2%          Conservative:      16
Best Local Similarity:    29.2%          Mismatches:        26
Query Match:              18.9%          Indels:            21
DB:                       5              Gaps:               5

US-10-757-745-2_COPY_54_140 (1-87) x AAS87499 (1-2228)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||            |||||||:::   |||   ||||||
Db       1028 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 1075

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||     |||:::|||||||||||||:::|||             |||
Db       1076 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 1135
```

```
Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||               |||   :::
Db       1136 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 1180

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       1181 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 1240

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       1241 TCTAAAGCAGAGGGTATTAAGAACTAT 1267
```

RESULT 26
AAS82908/c
ID AAS82908 standard; cDNA; 2717 BP.
XX
AC AAS82908;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #18712.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG18721.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 18712; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 2717 BP; 755 A; 606 C; 559 G; 797 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.914	Length:	2717
Score:	84.50	Matches:	26
Percent Similarity:	47.2%	Conservative:	16
Best Local Similarity:	29.2%	Mismatches:	26
Query Match:	18.9%	Indels:	21
DB:	5	Gaps:	5

US-10-757-745-2_COPY_54_140 (1-87) x AAS82908 (1-2717)

```
Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||           ||||||:::   |||   ||||||
Db       2109 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 2062

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::|||||||||||:::|||             |||
Db       2061 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 2002

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||               |||   :::
Db       2001 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 1957

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       1956 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 1897

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       1896 TCTAAAGCAGAGGGTATTAAGAACTAT 1870
```

RESULT 27
AAS75995
ID AAS75995 standard; cDNA; 4002 BP.
XX
AC AAS75995;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #11799.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG11808.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 11799; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

```
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 4002 BP; 1159 A; 837 C; 891 G; 1115 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.53          Length:         4002
Score:                    84.50         Matches:        26
Percent Similarity:       47.2%         Conservative:   16
Best Local Similarity:    29.2%         Mismatches:     26
Query Match:              18.9%         Indels:         21
DB:                       5             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x AAS75995 (1-4002)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||             ||||||:::   |||   ||||||
Db       1435 GAACCTACTACTAAAGAGTCT------------AGACCACAGGCTATAAAAGAACCTGTG 1482

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||       |||:::|||||||||||||:::|||           |||
Db       1483 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 1542

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                  :::      :::|||   :::   :::|||               |||   :::
Db       1543 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 1587

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                 :::
Db       1588 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 1647

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                :::|||   |||:::   :::|||
Db       1648 TCTAAAGCAGAGGGTATTAAGAACTAT 1674

RESULT 28
AAS87489/c
ID   AAS87489 standard; cDNA; 4890 BP.
XX
AC   AAS87489;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #23293.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG23302.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 23293; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 4890 BP; 1388 A; 1100 C; 986 G; 1416 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2             Length:         4890
Score:                    84.50         Matches:        26
Percent Similarity:       47.2%         Conservative:   16
Best Local Similarity:    29.2%         Mismatches:     26
Query Match:              18.9%         Indels:         21
DB:                       5             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x AAS87489 (1-4890)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||             ||||||:::   |||   ||||||
Db       2181 GAACCTACTACTAAAGAGTCT------------AGACCACAGGCTATAAAAGAACCTGTG 2134

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||       |||:::|||||||||||||:::|||           |||
Db       2133 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 2074

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                  :::      :::|||   :::   :::|||               |||   :::
Db       2073 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 2029

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                 :::
Db       2028 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 1969

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                :::|||   |||:::   :::|||
Db       1968 TCTAAAGCAGAGGGTATTAAGAACTAT 1942

RESULT 29
AAS82911
ID   AAS82911 standard; cDNA; 6448 BP.
```

```
XX
AC   AAS82911;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #18715.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG18724.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 18715; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 6448 BP; 1981 A; 1307 C; 1393 G; 1767 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.89          Length:         6448
Score:                  84.50         Matches:        26
Percent Similarity:     47.2%         Conservative:   16
Best Local Similarity:  29.2%         Mismatches:     26
Query Match:            18.9%         Indels:         21
DB:                     5             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x AAS82911 (1-6448)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||           ||||||:::   |||   ||||||
Db       1028 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 1075

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::|||||||||||||:::|||            |||
Db       1076 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 1135

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||                |||   :::
Db       1136 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 1180

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                  :::
Db       1181 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 1240

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
              :::|||   |||:::   :::|||
Db       1241 TCTAAAGCAGAGGGTATTAAGAACTAT 1267

RESULT 30
AAS87485
ID   AAS87485 standard; cDNA; 7197 BP.
XX
AC   AAS87485;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #23289.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG23298.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 23289; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
```

```
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 7197 BP; 2183 A; 1419 C; 1640 G; 1955 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.35          Length:         7197
Score:                  84.50         Matches:        26
Percent Similarity:     47.2%         Conservative:   16
Best Local Similarity:  29.2%         Mismatches:     26
Query Match:            18.9%         Indels:         21
DB:                     5             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x AAS87485 (1-7197)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||            |||||:::   |||   ||||||
Db       5158 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 5205

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::|||||||||||||:::|||           |||
Db       5206 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 5265

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||                |||   :::
Db       5266 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 5310

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||               :::
Db       5311 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 5370

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                ::::|||   |||:::   :::|||
Db       5371 TCTAAAGCAGAGGGTATTAAGAACTAT 5397

RESULT 31
AAS82893/c
ID   AAS82893 standard; cDNA; 7556 BP.
XX
AC   AAS82893;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #18697.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG18706.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 18697; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 7556 BP; 2135 A; 1725 C; 1515 G; 2181 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.57          Length:         7556
Score:                  84.50         Matches:        26
Percent Similarity:     47.2%         Conservative:   16
Best Local Similarity:  29.2%         Mismatches:     26
Query Match:            18.9%         Indels:         21
DB:                     5             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x AAS82893 (1-7556)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||            |||||:::   |||   ||||||
Db       4740 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 4693

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::|||||||||||||:::|||           |||
Db       4692 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 4633

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||                |||   :::
Db       4632 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 4588

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||               :::
Db       4587 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 4528

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                ::::|||   |||:::   :::|||
Db       4527 TCTAAAGCAGAGGGTATTAAGAACTAT 4501

RESULT 32
AAS87486/c
ID   AAS87486 standard; cDNA; 7731 BP.
XX
AC   AAS87486;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #23290.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
```

```
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG23299.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 23290; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 7731 BP; 2158 A; 1663 C; 1546 G; 2364 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.68          Length:        7731
Score:                  84.50         Matches:       26
Percent Similarity:     47.2%         Conservative:  16
Best Local Similarity:  29.2%         Mismatches:    26
Query Match:            18.9%         Indels:        21
DB:                     5             Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x AAS87486 (1-7731)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||            |||||||:::   |||   ||||||
Db       3801 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 3754

Qy         29 ------------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                        |||      |||:::|||||||||||||:::|||          |||
Db       3753 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 3694

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||                |||   :::
Db       3693 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 3649

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                :::
Db       3648 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 3589

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       3588 TCTAAAGCAGAGGGTATTAAGAACTAT 3562
```

```
RESULT 33
AAS87514
ID   AAS87514 standard; cDNA; 11742 BP.
XX
AC   AAS87514;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #23318.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG23327.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 23318; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 11742 BP; 3484 A; 2354 C; 2638 G; 3266 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.43          Length:        11742
Score:                  84.50         Matches:       26
Percent Similarity:     47.2%         Conservative:  16
Best Local Similarity:  29.2%         Mismatches:    26
Query Match:            18.9%         Indels:        21
DB:                     5             Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x AAS87514 (1-11742)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
```

```
Db      10954 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 11001

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44

Db      11002 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 11061

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64

Db      11062 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 11106

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77

Db      11107 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 11166

Qy         78 GluArgAlaArgGlyValCysSerTyr 86

Db      11167 TCTAAAGCAGAGGGTATTAAGAACTAT 11193
```

```
RESULT 34
AAS75996
ID   AAS75996 standard; cDNA; 12036 BP.
XX
AC   AAS75996;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #11800.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG11809.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 11800; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 12036 BP; 3483 A; 2390 C; 2747 G; 3416 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:                6.64          Length:          12036
Score:                    84.50         Matches:         26
Percent Similarity:       47.2%         Conservative:    16
Best Local Similarity:    29.2%         Mismatches:      26
Query Match:              18.9%         Indels:          21
DB:                       5             Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x AAS75996 (1-12036)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28

Db       4558 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 4605

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44

Db       4606 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 4665

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64

Db       4666 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 4710

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77

Db       4711 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 4770

Qy         78 GluArgAlaArgGlyValCysSerTyr 86

Db       4771 TCTAAAGCAGAGGGTATTAAGAACTAT 4797
```

```
RESULT 35
AAS75998
ID   AAS75998 standard; cDNA; 23157 BP.
XX
AC   AAS75998;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #11802.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG11811.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits to assess
PT   biodiversity.
```

```
XX
PS   Claim 1; SEQ ID NO 11802; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 23157 BP; 6796 A; 4596 C; 5194 G; 6571 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                15.9          Length:        23157
Score:                    84.50         Matches:       26
Percent Similarity:       47.2%         Conservative:  16
Best Local Similarity:    29.2%         Mismatches:    26
Query Match:              18.9%         Indels:        21
DB:                       5             Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x AAS75998 (1-23157)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||            ||||||:::   |||   ||||||
Db      15505 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 15552

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::|||             |||
Db      15553 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 15612

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                  :::      :::|||   :::   :::|||                 |||   :::
Db      15613 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 15657

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db      15658 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 15717

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                ::|||   |||:::   :::|||
Db      15718 TCTAAAGCAGAGGGTATTAAGAACTAT 15744

RESULT 36
AAS82912
ID   AAS82912 standard; cDNA; 32313 BP.
XX
AC   AAS82912;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #18716.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG18725.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 18716; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 32313 BP; 9408 A; 6372 C; 7230 G; 9303 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                24.8          Length:        32313
Score:                    84.50         Matches:       26
Percent Similarity:       47.2%         Conservative:  16
Best Local Similarity:    29.2%         Mismatches:    26
Query Match:              18.9%         Indels:        21
DB:                       5             Gaps:          5

US-10-757-745-2_COPY_54_140 (1-87) x AAS82912 (1-32313)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||            ||||||:::   |||   ||||||
Db      24660 GAACCTACTACTAAAGAGTCT-----------AGACCACAGGCTATAAAAGAACCTGTG 24707

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::|||             |||
Db      24708 AGCCAGTCTGAAAAACGGGATGTTAACCTTACTAATGAGGACACCATCAAGCAACCTACT 24767

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                  :::      :::|||   :::   :::|||                 |||   :::
Db      24768 GAATCTGTTAAAGAAACTGAAACTTCTGCAAAAGAA--------------AGCACAGTT 24812

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db      24813 ACAGAAGAATTAAGAGAAGGTATCGATGCTGTTTACCCCTCATTGGTAGGTACTGCTGAT 24872

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
```

```
                  :::|||   |||:::   :::|||
Db      24873 TCTAAAGCAGAGGGTATTAAGAACTAT 24899
```

RESULT 37
ADF90711
ID   ADF90711 standard; DNA; 6115 BP.
XX
AC   ADF90711;
XX
DT   26-FEB-2004 (first entry)
XX
DE   Human hepatic-fibrosis disease marker SEQ ID 173.
XX
KW   Hepatic fibrosis; marker; chronic hepatitis; liver cirrhosis;
KW   hepatic carcinoma; human; ds.
XX
OS   Homo sapiens.
XX
PN   JP2003259877-A.
XX
PD   16-SEP-2003.
XX
PF   11-MAR-2002; 2002JP-00065013.
XX
PR   11-MAR-2002; 2002JP-00065013.
XX
PA   (SUMU) SUMITOMO SEIYAKU KK.
XX
DR   WPI; 2003-821598/77.
XX
PT   Hepatic fibrosis disease markers comprising polynucleotides or
PT   antibodies, useful for improved diagnosis, screening and developing drugs
PT   to treat hepatitis, to control cirrhosis and carcinoma.
XX
PS   Claim 1; SEQ ID NO 173; 313pp; Japanese.
XX
CC   The present invention relates to hepatic-fibrosis disease markers
CC   (ADF90539-ADF90871) and related proteins (ADF90872-ADF90917). The
CC   sequences are useful for detecting and treating hepatic fibrosis caused
CC   by alcohol consumption, virus infection, etc., and the associated chronic
CC   hepatitis, etc. leading to liver cirrhosis and hepatic carcinoma. The
CC   markers allow the cause of hepatic fibrosis to be clarified (diagnostic
CC   precision), so more suitable treatments can be developed and given.
XX
SQ   Sequence 6115 BP; 1454 A; 1561 C; 1568 G; 1484 T; 0 U; 48 Other;

Alignment Scores:

Pred. No.:	4.31	Length:	6115
Score:	83.00	Matches:	25
Percent Similarity:	47.8%	Conservative:	7
Best Local Similarity:	37.3%	Mismatches:	33
Query Match:	18.5%	Indels:	2
DB:	10	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x ADF90711 (1-6115)

```
Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSer 26
              |||   |||  ||||||:::          ||||||:::              |||
Db       3973 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCACTAGGATCTGC 4032

Qy         27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys 46
              :::      |||   |||   |||      |||:::|||         |||||||||
Db       4033 CAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTCC 4092

Qy         47 ------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                    |||  ||||||   |||       |||   ||||||:::|||:::
Db       4093 CCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTGT 4152

Qy         65 ThrTrpAsnIleAspGlyLeu 71
                |||   :::     |||
Db       4153 GTCTGGACTTTGAGCAGGCTT 4173
```

RESULT 38

AEB47090
ID   AEB47090 standard; DNA; 37741 BP.
XX
AC   AEB47090;
XX
DT   06-OCT-2005 (first entry)
XX
DE   Chimpanzee adenovirus, ChAd3, genomic DNA.
XX
KW   cytostatic; vaccine; gene therapy; pharmaceutical; cancer; neoplasm;
KW   immunogenicity; ChAd3; ds; gene.
XX
OS   Chimpanzee adenovirus.
XX
PN   WO2005071093-A2.
XX
PD   04-AUG-2005.
XX
PF   18-JAN-2005; 2005WO-EP000558.
XX
PR   23-JAN-2004; 2004US-0538799P.
XX
PA   (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
PI   Cirillo A, Colloca S, Ercole BB, Meola A, Nicosia A, Sporeno E;
XX
DR   WPI; 2005-534014/54.
XX
PT   New isolated chimpanzee nucleic acid sequence, useful for generating
PT   chimpanzee adenoviral vectors useful as vaccine carriers to elicit
PT   protective and/or therapeutic immune responses.
XX
PS   Claim 1; SEQ ID NO 1; 356pp; English.
XX
CC   The invention relates to an isolated chimpanzee nucleic acid sequence.
CC   The chimpanzee nucleic acid sequences are useful for generating
CC   chimpanzee adenoviral vectors in human E1-expressing cells lines, which
CC   are useful as vaccine carriers to elicit protective and/or therapeutic
CC   immune responses. The vaccines and pharmaceutical compositions are useful
CC   for the prevention and treatment of cancer. The chimpanzee nucleic acid
CC   sequences are also useful for boosting an antigen-specific immune
CC   response in a mammal, eliciting an immune response in a naive mammal,
CC   inducing an immune response against an antigen derived from an infectious
CC   agent, or breaking host tolerance to a self-antigen. The present sequence
CC   represents a chimpanzee adenovirus genomic DNA.
XX
SQ   Sequence 37741 BP; 7853 A; 11639 C; 11107 G; 7142 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	48.7	Length:	37741
Score:	83.00	Matches:	31
Percent Similarity:	42.4%	Conservative:	11
Best Local Similarity:	31.3%	Mismatches:	22
Query Match:	18.5%	Indels:	35
DB:	14	Gaps:	5

US-10-757-745-2_COPY_54_140 (1-87) x AEB47090 (1-37741)

```
Qy          2 GluArgAlaLeuAsnSerTyrPhe---------------------------------- 9
              |||:::|||:::::|||:::|||
Db        763 GAGCAGGCGGTTTCCAGTTTTTTTCCAGAGTCCATGTTGTTGGCCAGCCAGGAGGGGGTC 822

Qy         10 ------------GluProProValGluGluSer------------AlaLeuGluArgArg 21
                         ||||||:::   :::|||                |||   |||:::
Db        823 GAACTTGAGACCCCTCCTCCGATCGTGGATTCCCCCGATCCGCCGCAGCTGACTAGGCAG 882

Qy         22 ProGluThrIleSerGluProLysThr--------------------TyrValAspLeu 34
              ||||||      :::   |||:::|||                          :::||||||
Db        883 CCCGAGCGCTGTGCGGGACCTGAGACTATGCCCCAGCTGCTACCTGAGGTGATCGATCTC 942

Qy         35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
              |||      |||                 |||          ||||||||||||   :::
Db        943 ACCTGTAATGAG-----------------TCTGGTTTTCCACCCAGCGAGGAT---GAG 981
```

```
Qy         55 GlnGluAsnGlySerMetPheSerLeu-IleThrTrpAsnIleAspGlyLeuAsp 72
Db        982 GACGAAGAGGGTGAGCAGTTTGTGTTAGATTCTGTGGAACAACCCGGGCGAGGAT 1036

RESULT 39
ADD34445/c
ID   ADD34445 standard; DNA; 826 BP.
XX
AC   ADD34445;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Mouse mitochondrial DNA sequence SEQ ID NO:2223.
XX
KW   ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW   mitochondrial disease; oxidative phosphorylation dysfunction;
KW   oxidative stress; apoptosis; aging.
XX
OS   Mus musculus.
XX
PN   WO2003020220-A2.
XX
PD   13-MAR-2003.
XX
PF   30-AUG-2002; 2002WO-US027886.
XX
PR   30-AUG-2001; 2001US-0316323P.
PR   31-AUG-2001; 2001CA-02356540.
XX
PA   (UYEM-) UNIV EMORY.
XX
PI   Wallace DC,  Levy S,  Kerstann K,  Procaccio V;
XX
DR   WPI; 2003-300821/29.
XX
PT   Array containing probes for genes involved in mitochondrial biology,
PT   useful for determining mitochondrial biology gene expression profiles for
PT   use in diagnosing pathologies and identifying biochemical pathways.
XX
PS   Claim 2; SEQ ID NO 2223; 201pp; English.
XX
CC   The invention relates to a novel array comprising at least two isolated
CC   nucleotide molecules, each molecule having a sequence capable of uniquely
CC   hybridising to a nucleic acid molecule which is an expression product of
CC   a gene involved in mitochondrial biology. The array comprises two or more
CC   isolated nucleic acid molecules or spots, each molecule having a sequence
CC   chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC   of the invention is useful for determining an expression profile of a
CC   mouse or human sample containing nucleic acid, by contacting the array
CC   with the sample under conditions allowing selective hybridisation, and
CC   measuring hybridisation of nucleic acid in the sample to the array to
CC   produce an expression profile. The array is also useful for determining
CC   an expression profile of a first labelled sample containing nucleic acid
CC   relative to a second, differently labelled sample containing nucleic
CC   acid. The second sample is a reference or a standard. An array is useful
CC   for determining an expression profile diagnostic of an energy-metabolism-
CC   related physiological condition. An array of the invention is useful for
CC   determining mitochondrial biology gene expression profiles of organisms,
CC   such as human, mice and closely related species, tissue and organs of
CC   such organisms, which are useful for determining expression profiles
CC   diagnostic of energy metabolism-related physiological conditions,
CC   diagnosing such physiological conditions, identifying biochemical
CC   pathways, genes, and mutations involved in such physiological conditions,
CC   identifying therapeutic agents useful for preventing and/or treating such
CC   physiological conditions, evaluating and/or monitoring the efficacy of
CC   such therapies, and creating and identifying animal models of human
CC   energy metabolism-related physiological conditions. An array is also
CC   useful for defining expression signatures or profiles for mitochondrial
CC   diseases, as well as distinguishing clinical disorders that result from
CC   oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC   apoptosis and aging. An array of the invention contains probes of genes
CC   not previously recognised to participate in mitochondrial biology. The
CC   sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC   clones used to make the probes of the invention. Some sequences are not
CC   present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC   1906, 2408 and 2643.
XX
SQ   Sequence 826 BP; 210 A; 189 C; 149 G; 278 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.478          Length:          826
Score:                    81.50          Matches:         23
Percent Similarity:       46.7%          Conservative:    20
Best Local Similarity:    25.0%          Mismatches:      28
Query Match:              18.2%          Indels:          21
DB:                       10             Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x ADD34445 (1-826)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSer---------------AlaLeuGlu 19
Db        801 TTGCATCAAGTATGGGAACCAAAAGTTTCTGAGACCCTTCAAGAGTTGCCTTCAATGGAG 742

Qy         20 ArgArgProGluThrIleSerGlu------ProLysThrTyrValAspLeuThrAsnGlu 37
Db        741 AAAATCCCACACTCTCTTGATAATCATCTTCCTGATACACACATAGATCTAACGAAAGAT 682

Qy         38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
Db        681 TCAGCCACTGAGACC-----------------------------------AAAAGCTTG 658

Qy         58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn------Asn 75
Db        657 GGGGAGCTAATGGAAGTAACAGTTTTAAATGTTGATCACTTGGAATGTTCTCAAACCAAC 598

Qy         76 LeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
Db        597 TTAGATCAAGATGCAGAGATAACATGTAGTTCTTTA 562

RESULT 40
AAC81290
ID   AAC81290 standard; cDNA; 3321 BP.
XX
AC   AAC81290;
XX
DT   23-FEB-2001  (first entry)
XX
DE   Mouse apoptosis-related protein cDNA, SEQ ID NO:1.
XX
KW   Apoptosis-associated protein; mouse; murine; pro-apoptotic;
KW   Fas-mediated apoptosis induction; drug screening; inhibitor; inducer;
KW   Escherichia coli SANK 70599; FERM BP-6707; ss.
XX
OS   Mus musculus.
XX
PN   WO200065044-A1.
XX
PD   02-NOV-2000.
XX
PF   21-APR-2000; 2000WO-JP002615.
XX
PR   23-APR-1999;   99JP-00117103.
XX
PA   (SANY ) SANKYO CO LTD.
XX
PI   Yonehara S,  Imai Y;
XX
DR   WPI; 2000-687337/67.
DR   P-PSDB; AAB29693.
XX
PT   DNA and encoded apoptosis-associated proteins, applicable in screening
PT   apoptosis inhibitors and apoptosis-induction promoters for treating
PT   diseases due to abnormal Fas-mediated apoptosis.
XX
PS   Claim 5; Page 46-55; 109pp; Japanese.
```

XX
CC The invention relates to a novel mouse apoptosis-associated protein
CC (AAB29693) and to nucleic acids encoding it (AAC81290). The mouse
CC apoptosis- associated protein potentiates the induction of Fas-mediated
CC apoptosis. The invention also relates to variants of the protein which
CC retain activity, recombinant vectors and host cells comprising DNA
CC encoding mouse apoptosis-associated protein, recombinant production of
CC the protein, an antibody which specifically binds the protein, and
CC methods of screening compounds for their ability to promote or inhibit
CC apoptosis induction. The invention additionally discloses Escherichia
CC coli SANK 70599 (FERM BP-6707) which contains a plasmid encoding mouse
CC apoptosis- associated protein. The mouse apoptosis-associated protein and
CC DNA encoding it may be used to screen potential apoptosis inhibitors and
CC promoters of apoptosis induction. The compounds thus identified can be
CC used in the treatment of diseases associated with abnormal Fas-mediated
CC apoptosis. The present sequence represents cDNA encoding mouse apoptosis-
CC associated protein
XX
SQ Sequence 3321 BP; 1163 A; 578 C; 719 G; 861 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.05	Length:	3321
Score:	81.50	Matches:	23
Percent Similarity:	46.7%	Conservative:	20
Best Local Similarity:	25.0%	Mismatches:	28
Query Match:	18.2%	Indels:	21
DB:	3	Gaps:	4

US-10-757-745-2_COPY_54_140 (1-87) x AAC81290 (1-3321)

```
Qy          5 LeuAsnSerTyrPheGluProProValGluGluSer---------------AlaLeuGlu 19
              |||:::      :::||||||   |||   |||:::                :::::|||
Db       2320 TTGCATCAAGTATGGGAACCAAAAGTTTCTGAGAACCTTCAAGAGTTGCCTTCAATGGAG 2379

Qy         20 ArgArgProGluThrIleSerGlu------ProLysThrTyrValAspLeuThrAsnGlu 37
              :::  |||  ::::::          |||   |||::::::|||||||||   :::
Db       2380 AAAATCCCACACTCTCTTGATAATCATCTTCCTGATACACACATAGATCTAACGAAAGAT 2439

Qy         38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
                 |||::::::                                    :::
Db       2440 TCAGCCACTGAGACC-----------------------------------AAAAGCTTG 2463

Qy         58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn------Asn 75
              |||   :::     :::      |||:::|||   |||:::   :::      |||
Db       2464 GGGGAGCTAATGGAAGTAACAGTTTTAAATGTTGATCACTTGGAATGTTCTCAAACCAAC 2523

Qy         76 LeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||   :::   |||         ||||||  |||
Db       2524 TTAGATCAAGATGCAGAGATAACATGTAGTTCTTTA 2559
```

RESULT 41
AAC81309
ID AAC81309 standard; cDNA; 6763 BP.
XX
AC AAC81309;
XX
DT 23-FEB-2001 (first entry)
XX
DE Mouse FLASH protein cDNA, SEQ ID NO:21.
XX
KW FLASH protein; FLICE-associated huge protein;
KW apoptosis-associated protein; mouse; murine; pro-apoptotic;
KW Fas-mediated apoptosis induction; drug screening; inhibitor; inducer;
KW Escherichia coli SANK 70599; FERM BP-6707; ss.
XX
OS Mus musculus.
XX
PN WO200065044-A1.
XX
PD 02-NOV-2000.
XX
PF 21-APR-2000; 2000WO-JP002615.
XX
PR 23-APR-1999; 99JP-00117103.
XX
PA (SANY) SANKYO CO LTD.
XX
PI Yonehara S, Imai Y;
XX
DR WPI; 2000-687337/67.
DR P-PSDB; AAB29694.
XX
PT DNA and encoded apoptosis-associated proteins, applicable in screening
PT apoptosis inhibitors and apoptosis-induction promoters for treating
PT diseases due to abnormal Fas-mediated apoptosis.
XX
PS Example; Page 72-90; 109pp; Japanese.
XX
CC The invention relates to a novel mouse apoptosis-associated protein
CC (AAB29693) and to nucleic acids encoding it (AAC81290). The mouse
CC apoptosis- associated protein potentiates the induction of Fas-mediated
CC apoptosis. The invention also relates to variants of the protein which
CC retain activity, recombinant vectors and host cells comprising DNA
CC encoding mouse apoptosis-associated protein, recombinant production of
CC the protein, an antibody which specifically binds the protein, and
CC methods of screening compounds for their ability to promote or inhibit
CC apoptosis induction. The invention additionally discloses Escherichia
CC coli SANK 70599 (FERM BP-6707) which contains a plasmid encoding mouse
CC apoptosis- associated protein. The mouse apoptosis-associated protein and
CC DNA encoding it may be used to screen potential apoptosis inhibitors and
CC promoters of apoptosis induction. The compounds thus identified can be
CC used in the treatment of diseases associated with abnormal Fas-mediated
CC apoptosis. The present sequence represents cDNA encoding mouse FLASH
CC (FLICE-associated huge) protein which is involved in Fas-mediated
CC activation of caspase-8 during apoptosis
XX
SQ Sequence 6763 BP; 2286 A; 1235 C; 1530 G; 1712 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.87	Length:	6763
Score:	81.50	Matches:	23
Percent Similarity:	46.7%	Conservative:	20
Best Local Similarity:	25.0%	Mismatches:	28
Query Match:	18.2%	Indels:	21
DB:	3	Gaps:	4

US-10-757-745-2_COPY_54_140 (1-87) x AAC81309 (1-6763)

```
Qy          5 LeuAsnSerTyrPheGluProProValGluGluSer---------------AlaLeuGlu 19
              |||:::      :::||||||   |||   |||:::                :::::|||
Db       5140 TTGCATCAAGTATGGGAACCAAAAGTTTCTGAGAACCTTCAAGAGTTGCCTTCAATGGAG 5199

Qy         20 ArgArgProGluThrIleSerGlu------ProLysThrTyrValAspLeuThrAsnGlu 37
              :::  |||  ::::::          |||   |||::::::|||||||||   :::
Db       5200 AAAATCCCACACTCTCTTGATAATCATCTTCCTGATACACACATAGATCTAACGAAAGAT 5259

Qy         38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
                 |||::::::                                    :::
Db       5260 TCAGCCACTGAGACC-----------------------------------AAAAGCTTG 5283

Qy         58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsn------Asn 75
              |||   :::     :::      |||:::|||   |||:::   :::      |||
Db       5284 GGGGAGCTAATGGAAGTAACAGTTTTAAATGTTGATCACTTGGAATGTTCTCAAACCAAC 5343

Qy         76 LeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
              |||   :::   |||         ||||||  |||
Db       5344 TTAGATCAAGATGCAGAGATAACATGTAGTTCTTTA 5379
```

RESULT 42
AAD41820/c
ID AAD41820 standard; cDNA; 1473 BP.
XX
AC AAD41820;
XX
DT 30-OCT-2002 (first entry)
XX

```
DE   Human transporter protein encoding cDNA.
XX
KW   Human; transporter protein; immune response; pharmacogenomic analysis;
KW   tissue typing; gene therapy; transgenic animal; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1. .189
FT                   /*tag=  a
FT   CDS             190. .1473
FT                   /*tag=  b
FT                   /product= "Human transporter protein"
XX
PN   WO200248363-A2.
XX
PD   20-JUN-2002.
XX
PF   12-DEC-2001; 2001WO-US047461.
XX
PR   12-DEC-2000; 2000US-0254588P.
PR   02-APR-2001; 2001US-00822859.
PR   13-APR-2001; 2001US-00833789.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Webster M,  Di Francesco V,  Beasley EM;
XX
DR   WPI; 2002-500534/53.
DR   P-PSDB; AAE25571.
XX
PT   Novel human transporter proteins and nucleic acids, useful for
PT   identifying modulators useful for treating disease or condition mediated
PT   by the protein.
XX
PS   Claim 4; Page 85; 96pp; English.
XX
CC   The invention relates to human transporter proteins and their
CC   corresponding nucleic acids. Human transporter peptide is useful for
CC   identifying a modulator that binds to transporter protein. It is used to
CC   raise antibodies or to elicit another immune response. It is also useful
CC   to provide a target for diagnosing a disease or predisposition to disease
CC   mediated by it. It is also useful in pharmacogenomic analysis. The
CC   antibodies of the invention is used to assess expression in disease
CC   states such as in active stages of the disease or in an individual with a
CC   predisposition toward disease related to the protein's function. They are
CC   also used to assess normal and aberrant subcellular localisation of cells
CC   in the various tissues in an organism, and for pharmacogenomic analysis,
CC   tissue typing and inhibiting protein function. The nucleic acid molecules
CC   of the invention are useful for monitoring the effectiveness of
CC   modulating compounds on the expression or activity of the transporter
CC   gene in clinical trials or in a treatment regimen, diagnostic assays for
CC   qualitative changes in transporter nucleic acid expression, and
CC   particularly in quantitative changes that lead to pathology. They are
CC   also used in gene therapy. Genetically engineered host cells comprising
CC   transporter DNA is used to produce non-human transgenic animals. The
CC   present sequence is human transporter protein encoding cDNA. Human
CC   transporter gene is located at chromosome 5
XX
SQ   Sequence 1473 BP; 321 A; 421 C; 341 G; 390 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.26          Length:         1473
Score:                  79.00         Matches:        23
Percent Similarity:     43.3%         Conservative:   6
Best Local Similarity:  34.3%         Mismatches:     36
Query Match:            17.6%         Indels:         2
DB:                     6             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x AAD41820 (1-1473)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                        |||      |||            |||

Db       1289 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1230

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||          ::::::|||      |||  |||      |||      |||  |||

Db       1229 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1176

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::            |||:::      :::      ||||||   |||

Db       1175 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1116

Qy         61 PheSerLeuIleThrTrpAsn 67
                 |||      |||||||||

Db       1115 ATCTCAGCTGGGACGTGGAAC 1095

RESULT 43
ABZ80231/c
ID   ABZ80231 standard; cDNA; 1703 BP.
XX
AC   ABZ80231;
XX
DT   02-JUN-2003  (first entry)
XX
DE   Human tramdorin 2 encoding cDNA SEQ ID NO:4.
XX
KW   Neuroprotective; nootropic; cerebroprotective; analgesic; gene therapy;
KW   central nervous system disorder; CNS disorder; multiple sclerosis;
KW   nerve injury; neuropathic pain; stroke; trauma; non-CNS disorder; tramd;
KW   tramdorin; human; tramdorin 2; chromosome 5; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             255. .1667
FT                   /*tag=  a
FT                   /product= "tramdorin 2"
XX
PN   WO2003016502-A2.
XX
PD   27-FEB-2003.
XX
PF   21-AUG-2002; 2002WO-US026637.
XX
PR   21-AUG-2001; 2001US-0313907P.
PR   21-AUG-2002; 2002US-00225810.
XX
PA   (MCLA-) MCLAUGHLIN RES INST.
XX
PI   Bermingham JR;
XX
DR   WPI; 2003-278567/27.
DR   P-PSDB; ABP96438.
XX
PT   New nucleic acid sequence encoding tramdorins, e.g. mouse tramd 1, mouse
PT   tramd 2, mouse tramd 3, human tramd 1, human tramd 2, human tramd 3 or
PT   rat tramd 1, useful for treating CNS, e.g. stroke, multiple sclerosis,
PT   trauma, neuropathic pain.
XX
PS   Claim 1; Fig 11; 177pp; English.
XX
CC   The present invention describes an isolated nucleic acid sequence
CC   comprising a cDNA sequence encoding mouse tramdorin (tramd) 2, mouse
CC   tramd 3, human tramd 1, human tramd 2, human tramd 3 or rat tramd 1, or
CC   the genomic sequence of mouse tramd 1 or mouse tramd 3. Mouse tramd 1 is
CC   located to chromosome 11, whereas human tramd 1 is located to chromosome
CC   5q31-33. The tramd sequences have neuroprotective, nootropic, analgesic
CC   and cerebroprotective activities, and can be used in gene therapy. The
CC   nucleic acid sequences are useful for diagnosing and treating central
CC   nervous system (CNS) disorders such as multiple sclerosis, nerve injury,
CC   neuropathic pain, stroke or trauma, and non-CNS disorders. The present
CC   sequence encodes human tramd 2, which is given in the exemplification of
CC   the present invention
XX
SQ   Sequence 1703 BP; 371 A; 499 C; 386 G; 447 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:                2.74          Length:          1703
Score:                    79.00         Matches:         23
Percent Similarity:       43.3%         Conservative:    6
Best Local Similarity:    34.3%         Mismatches:      36
Query Match:              17.6%         Indels:          2
DB:                       8             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x ABZ80231 (1-1703)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                        |||     |||           |||
Db       1483 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1424

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||          ::::::|||      |||   |||       |||       |||   |||
Db       1423 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1370

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::             |||:::     :::     ||||||   |||
Db       1369 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1310

Qy         61 PheSerLeuIleThrTrpAsn 67
              |||     |||||||||
Db       1309 ATCTCAGCTGGGACGTGGAAC 1289

RESULT 44
ADR07759/c
ID   ADR07759 standard; cDNA; 2456 BP.
XX
AC   ADR07759;
XX
DT   04-NOV-2004  (first entry)
XX
DE   Full length human cDNA useful for treating neurological disease Seq 1265.
XX
KW   gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
KW   osteoporosis; neurological disease; Alzheimer's disease;
KW   Parkinson's disease; dementia; short memory; cancer;
KW   sense or motor function; emotional reaction; fear response; panic;
KW   osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW   tranquiliser.
XX
OS   Homo sapiens.
XX
PN   EP1447413-A2.
XX
PD   18-AUG-2004.
XX
PF   12-FEB-2004; 2004EP-00003145.
XX
PR   14-FEB-2003; 2003JP-00102207.
PR   09-MAY-2003; 2003JP-00131452.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Yamamoto J,  Nishikawa T,  Isono Y,  Sugiyama T,  Otsuki T;
PI   Wakamatsu A,  Ishii S,  Nagai K,  Irie R;
XX
DR   WPI; 2004-583265/57.
DR   P-PSDB; ADR09715.
XX
PT   New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT   Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS   Claim 1; SEQ ID NO 1265; 2686pp; English.
XX
CC   This invention relates to novel, isolated full length human cDNA
CC   molecules and the encoded proteins thereof. Specifically, it refers to
CC   cDNA clones obtained by an oligo-capping method, where none of these
CC   clones are identical to any known human mRNAs. The present invention
CC   describes an immunoassay to identify agonists and antagonists, as well as
CC   antibodies, antisense molecules and siRNAs that can all be used to bind
CC   to and modulate expression of the cDNA molecules. As such, these
CC   molecules are useful for diagnostic markers or therapeutic targets for
CC   the various diseases or morbid states. In particular, they are useful in
CC   gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC   disease, Parkinson's disease, dementia, short memory and various cancers,
CC   as well as for maintaining equilibrium of sense or motor function, and
CC   for treating emotional reaction, fear response and panic. Accordingly,
CC   they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC   cytostatic and tranquiliser activities. This polynucleotide is a full
CC   length human cDNA sequence of the invention. NOTE: This sequence is not
CC   given in the sequence listing of the specification but can be obtained on
CC   CD-ROM from the European Patent Office, Vienna Sub-office.
XX
SQ   Sequence 2456 BP; 590 A; 677 C; 526 G; 663 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.46          Length:          2456
Score:                    79.00         Matches:         23
Percent Similarity:       43.3%         Conservative:    6
Best Local Similarity:    34.3%         Mismatches:      36
Query Match:              17.6%         Indels:          2
DB:                       13            Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x ADR07759 (1-2456)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                        |||     |||           |||
Db       1641 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 1582

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||          ::::::|||      |||   |||       |||       |||   |||
Db       1581 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 1528

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::             |||:::     :::     ||||||   |||
Db       1527 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 1468

Qy         61 PheSerLeuIleThrTrpAsn 67
              |||     |||||||||
Db       1467 ATCTCAGCTGGGACGTGGAAC 1447

RESULT 45
ADQ67469/c
ID   ADQ67469 standard; cDNA; 3558 BP.
XX
AC   ADQ67469;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Novel human cDNA sequence #2442.
XX
KW   ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW   cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW   neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW   cancer.
XX
OS   Homo sapiens.
XX
PN   EP1440981-A2.
XX
PD   28-JUL-2004.
XX
PF   21-JAN-2004; 2004EP-00001196.
XX
PR   21-JAN-2003; 2003JP-00102206.
PR   09-MAY-2003; 2003JP-00131392.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Sugiyama T,  Otsuki T,  Wakamatsu A,  Sato H,  Ishii S;
PI   Yamamoto J,  Isono Y,  Nagai K,  Irie R;
XX
```

```
DR   WPI; 2004-535376/52.
DR   P-PSDB; ADQ67776.
XX
PT   Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT   Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS   Claim 1; SEQ ID NO 4630; 2449pp; English.
XX
CC   The invention relates to 2495 novel polynucleotides (I) and their encoded
CC   polypeptides, sequences hybridizing to these nucleotides, sequences
CC   encoding partial polypeptides and sequences having 70% or 90% identity to
CC   the nucleotide and protein sequences. The nucleotides and polypeptides
CC   are useful as diagnostic markers or therapeutic target for the diseases
CC   or morbid states. They are also useful for treating osteoporosis,
CC   neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC   dementia and various cancers. This sequence corresponds to a nucleotide
CC   sequence of the invention.
XX
SQ   Sequence 3558 BP; 851 A; 941 C; 740 G; 1026 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              7.3          Length:        3558
Score:                  79.00        Matches:       23
Percent Similarity:     43.3%        Conservative:  6
Best Local Similarity:  34.3%        Mismatches:    36
Query Match:            17.6%        Indels:        2
DB:                     12           Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x ADQ67469 (1-3558)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||   ||||||                        |||       |||          |||
Db       3129 ATGATGAGAGCCAGGGCGCTGCTGCTCACGGAGCCTACCAGGGAGATGACCAAGTCCAGG 3070

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||          ::::::|||      |||   |||      |||        |||   |||
Db       3069 CGGGGGATGAGGATGGCTGAG------ACACAGGTTAGACAGACCAAGGCTGAGCGGACA 3016

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||   :::|||:::            |||:::     :::      ||||||   |||
Db       3015 GACAGGTCTACAAACAGTGCCCAGCTCTCTGACACTTGGGAGATGGCAAACGGGATGATG 2956

Qy         61 PheSerLeuIleThrTrpAsn 67
                |||      |||||||||
Db       2955 ATCTCAGCTGGGACGTGGAAC 2935

RESULT 46
ABQ55431
ID   ABQ55431 standard; cDNA; 480 BP.
XX
AC   ABQ55431;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human ovarian antigen HNOJL59 cDNA, SEQ ID NO:1311.
XX
KW   Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW   ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW   infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200200677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US018569.
XX
PR   07-JUN-2000; 2000US-0209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   P-PSDB; ABP42354.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT   useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT   cancer), immune disorders, cardiovascular disorders and neurological
PT   diseases.
XX
PS   Claim 1; SEQ ID NO 1311; 2922pp; English.
XX
CC   The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC   ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC   encompasses polypeptides 90% identical and polynucleotides 95% identical
CC   to the sequences of the invention. The invention additionally relates to
CC   recombinant vectors and host cells comprising human ovarian antigen
CC   polynucleotides, antibodies against human ovarian antigens, and the use
CC   of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC   treating, prognosing or preventing various ovary and/or breast-related
CC   disorders. Such conditions include ovarian cancer and breast cancer, and
CC   metastatic tumours of ovarian or breast origin, reproductive system
CC   disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC   polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC   disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC   shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC   vaginitis), immune disorders (e.g., congenital and acquired
CC   immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC   blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC   respiratory disorders, neurological disorders, gastrointestinal disorders
CC   and urinary system disorders. Ovarian antigen polypeptides and
CC   polynucleotides may also be used in screening for compounds which
CC   modulate ovarian antigen expression or activity. The polynucleotides may
CC   further be used for gene therapy, chromosome mapping, in the
CC   identification of individuals and in forensic analysis, and the
CC   polypeptides may be used as food additives or to prepare antibodies
CC   useful in disease diagnosis, drug targeting and phenotyping. The present
CC   sequence represents cDNA encoding a human ovarian antigen of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 480 BP; 92 A; 156 C; 107 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0.692        Length:        480
Score:                  78.00        Matches:       26
Percent Similarity:     48.5%        Conservative:  7
Best Local Similarity:  38.2%        Mismatches:    32
Query Match:            17.4%        Indels:        3
DB:                     6            Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x ABQ55431 (1-480)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db        113 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 172

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
                :::      |||   |||   |||       |||:::|||          |||||||||
Db        173 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 232

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||         |||   ||||||:::|||:::
Db        233 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 292

Qy         64 eThrTrpAsnIleAspGlyLeu 71
```

```
                    |||   :::      |||
Db        293 TGTCTGGACTTTGAGCAGGCTT 314
```

RESULT 47
ADA24517
ID ADA24517 standard; cDNA; 1142 BP.
XX
AC ADA24517;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human cDNA differential expressed in adipose tissue, INCYTE1453266.4.
XX
KW ss; differential expression; adipose tissue; cytostatic; hypotensive;
KW antiarteriosclerotic; antidiabetic; anorectic; gene therapy;
KW peroxisome proliferator-activated receptor gamma; PPARgamma;
KW diabetes mellitus; obesity; hypertension; atherosclerosis; breast cancer;
KW prostate cancer; colon cancer; polycystic ovarian syndrome.
XX
OS Homo sapiens.
XX
PN US2003096272-A1.
XX
PD 22-MAY-2003.
XX
PF 29-JUL-2002; 2002US-00208408.
XX
PR 30-JUL-2001; 2001US-0308868P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Schebye XM;
XX
DR WPI; 2003-606416/57.
XX
PT New combination comprising several cDNAs, useful for preparing a
PT composition for diagnosing or treating diabetes mellitus, obesity,
PT hypertension, atherosclerosis, or cancer of the breast, prostate or
PT colon.
XX
PS Claim 1; Page 61-62; 84pp; English.
XX
CC Then invention relates to a new combination comprising 55 cDNAs (ADA24485
CC -ADA24539)or their complements that are differentially regulated in an
CC adipose sample. Also included are detecting differential expression of
CC one or more cDNAs in a sample containing nucleic acids, screening several
CC molecules or compounds to identify a ligand that specifically binds a
CC cDNA, a vector comprising the cDNA, a host cell containing the vector,
CC producing a protein, screening several molecules or compounds, producing
CC an antibody and the isolated antibody. The cDNAs comprise sequences which
CC are upregulated or downregulated in response to peroxisome proliferator-
CC activated receptor gamma (PPARgamma) agonist. The combination comprising
CC several cDNAs is useful for preparing a composition for diagnosing or
CC treating diabetes mellitus, obesity, hypertension, atherosclerosis,
CC cancer of the breast, prostate or colon, or polycystic ovarian syndrome.
XX
SQ Sequence 1142 BP; 274 A; 309 C; 297 G; 262 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.2	Length:	1142
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	9	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x ADA24517 (1-1142)

```
Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::    |||       |||
Db        836 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 895

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
```

```
                :::      |||   |||   |||      |||:::|||          |||||||||
Db        896 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 955

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||          |||   ||||||:::|||:::
Db        956 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1015

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                |||   :::      |||
Db       1016 TGTCTGGACTTTGAGCAGGCTT 1037
```

RESULT 48
ABZ77286
ID ABZ77286 standard; cDNA; 3672 BP.
XX
AC ABZ77286;
XX
DT 28-MAY-2003 (first entry)
XX
DE Nucleotide sequence of human IGFBP5 polypeptide.
XX
KW Human; angiogenesis; wound healing; retinopathy; ischemia; inflammation;
KW microvasculopathy; bone healing; skin inflammation; HOG3; HOG8; HOG18;
KW follicular development; CA9; HXB; IGFBP5; HFARP; STC1; mig-6; SSR4;
KW cancer; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003010205-A1.
XX
PD 06-FEB-2003.
XX
PF 26-JUL-2002; 2002WO-US023786.
XX
PR 26-JUL-2001; 2001US-0307600P.
PR 24-JUL-2002; 2002US-00201642.
XX
PA (UYDU-) UNIV DUKE MEDICAL CENT.
XX
PI Riggins GJ, Lal A;
XX
DR WPI; 2003-239423/23.
XX
PT Inhibiting angiogenesis for treating wound healing, retinopathy,
PT ischemia, inflammation, microvasculopathy, bone healing, skin
PT inflammation or follicular development by providing to a subject an
PT antisense polynucleotide.
XX
PS Claim 1; Page 56-57; 66pp; English.
XX
CC The present sequence encodes an IGFBP5 polypeptide. It is used in the
CC method of the invention. The specification describes a method modulating
CC angiogenesis associated with wound healing, retinopathy, ischemia,
CC inflammation, microvasculopathy, bone healing, skin inflammation or
CC follicular development. The method comprises providing to a subject HOG3,
CC HOG8, HOG18, CA9, HXB, IGFBP5, HFARP, STC1, mig-6 or SSR4. The methods,
CC antisense polynucleotides, polypeptides and antibodies are useful for
CC treating wound healing, retinopathy, ischemia, inflammation,
CC microvasculopathy, bone healing, skin inflammation or follicular
CC development, or cancer such as breast, colon or lung cancer, or
CC glioblastoma
XX
SQ Sequence 3672 BP; 897 A; 975 C; 906 G; 894 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	10.4	Length:	3672
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	10	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x ABZ77286 (1-3672)

```
Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::    |||      |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
              :::      |||   |||   |||      |||:::|||          |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||          |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                |||   :::      |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 49
ADR24529
ID   ADR24529 standard; DNA; 3672 BP.
XX
AC   ADR24529;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Breast cancer prognosis marker #390.
XX
KW   ds; breast cancer; prognosis; gene expression; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004065545-A2.
XX
PD   05-AUG-2004.
XX
PF   15-JAN-2004; 2004WO-US001100.
XX
PR   15-JAN-2003; 2003US-00342887.
XX
PA   (ROSE-) ROSETTA INPHARMATICS LLC.
PA   (NECA-) NETHERLANDS CANCER INST.
XX
PI   Van't Veer LJ,  He Y;
XX
DR   WPI; 2004-593473/57.
XX
PT   Classifying a breast cancer patient according to prognosis comprises
PT   determining the similarity between the level of expression of each of
PT   five genes in a cell sample taken from patient, to control levels.
XX
PS   Disclosure; SEQ ID NO 390; 226pp; English.
XX
CC   The invention relates to a method of classifying a breast cancer patient
CC   according to prognosis by determining the similarity between the level of
CC   expression of each of five genes for which markers are listed in the
CC   specification, in a cell sample taken from the breast cancer patient, to
CC   control levels of expression for each respective five genes to obtain a
CC   patient similarity value. The methods are useful for classifying a breast
CC   cancer patient according to prognosis. Kits and computer program products
CC   are useful for data analysis using the diagnostic, prognostic and
CC   statistical methods of the invention. This sequence corresponds to a
CC   marker used in the method of the invention.
XX
SQ   Sequence 3672 BP; 897 A; 975 C; 906 G; 894 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              10.4          Length:         3672
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     13            Gaps:           1
```

```
US-10-757-745-2_COPY_54_140 (1-87) x ADR24529 (1-3672)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::    |||      |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
              :::      |||   |||   |||      |||:::|||          |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||          |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                |||   :::      |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 50
ADY61760
ID   ADY61760 standard; DNA; 3672 BP.
XX
AC   ADY61760;
XX
DT   19-MAY-2005  (first entry)
XX
DE   Human gene marker indicating poor prognosis for breast cancer Seq 76.
XX
KW   gene; ds; prognosis; breast cancer; diagnosis; RNA interference;
KW   screening; gene therapy; cytostatic.
XX
OS   Homo sapiens.
XX
PN   US2005054826-A1.
XX
PD   10-MAR-2005.
XX
PF   18-MAY-2004; 2004US-00848755.
XX
PR   19-MAY-2003; 2003US-0471842P.
XX
PA   (ROSE-) ROSETTA INPHARMATICS LLC.
XX
PI   Mao M;
XX
DR   WPI; 2005-213102/22.
XX
PT   New purified diaphanous-3 (DIAPH3) peptides and nucleic acids, useful as
PT   targets for screening assays designed to identify inhibitors or
PT   modulators of DIAPH3 activity, which are useful for treating breast
PT   cancer.
XX
PS   Disclosure; SEQ ID NO 76; 54pp; English.
XX
CC   This invention relates to a novel human diaphanous-3 (DIAPH3) protein and
CC   the encoding polynucleotide thereof. Specifically, it refers to the use
CC   of DIAPH3 cDNA in the prognosis of breast cancer, such that the DIAPH3
CC   protein can be used as a target for anti-cancer drugs or for the
CC   identification of molecules with anti-cancer activity. The present
CC   invention provides antibodies that specifically bind to the DIAPH3
CC   protein, as well as methods to identify an agent that modulates the
CC   binding of DIAPH3 to a binding partner. In particular it describes
CC   measuring the expression level of DIAPH3 using hybridization oligos and
CC   comparing this with a control expression level, where a raised expression
CC   level indicates a poor prognosis with respect to breast cancer. As such,
CC   it also provides small interfering RNA oligos that can hybridize to the
CC   DIAPH3 gene and inhibit expression thereof. The peptide and nucleic acid
CC   sequences of this invention are useful as markers to differentiate
CC   patients with a good prognosis from those with poor prognosis for breast
CC   cancer and are useful as targets for screening assays designed to
CC   identify inhibitors or modulators of DIAPH3 activity, which in turn are
CC   used for treating breast cancer. Accordingly, they can be used in gene
CC   therapy and exhibit cytostatic activity. This polynucleotide sequence is
```

```
CC   a human gene marker that when expression is upregulated in a patient
CC   sample it is correlated with a poor prognosis for breast cancer, given in
CC   an exemplification of the invention.
XX
SQ   Sequence 3672 BP; 897 A; 975 C; 906 G; 894 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                10.4          Length:          3672
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       14            Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x ADY61760 (1-3672)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||      |||
Db       1760 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||  |||   |||      |||:::|||           |||||||||
Db       1820 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1879

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||         |||   ||||||:::|||:::
Db       1880 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1939

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                |||   :::      |||
Db       1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 51
ACN37824
ID   ACN37824 standard; cDNA; 4180 BP.
XX
AC   ACN37824;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Tumour-associated antigenic target (TAT) cDNA DNA324220, SEQ ID NO:960.
XX
KW   Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW   tumour; diagnosis; cell proliferative disorder; breast cancer;
KW   colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW   central nervous system cancer; bladder cancer; pancreatic cancer;
KW   cervical cancer; melanoma; leukaemia; hybridisation probe;
KW   chromosome identification; chromosome mapping; gene mapping;
KW   gene therapy; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004030615-A2.
XX
PD   15-APR-2004.
XX
PF   29-SEP-2003; 2003WO-US028547.
XX
PR   02-OCT-2002; 2002US-0414971P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wu TD,  Zhang Z,  Zhou Y;
XX
DR   WPI; 2004-347921/32.
DR   P-PSDB; ABM80380.
XX
PT   New tumor-associated antigenic target polypeptides and nucleic acids,
PT   useful in preparing a medicament for treating or detecting a
PT   proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT   prostate cancer or tumor.
XX
PS   Claim 1; SEQ ID NO 960; 7273pp; English.
XX
CC   The invention relates to human tumour-associated antigenic target (TAT)
CC   polypeptides, and their related nucleic acids. The TAT polypeptides are
CC   overexpressed in cancer tissues compared to normal tissues, and may thus
CC   serve as effective targets for the diagnosis and treatment of cancer in
CC   mammals. The invention also relates to nucleic acid and polypeptide
CC   sequences at least 80% identical to the TAT nucleic acids and
CC   polypeptides; expression vectors and host cells comprising a TAT nucleic
CC   acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC   molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC   TAT polypeptide; and methods and compositions for the treatment or
CC   diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC   antibodies, antagonists, binding molecules and compositions are useful
CC   for diagnosing or treating a cell proliferative disorder associated with
CC   increased TAT expression, particularly cancers such as breast cancer,
CC   colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC   cancer, pancreatic cancer, cervical cancer, cancers of the central
CC   nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC   used as hybridisation probes, in chromosome and gene mapping, in
CC   chromosome identification and in gene therapy. The present sequence
CC   represents a TAT nucleic acid of the invention
XX
SQ   Sequence 4180 BP; 1046 A; 1077 C; 1061 G; 996 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                12.4          Length:          4180
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       13            Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x ACN37824 (1-4180)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||      |||
Db       2277 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 2336

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||  |||   |||      |||:::|||           |||||||||
Db       2337 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 2396

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||         |||   ||||||:::|||:::
Db       2397 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 2456

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                |||   :::      |||
Db       2457 TGTCTGGACTTTGAGCAGGCTT 2478

RESULT 52
AAL60051
ID   AAL60051 standard; cDNA; 4194 BP.
XX
AC   AAL60051;
XX
DT   27-AUG-2003  (first entry)
XX
DE   Human Pc011 cDNA.
XX
KW   Human; differentially regulated protein; prevention; therapy; vaccine;
KW   prostate cancer; gene therapy; Pc011; chromosome 2q22-q23; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1588. .2019
FT                   /*tag=  a
FT                   /product= "Human Pc011 protein"
XX
PN   WO2003040331-A2.
XX
PD   15-MAY-2003.
```

XX
PF 07-NOV-2002; 2002WO-US035563.
XX
PR 07-NOV-2001; 2001US-0331041P.
PR 07-NOV-2001; 2001US-0331042P.
PR 18-DEC-2001; 2001US-0340251P.
PR 07-JAN-2002; 2002US-0344791P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Jay G, Kovacs KF, Fan W;
XX
DR WPI; 2003-449451/42.
DR P-PSDB; AAO29557.
XX
PT New polynucleotide for diagnosing, staging, monitoring, prognosticating,
PT preventing or treating, or determining the predisposition to, diseases or
PT conditions such as prostate cancer, and for research or forensic science.
XX
PS Claim 29; Page 106-108; 100pp; English.
XX
CC The present invention relates to novel differentially regulated genes and
CC polypeptides encoded by them. Sequences of the invention are useful in
CC diagnosing, staging, monitoring, prognosticating, preventing, treating or
CC determining the predisposition to diseases or conditions such as prostate
CC cancer. They may be used as molecular markers, drug targets, vaccines, in
CC gene therapy, research, clinical medicine or forensic science. The
CC present sequence is a differentially regulated prostate cDNA, Pc011 used
CC in the invention. Pc011 gene is located on chromosome 2q22-q23
XX
SQ Sequence 4194 BP; 1073 A; 1076 C; 1056 G; 989 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	12.4	Length:	4194
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	10	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x AAL60051 (1-4194)

```
Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db       2254 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 2313

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||         |||||||||
Db       2314 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 2373

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                 |||   ||||||   |||          |||   ||||||:::|||:::
Db       2374 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 2433

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                ||||   :::      |||
Db       2434 TGTCTGGACTTTGAGCAGGCTT 2455
```

RESULT 53
AAZ41297
ID AAZ41297 standard; cDNA; 4750 BP.
XX
AC AAZ41297;
XX
DT 18-JAN-2000 (first entry)
XX
DE Human normal ovarian tissue derived cDNA 76.
XX
KW Human; ovary; screening; ovarian cancer; treatment; ss.
XX
OS Homo sapiens.
XX
PN DE19816395-A1.
XX
PD 07-OCT-1999.
XX
PF 03-APR-1998; 98DE-01016395.
XX
PR 03-APR-1998; 98DE-01016395.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR WPI; 1999-552352/47.
XX
PT Nucleic acid sequences potentially useful in diagnosis or therapy of
PT ovarian cancer.
XX
PS Claim 3; Page 186-187; 274pp; German.
XX
CC This invention describes novel nucleic acid sequences that are highly
CC expressed in normal ovary tissue. Artificial chromosomes and cosmid
CC clones containing the sequences can be used as gene transfer vehicles.
CC The sequences can be used to produce DNA fragments containing full-length
CC genes. Host cells transformed with the sequences can be used to produce
CC polypeptides or polypeptide fragments, which can be used to screen phage
CC displays for polypeptides that bind to them, or as tools for identifying
CC agents active against ovarian cancer, or to prepare medicaments for
CC treating ovarian cancer. The cDNA sequences can be used to obtain genomic
CC genes, their promoters, enhancers, silencers, exon structures, intron
CC structures and their splice variants. AAZ41222-Z41324 represent cDNA
CC sequences derived from normal human ovarian tissue and which encode the
CC protein fragments represented in AAY59724-Y59837
XX
SQ Sequence 4750 BP; 1237 A; 1203 C; 1186 G; 1124 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	14.7	Length:	4750
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	2	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x AAZ41297 (1-4750)

```
Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db       2521 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 2580

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||         |||||||||
Db       2581 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 2640

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                 |||   ||||||   |||          |||   ||||||:::|||:::
Db       2641 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 2700

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                ||||   :::      |||
Db       2701 TGTCTGGACTTTGAGCAGGCTT 2722
```

RESULT 54
ABX77514
ID ABX77514 standard; cDNA; 5511 BP.
XX
AC ABX77514;
XX
DT 09-APR-2003 (first entry)
XX
DE Differentially expressed breast cancer associated cDNA #9.
XX
KW Breast cancer; differential gene expression; BC-cDNA;
KW breast cancer diagnosis; breast cancer monitoring;
KW breast cancer treatment; breast cancer staging; gene; ss.

XX
OS Homo sapiens.
XX
PN US2002156263-A1.
XX
PD 24-OCT-2002.
XX
PF 04-OCT-2001; 2001US-00974298.
XX
PR 05-OCT-2000; 2000US-0238331P.
XX
PA (CHEN/) CHEN H.
XX
PI Chen H;
XX
DR WPI; 2003-182653/18.
XX
PT New cDNAs, which are differentially expressed in (metastatic) breast
PT cancer useful for diagnosing or staging, breast cancer, or for monitoring
PT the treatment of breast cancer in an individual.
XX
PS Claim 1; SEQ ID NO 13; 30pp; English.
XX
CC The invention describes a combination of cDNAs (designated BC-cDNAs),
CC which are differentially expressed in breast cancer. The combination
CC includes 152 cDNA sequences, or their complements. The protein encoded by
CC any of these BC-cDNAs is useful for screening several molecules or
CC compounds to identify at least one ligand that specifically binds the
CC protein, producing or preparing polyclonal or monoclonal antibodies, or
CC purifying antibodies from a sample. The antibodies, which specifically
CC bind the protein differentially expressed in breast cancer is useful for
CC detecting the expression of a protein in a sample. The BC-cDNAs are also
CC useful for diagnosing, monitoring the treatment of, or staging, breast
CC cancer. This sequence represents a differentially expressed breast cancer
CC associated cDNA. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=20020156263
XX
SQ Sequence 5511 BP; 1320 A; 1529 C; 1428 G; 1233 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	17.9	Length:	5511
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	8	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x ABX77514 (1-5511)

```
Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::   |||      |||
Db       4092 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4151

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::     |||  |||   |||      |||:::|||          |||||||||
Db       4152 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4211

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||         |||   ||||||:::|||:::
Db       4212 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4271

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                |||   :::     |||
Db       4272 TGTCTGGACTTTGAGCAGGCTT 4293
```

RESULT 55
ABX63285
ID ABX63285 standard; cDNA; 5511 BP.
XX
AC ABX63285;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human cDNA #285 differentially expressed in activated vascular tissue.
XX
KW Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiant;
KW hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;
KW gene therapy; vascular disease; cancer; coronary; artery disease;
KW hypertension; diabetes; pre-eclampsia; restenosis;
KW ischaemia-reperfusion injury; stroke.
XX
OS Homo sapiens.
XX
PN US2002137081-A1.
XX
PD 26-SEP-2002.
XX
PF 08-JAN-2002; 2002US-00044090.
XX
PR 28-JUL-2000; 2000US-0222469P.
PR 08-JAN-2001; 2001US-0260483P.
XX
PA (BAND/) BANDMAN O.
XX
PI Bandman O;
XX
DR WPI; 2003-110597/10.
XX
PT Combination for diagnosing, staging, treating, or monitoring the
PT progression of treatment of a vascular disease, e.g. atherosclerosis,
PT comprises several cDNAs that are differentially expressed in activated
PT vascular tissue.
XX
PS Claim 1; Page; 18pp; English.
XX
CC This invention relates to a combination comprising several cDNAs that are
CC differentially expressed in activated vascular tissue. The invention also
CC discloses a high throughput method for detecting differentially expressed
CC cDNAs in a sample. The cDNAs of the invention may have
CC antiarteriosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;
CC gynaecological; vasotropic and cerebroprotective activities and may be
CC used in gene therapy. The cDNAs of the invention may be used in a high-
CC throughput methods for detecting differential expression of one or more
CC cDNAs in a sample, or screening several molecules or compounds to
CC identify a molecule or compound that specifically binds a cDNA of the
CC invention. A protein encoded by the cDNA may be used to screen several
CC molecules or compounds to identify a ligand that specifically binds to
CC the protein, or to produce or purify an antibody to the protein that can
CC be used to detect a protein in a sample or purify a natural or
CC recombinant protein from a sample. The nucleotides may be useful for
CC diagnosing, staging, treating, or monitoring the progression of treatment
CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery
CC disease, hypertension, diabetes, pre-eclampsia, ischaemia- reperfusion
CC injury, restenosis, or stroke. The cDNAs can also be used for large-scale
CC genetic or gene expression analysis of several new nucleic acid
CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for
CC diagnosing pre-pathologic disorders, and chronic or acute diseases
CC associated with abnormalities in the expression, amount or distribution
CC of the protein. The present sequence represents a cDNA of the invention
CC that is differentially expressed in activated vascular tissue. Note: The
CC sequence data for this patent did not form part of the specification, but
CC was obtained in electronic format directly from USPTO at
CC http.seqdata.uspto.gov/sequence.html?DocID=20020137081
XX
SQ Sequence 5511 BP; 1320 A; 1529 C; 1428 G; 1233 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	17.9	Length:	5511
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	8	Gaps:	1

```
US-10-757-745-2_COPY_54_140 (1-87) x ABX63285 (1-5511)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::   |||       |||
Db       4092 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4151

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||          |||||||||
Db       4152 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4211

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||         |||   ||||||:::|||:::
Db       4212 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4271

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       4272 TGTCTGGACTTTGAGCAGGCTT 4293

RESULT 56
ACH04033
ID   ACH04033 standard; cDNA; 5511 BP.
XX
AC   ACH04033;
XX
DT   26-SEP-2003  (first entry)
XX
DE   Human cDNA differentially expressed in lung cancer #238.
XX
KW   Gene therapy; emphysema; ss; gene; chronic obstructive pulmonary disease;
KW   respiratory disorder; lung cancer; asthma; human.
XX
OS   Homo sapiens.
XX
PN   US2003065157-A1.
XX
PD   03-APR-2003.
XX
PF   04-APR-2002; 2002US-00116802.
XX
PR   04-APR-2001; 2001US-0281593P.
XX
PA   (LASE/) LASEK A W.
XX
PI   Lasek AW;
XX
DR   WPI; 2003-540803/51.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   respiratory disorders, useful for diagnosing or treating respiratory
PT   disorders e.g., lung cancer, chronic obstructive pulmonary disease,
PT   emphysema or asthma.
XX
PS   Claim 1; Page; 39pp; English.
XX
CC   The invention relates to a combination comprising cDNAs or their
CC   complements that are differentially expressed in respiratory disorder.
CC   The combination is useful for preparing a composition for diagnosing or
CC   treating respiratory disorders e.g. lung cancer, chronic obstructive
CC   pulmonary disease, emphysema or asthma. The present sequence represents
CC   human cDNA differentially expressed during lung cancer
XX
SQ   Sequence 5511 BP; 1320 A; 1529 C; 1428 G; 1233 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              17.9          Length:         5511
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     9             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x ACH04033 (1-5511)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::   |||       |||
Db       4092 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4151

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||          |||||||||
Db       4152 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4211

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||         |||   ||||||:::|||:::
Db       4212 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4271

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       4272 TGTCTGGACTTTGAGCAGGCTT 4293

RESULT 57
ADJ56346
ID   ADJ56346 standard; cDNA; 5511 BP.
XX
AC   ADJ56346;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human cDNA differentially expressed in MYCN activated cells SeqID 152.
XX
KW   human; differential expression; transactivator; proto-oncogene;
KW   neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
KW   MYCN activated cell.
XX
OS   Homo sapiens.
XX
PN   US2003119009-A1.
XX
PD   26-JUN-2003.
XX
PF   25-FEB-2002; 2002US-00084817.
XX
PR   23-FEB-2001; 2001US-0270784P.
XX
PA   (STUA/) STUART S G.
PA   (NUCH/) NUCHTERN J G.
PA   (PLON/) PLON S E.
PA   (SHOH/) SHOHET J M.
XX
PI   Stuart SG,  Nuchtern JG,  Plon SE,  Shohet JM;
XX
DR   WPI; 2003-635698/60.
XX
PT   New genes regulated by MYCN activation, useful in gene therapy,
PT   particularly for treating a subject with e.g. neuroblastoma or other
PT   cancers, or for diagnosing, staging or monitoring the treatment of the
PT   cancer.
XX
PS   Claim 1; SEQ ID NO 152; 27pp; English.
XX
CC   This invention relates to novel isolated cDNAs that are differentially
CC   expressed in MYCN activated cells. Specifically, it refers to
CC   polynucleotide sequences that exhibit differential expression patterns in
CC   cells activated by the transactivator MYCN, where MYCN is a proto-
CC   oncogene that is amplified in neuroblastoma cells and is common in small
CC   cell lung cancers. The present invention describes these cDNA molecules
CC   as useful for in hybridisation assays to detect expression of nucleic
CC   acids (or complementary nucleic acids) in a present in a given sample, as
CC   well as for screening assays by identifying molecules or compounds that
CC   specifically bind the cDNA as a ligand and modulate function or activity.
CC   Accordingly, these compositions exhibit cytostatic activity and can also
CC   be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC   that is differentially expressed in MYCN activated cells, given in an
CC   exemplification of the invention. NOTE: This sequence does not appear in
CC   the printed specification but has been obtained in electronic format from
CC   the US Patent Office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
```

```
XX
SQ   Sequence 5511 BP; 1320 A; 1529 C; 1428 G; 1233 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                17.9          Length:          5511
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       10            Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x ADJ56346 (1-5511)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::   |||       |||
Db       4092 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4151

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
              :::      |||   |||   |||      |||:::|||         |||||||||
Db       4152 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4211

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||         |||   ||||||:::|||:::
Db       4212 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4271

Qy         64 eThrTrpAsnIleAspGlyLeu 71
              |||   :::      |||
Db       4272 TGTCTGGACTTTGAGCAGGCTT 4293

RESULT 58
ADF90711/c
ID   ADF90711 standard; DNA; 6115 BP.
XX
AC   ADF90711;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Human hepatic-fibrosis disease marker SEQ ID 173.
XX
KW   Hepatic fibrosis; marker; chronic hepatitis; liver cirrhosis;
KW   hepatic carcinoma; human; ds.
XX
OS   Homo sapiens.
XX
PN   JP2003259877-A.
XX
PD   16-SEP-2003.
XX
PF   11-MAR-2002; 2002JP-00065013.
XX
PR   11-MAR-2002; 2002JP-00065013.
XX
PA   (SUMU ) SUMITOMO SEIYAKU KK.
XX
DR   WPI; 2003-821598/77.
XX
PT   Hepatic fibrosis disease markers comprising polynucleotides or
PT   antibodies, useful for improved diagnosis, screening and developing drugs
PT   to treat hepatitis, to control cirrhosis and carcinoma.
XX
PS   Claim 1; SEQ ID NO 173; 313pp; Japanese.
XX
CC   The present invention relates to hepatic-fibrosis disease markers
CC   (ADF90539-ADF90871) and related proteins (ADF90872-ADF90917). The
CC   sequences are useful for detecting and treating hepatic fibrosis caused
CC   by alcohol consumption, virus infection, etc., and the associated chronic
CC   hepatitis, etc. leading to liver cirrhosis and hepatic carcinoma. The
CC   markers allow the cause of hepatic fibrosis to be clarified (diagnostic
CC   precision), so more suitable treatments can be developed and given.
XX
SQ   Sequence 6115 BP; 1454 A; 1561 C; 1568 G; 1484 T; 0 U; 48 Other;

Alignment Scores:

Pred. No.:                20.5          Length:          6115
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       10            Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x ADF90711 (1-6115)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::   |||       |||
Db       2236 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 2177

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
              :::      |||   |||   |||      |||:::|||         |||||||||
Db       2176 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 2117

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||         |||   ||||||:::|||:::
Db       2116 CCCCCACATTCCTCCCAGCCTGGACCTGATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 2057

Qy         64 eThrTrpAsnIleAspGlyLeu 71
              |||   :::      |||
Db       2056 TGTCTGGACTTTGAGCAGGCTT 2035

RESULT 59
ADP07274
ID   ADP07274 standard; DNA; 6316 BP.
XX
AC   ADP07274;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human IGFBP5 DNA.
XX
KW   ds; proliferative disease; breast; methylation; CpG; bisulfite; human.
XX
OS   Homo sapiens.
XX
PN   DE10255104-A1.
XX
PD   11-MAR-2004.
XX
PF   26-NOV-2002; 2002DE-01055104.
XX
PR   27-AUG-2002; 2002DE-01039313.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Maier S;
XX
DR   WPI; 2004-284340/27.
DR   GENBANK; NM_000599.
XX
PT   Analyzing proliferative diseases of breast cells, useful e.g. for
PT   diagnosis, prognosis and treatment of breast cancer, by determining
PT   methylation status of specific genes.
XX
PS   Claim 1; Page; 22pp; German.
XX
CC   This invention describes a novel method of analysing proliferative
CC   diseases of breast cells by determining the methylation status of certain
CC   genes. The invention also describes nucleic acids or their complements,
CC   oligomers, especially oligonucleotides or peptide nucleic acid oligomers,
CC   that hybridise to, or are identical with, any of the nucleic acids, the
CC   preparation of an oligomer array for analysing proliferative diseases of
CC   breast cells that are associated with the methylation status of CpG
CC   dinucleotides of the genes by bonding at least one oligomer to a solid
CC   phase and a kit comprising a bisulfite reagent and the oligomer. The
CC   nucleic acids are genomic sequences (5'- and/or regulatory and/or CpG-
CC   rich regions). The base sequence of the oligomer includes at least one
CC   CpG island, especially with C in the middle third of the sequence. The
CC   process involves a genomic DNA sample treated chemically, specifically
```

```
CC   with a bisulfite reagent, to convert non-methylated C to uracil or some
CC   other base having base-pairing properties different from those of C.
CC   Fragments of the treated DNA are then amplified, using primers and a
CC   polymerase and the methylation status of the genomic CpG dinucleotides is
CC   determined by analysis of the amplicons, particularly by hybridisation to
CC   the oligomer, optionally with extension of the hybridised oligomer by at
CC   least one base, or detection is by sequencing. The amplification may use
CC   methylation-specific primers. Alternatively, genomic DNA is extracted
CC   from a sample and digested with methylation-specific restriction enzymes,
CC   then the digestion fragments detected, optionally after amplification. In
CC   either method, more than 10 fragments of 100-200 bp are amplified in a
CC   single reaction vessel, using a heat-resistant DNA polymerase in PCR. The
CC   amplicons carry detectable markers, e.g. fluorophores, radioisotopes
CC   and/or releasable fragments of known mass that can be detected by mass
CC   spectrometry. The method is used for characterisation, classification,
CC   differentiation, staging, phase-estimation, diagnosis and/or therapy of
CC   proliferative diseases of breast cells. The method provides very specific
CC   classification of proliferative diseases, allowing better treatment. It
CC   can both characterise methylation status and detect single-nucleotide
CC   polymorphisms. This sequence represents human gene used to illustrate the
CC   method of the invention. NOTE: This sequence does not appear in the
CC   printed specification but has been retrieved from Genbank.
XX
SQ   Sequence 6316 BP; 1611 A; 1680 C; 1554 G; 1471 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              21.4          Length:         6316
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     12            Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x ADP07274 (1-6316)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::   |||      |||
Db       4103 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4162

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||         |||||||||
Db       4163 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4222

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||         |||   ||||||:::|||:::
Db       4223 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4282

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       4283 TGTCTGGACTTTGAGCAGGCTT 4304

RESULT 60
ADP07323
ID   ADP07323 standard; DNA; 6316 BP.
XX
AC   ADP07323;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human IGFBP5 DNA #2.
XX
KW   ds; proliferative disease; breast; methylation; CpG; bisulfite; human.
XX
OS   Homo sapiens.
XX
PN   DE10255104-A1.
XX
PD   11-MAR-2004.
XX
PF   26-NOV-2002; 2002DE-01055104.
XX
PR   27-AUG-2002; 2002DE-01039313.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Maier S;
XX
DR   WPI; 2004-284340/27.
DR   GENBANK; NM_000599.
XX
PT   Analyzing proliferative diseases of breast cells, useful e.g. for
PT   diagnosis, prognosis and treatment of breast cancer, by determining
PT   methylation status of specific genes.
XX
PS   Claim 1; Page; 22pp; German.
XX
CC   This invention describes a novel method of analysing proliferative
CC   diseases of breast cells by determining the methylation status of certain
CC   genes. The invention also describes nucleic acids or their complements,
CC   oligomers, especially oligonucleotides or peptide nucleic acid oligomers,
CC   that hybridise to, or are identical with, any of the nucleic acids, the
CC   preparation of an oligomer array for analysing proliferative diseases of
CC   breast cells that are associated with the methylation status of CpG
CC   dinucleotides of the genes by bonding at least one oligomer to a solid
CC   phase and a kit comprising a bisulfite reagent and the oligomer. The
CC   nucleic acids are genomic sequences (5'- and/or regulatory and/or CpG-
CC   rich regions). The base sequence of the oligomer includes at least one
CC   CpG island, especially with C in the middle third of the sequence. The
CC   process involves a genomic DNA sample treated chemically, specifically
CC   with a bisulfite reagent, to convert non-methylated C to uracil or some
CC   other base having base-pairing properties different from those of C.
CC   Fragments of the treated DNA are then amplified, using primers and a
CC   polymerase and the methylation status of the genomic CpG dinucleotides is
CC   determined by analysis of the amplicons, particularly by hybridisation to
CC   the oligomer, optionally with extension of the hybridised oligomer by at
CC   least one base, or detection is by sequencing. The amplification may use
CC   methylation-specific primers. Alternatively, genomic DNA is extracted
CC   from a sample and digested with methylation-specific restriction enzymes,
CC   then the digestion fragments detected, optionally after amplification. In
CC   either method, more than 10 fragments of 100-200 bp are amplified in a
CC   single reaction vessel, using a heat-resistant DNA polymerase in PCR. The
CC   amplicons carry detectable markers, e.g. fluorophores, radioisotopes
CC   and/or releasable fragments of known mass that can be detected by mass
CC   spectrometry. The method is used for characterisation, classification,
CC   differentiation, staging, phase-estimation, diagnosis and/or therapy of
CC   proliferative diseases of breast cells. The method provides very specific
CC   classification of proliferative diseases, allowing better treatment. It
CC   can both characterise methylation status and detect single-nucleotide
CC   polymorphisms. This sequence represents human gene used to illustrate the
CC   method of the invention. NOTE: This sequence does not appear in the
CC   printed specification but has been retrieved from Genbank.
XX
SQ   Sequence 6316 BP; 1611 A; 1680 C; 1554 G; 1471 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              21.4          Length:         6316
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     12            Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x ADP07323 (1-6316)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::   |||      |||
Db       4103 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4162

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||         |||||||||
Db       4163 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4222

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||         |||   ||||||:::|||:::
Db       4223 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4282
```

```
Qy          64 eThrTrpAsnIleAspGlyLeu 71
               |||   :::     |||
Db        4283 TGTCTGGACTTTGAGCAGGCTT 4304

RESULT 61
ABQ34415
ID   ABQ34415 standard; DNA; 894 BP.
XX
AC   ABQ34415;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 21006.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K,  Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 894 BP; 361 A; 301 C; 106 G; 126 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.17          Length:          894
Score:                    77.00         Matches:         24
Percent Similarity:       39.8%         Conservative:    9
Best Local Similarity:    28.9%         Mismatches:      24
Query Match:              17.2%         Indels:          26
DB:                       6             Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x ABQ34415 (1-894)

Qy           2 GluArgAlaLeuAsnSerTyr---------------------------------------- 8
               |||   |||||||||:::|||
Db         202 GAAAACGCGCTTAACGCGTACGTACGCGACGACGATTAACGACGCGCGAACAACGTAAAA 261

Qy           9 -----------------PheGluProProValGluGluSer------------------ 16
                               |||||||||         ::::::
Db         262 CGAAACGAAACAAAACTTTTCGAACCCACGAAACGACAAAACGAACGACCCCTACGACCG 321

Qy          17 AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsn 36
               :::       ||||||||||||   :::|||:::   ||||||          |||
Db         322 TCGACGACCCGACGACCCGAAATATTATCTAAAAAAAAAACG---ACAACCGCGACGACG 378

Qy          37 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 56
                  ||||||   :::|||   :::|||   :::|||                    |||
Db         379 ACGACTACAACGACAACAACCGAAACGAAAACTAACCCTAAAACAACAAACGACTCCGAA 438

Qy          57 AsnGlySer 59
               |||   |||
Db         439 AACGAATCT 447

RESULT 62
ABQ34414/c
ID   ABQ34414 standard; DNA; 894 BP.
XX
AC   ABQ34414;
XX
DT   12-JUL-2002  (first entry)
XX
DE   Oligonucleotide for detecting cytosine methylation SEQ ID NO 21005.
XX
KW   Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW   drug; side effect; cancer; central nervous system; cardiovascular;
KW   gastrointestinal; respiratory system; single nucleotide polymorphism;
KW   SNP; cell differentiation; ds.
XX
OS   Homo sapiens.
XX
PN   WO200218632-A2.
XX
PD   07-MAR-2002.
XX
PF   01-SEP-2001; 2001WO-EP010074.
XX
PR   01-SEP-2000; 2000DE-01043826.
PR   05-SEP-2000; 2000DE-01044543.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K,  Guetig D;
XX
DR   WPI; 2002-371829/40.
XX
PT   Determining the degree of cytosine methylation in genomic DNA, useful for
PT   diagnosis and prognosis, comprises selective hybridization of amplicons
PT   from chemically treated DNA.
XX
PS   Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC   This invention describes a novel method for determining the degree of
CC   methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC   genomic sample of DNA. The sample is treated chemically to convert
CC   cytosine (C) but not methylated C, to uracil, then part of the genomic
CC   DNA that contains the target C is amplified to form a labeled amplicon.
CC   The amplicon is hybridised to two classes, each with at least one member,
CC   of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC   degree of hybridisation to both classes is determined from the label on
CC   the amplicon. From the ratio of labels hybridised to the two classes of
CC   oligomers, the degree of methylation is calculated. The method is used:
CC   (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC   and of a wide range of diseases, e.g. cancer, disorders of the central
CC   nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
```

```
CC   particularly by detecting mutations or single nucleotide polymorphisms
CC   (SNP's); and (ii) for differentiation of cell or tissue types and for
CC   investigating cell differentiation. The method allows the methylation
CC   status of many C residues to be determined simultaneously. ABQ13410-
CC   ABQ54121 represent genomic DNA sequences used to illustrate the method
CC   for determining the degree of cytosine methylation described in the
CC   disclosure of the invention
XX
SQ   Sequence 894 BP; 126 A; 106 C; 301 G; 361 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.17          Length:         894
Score:                  77.00         Matches:        24
Percent Similarity:     39.8%         Conservative:   9
Best Local Similarity:  28.9%         Mismatches:     24
Query Match:            17.2%         Indels:         26
DB:                     6             Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x ABQ34414 (1-894)

Qy          2 GluArgAlaLeuAsnSerTyr----------------------------------------- 8
              |||   |||||||||:::|||
Db        693 GAAAACGCGCTTAACGCGTACGTACGCGACGACGATTAACGACGCGCGAACAACGTAAAA 634

Qy          9 ------------------PheGluProProValGluGluSer------------------ 16
                              |||||||||        ::::::
Db        633 CGAAACGAAACAAAACTTTTCGAACCCACGAAACGACAAAACGAACGACCCCTACGACCG 574

Qy         17 AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsn 36
              :::       ||||||||||||   :::|||:::   ||||||            |||
Db        573 TCGACGACCCGACGACCCGAAATATTATCTAAAAAAAAAACG---ACAACCGCGACGACG 517

Qy         37 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 56
                    ||||||   :::|||   :::|||   :::|||                  |||
Db        516 ACGACTACAACGACAACAACCGAAACGAAAACTAACCCTAAAACAACAAACGACTCCGAA 457

Qy         57 AsnGlySer 59
              |||   |||
Db        456 AACGAATCT 448

RESULT 63
AAD52973
ID   AAD52973 standard; DNA; 10506 BP.
XX
AC   AAD52973;
XX
DT   14-MAY-2003  (first entry)
XX
DE   Bacteriophage N4 virion RNA polymerase (vRNAP) DNA #1.
XX
KW   Virion RNA polymerase; nuclear magnetic resonance; NMR; microinjection;
KW   vRNAP; enzyme; gene; ds.
XX
OS   Bacteriophage N4.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .10503
FT                   /*tag=  a
FT                   /product= "Bacteriophage N4 vRNAP protein"
XX
PN   WO200295002-A2.
XX
PD   28-NOV-2002.
XX
PF   22-MAY-2002; 2002WO-US016295.
XX
PR   22-MAY-2001; 2001US-0292845P.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Kazmierczak KM,  Davydova EK,  Rothman-Denes LB;
XX
DR   WPI; 2003-140368/13.
DR   P-PSDB; AAE34688.
XX
PT   New nucleic acid encoding an N4 virion RNA polymerase for e.g.
PT   synthesizing RNAs of a desired sequence, RNAs for use as probes in
PT   hybridization studies or Southern or Northern blot analysis, and RNA:DNA
PT   hybrids.
XX
PS   Claim 2; Page 118-121; 165pp; English.
XX
CC   The invention relates to bacteriophage N4-coded virion RNA polymerase
CC   (vRNAP) and its nucleic acid. The nucleic acid is used to make an N4
CC   vRNAP which is useful: in the synthesis of RNAs of a desired sequence,
CC   RNAs for use as probes in hybridisation studies or Southern or Northern
CC   blot analysis, and RNA:DNA hybrids for nuclear magnetic resonance (NMR)
CC   structure determination; for in vitro studies of spliceosome assembly,
CC   splicing reactions and antisense experiments; for in vitro translation or
CC   microinjection; and for nucleic acid amplification. The present sequence
CC   is Bacteriophage N4 vRNAP DNA
XX
SQ   Sequence 10506 BP; 3035 A; 2070 C; 2390 G; 3011 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              92.3          Length:         10506
Score:                  75.50         Matches:        24
Percent Similarity:     46.1%         Conservative:   17
Best Local Similarity:  27.0%         Mismatches:     27
Query Match:            16.9%         Indels:         21
DB:                     8             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x AAD52973 (1-10506)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::           |||:::   |||   ||||||
Db       2851 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 2898

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                         |||     |||:::||||||||||||:::                |||
Db       2899 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 2958

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                    :::      :::|||   :::   :::|||               |||   :::
Db       2959 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3003

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                   :::
Db       3004 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3063

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                ::::|||   |||:::   :::|||
Db       3064 TCTAAAGCAGAGGGTATTAAGAACTAT 3090

RESULT 64
ADP09724
ID   ADP09724 standard; DNA; 10506 BP.
XX
AC   ADP09724;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Bacteriophage N4 virion RNA polymerase (vRNAP)-encoding DNA, SEQ:1.
XX
KW   Transcription product; target sequence; single-stranded promoter;
KW   hairpin RNA; short interfering RNA; siRNA; expression attenuation;
KW   nucleic acid library; RNA polymerase; detection; probe synthesis; RNAi;
KW   RNA interference; cell signalling; gene expression analysis;
KW   disease diagnosis; contaminant detection; replica plated microorganism;
KW   bacteriophage N4; RNA polymerase; vRNAP; gene; ds.
XX
OS   Bacteriophage N4.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .10503
FT                   /*tag=  a
```

FT /product= "Bacteriophage N4 virion RNA polymerase (vRNAP)"
XX
PN WO2004048594-A2.
XX
PD 10-JUN-2004.
XX
PF 21-NOV-2003; 2003WO-US037356.
XX
PR 21-NOV-2002; 2002US-0428013P.
XX
PA (EPIC-) EPICENTRE TECHNOLOGIES.
XX
PI Dahl GA, Jendrisak JJ, Davydova E, Rothman-Denes L, Gerdes S;
XX
DR WPI; 2004-450394/42.
DR P-PSDB; ADP09725.
XX
PT Making a transcription product of a target nucleic acid sequence, for
PT diagnosing diseases in plants or animals, comprises admixing RNA
PT polymerase, single-stranded transcription substrate and nucleoside
PT triphosphates.
XX
PS Claim 5; SEQ ID NO 1; 265pp; English.
XX
CC The invention relates to a method of making a transcription product
CC corresponding to a target nucleic acid using an RNA polymerase that is
CC able to transcribe RNA using a single-stranded promoter and using a
CC single-stranded DNA template. The method involves: producing a single-
CC stranded transcription substrate by operably joining a single-stranded
CC target nucleic acid to a single-stranded promoter sequence that can be
CC bound by the RNA polymerase; and incubating the single-stranded
CC transcription substrate, RNA polymerase and nucleoside triphosphates
CC (NTPs) to allow synthesis of the transcription product. In certain
CC embodiments the single-stranded target nucleic acid corresponds to or is
CC complementary to a target nucleic acid in a sample, and in other
CC embodiments the single-stranded target nucleic acid encodes a hairpin RNA
CC such as a short interfering RNA (siRNA). The RNA polymerase used in the
CC invention is preferably the bacteriophage N4 virion RNA polymerase
CC (vRNAP) or functional deletion mutants thereof (referred to as mini-
CC vRNAPs). The invention also relates to a method for obtaining additional
CC rounds of synthesis of transcription product by reverse transcribing the
CC transcription product, joining the produced first-strand cDNA to a single
CC stranded promoter sequence to form a second transcription substrate, and
CC incubating the second transcription substrate with an RNA polymerase that
CC can use a single-stranded promoter. The invention additionally relates to
CC a method for producing hairpin RNAs using the method of the invention; a
CC cell comprising such a hairpin RNA; a method for attenuating expression
CC of a target gene in a cell; a method of cloning a target nucleic acid; a
CC method of constructing a nucleic acid library comprising clones of
CC substantially all nucleic acids or all mRNAs within a sample; a circular
CC single-stranded DNA molecule made using a method of the invention; host
CC cells comprising the circular single-stranded DNA;and a signalling system
CC for detecting and/or quantifying analytes. The methods are useful in
CC research, diagnostic and therapeutic applications, such as preparing cDNA
CC corresponding to mRNA, making sense or anti-sense probes, detecting gene
CC or organism-specific sequences, cloning, cell signalling, or making RNA
CC for use in RNA interference (RNAi). The methods are useful for making
CC transcription products (e.g., hairpin RNA) corresponding to a target
CC nucleic acid sequence to detect target nucleic acids in living cells. The
CC methods are also useful for diagnosing diseases in plants and animals
CC (including humans) and for testing products such as food, blood and
CC tissue cultures for contaminants. The methods are additionally useful for
CC detecting cellular nucleic acids in whole cells from a specimen, such as
CC a fixed or paraffin-embedded section, or from microorganisms immobilised
CC on a solid support such as replica-plated bacteria or yeast. The present
CC sequence represents a specifically claimed polynucleotide encoding a full
CC -length bacteriophage N4 virion RNA polymerase (vRNAP).
XX
SQ Sequence 10506 BP; 3035 A; 2070 C; 2390 G; 3011 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 92.3 Length: 10506

Score: 75.50 Matches: 24
Percent Similarity: 46.1% Conservative: 17
Best Local Similarity: 27.0% Mismatches: 27
Query Match: 16.9% Indels: 21
DB: 12 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x ADP09724 (1-10506)

```
Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::|||||||:::            |||:::   |||   ||||||
Db       2851 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 2898

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::                  |||
Db       2899 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 2958

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||                 |||   :::
Db       2959 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3003

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       3004 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3063

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                ::::|||   |||:::   ::::|||
Db       3064 TCTAAAGCAGAGGGTATTAAGAACTAT 3090
```

RESULT 65
ADQ96719
ID ADQ96719 standard; DNA; 10506 BP.
XX
AC ADQ96719;
XX
DT 07-OCT-2004 (first entry)
XX
DE Bacteriophage N4 virion RNA polymerase (vRNAP) full-length DNA.
XX
KW environmental screening; genetic screening; pathogen detection;
KW gene detection; SNP detection; food; blood; tissue culture contamination;
KW ds; gene; virion RNA polymerase; vRNAP.
XX
OS Bacteriophage N4.
XX
FH Key Location/Qualifiers
FT CDS 1. .10503
FT /*tag= a
FT /product= "Bacteriophage N4 virion RNA polymerase (vRNAP)
FT full-length protein"
FT /note= "This CDS contains translation exceptions"
XX
PN WO2004059289-A2.
XX
PD 15-JUL-2004.
XX
PF 23-DEC-2003; 2003WO-US041046.
XX
PR 23-DEC-2002; 2002US-0436062P.
XX
PA (EPIC-) EPICENTRE TECHNOLOGIES.
PA (DAHL/) DAHL G A.
PA (JEND/) JENDRISAK J J.
PA (DAVY/) DAVYDOVA E.
PA (ROTH/) ROTHMAN-DENES L.
PA (GERD/) GERDES S.
XX
DR WPI; 2004-543544/52.
DR P-PSDB; ADQ96720.
XX
PT Detecting a target nucleic acid sequence, useful for diagnosing diseases
PT in plants and animals, comprises contacting target probes with a target
PT nucleic acid sequence and incubating under hybridization conditions.
XX

```
PS   Claim 35; SEQ ID NO 1; 296pp; English.
XX
CC   The invention relates to a novel method for detecting a target nucleic
CC   acid sequence which comprises contacting a linear single-stranded DNA
CC   target probe with a target nucleic acid sequence and incubating under
CC   hybridisation conditions such that the target-complementary sequences
CC   anneal adjacently to the target nucleic acid sequence to form a target
CC   probe-target complex and contacting the target probe-target complex with
CC   a ligase to form a transcription substrate. The method or kit of the
CC   invention may be useful for research or environmental and genetic
CC   screening, as well as for detection and quantification of a pathogen, an
CC   analyte, a gene or allele or a single nucleotide polymorphism (SNP) or
CC   other type of gene mutation or variant. The method and kit may also be
CC   utilised for diagnosing diseases in plants and animals, including humans
CC   and for testing products, such as food, blood, and tissue cultures for
CC   contaminants. The current sequence is that of the Bacteriophage N4 virion
CC   RNA polymerase (vRNAP) full-length DNA of the invention.
XX
SQ   Sequence 10506 BP; 3035 A; 2074 C; 2389 G; 3007 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                92.3          Length:         10506
Score:                    75.50         Matches:        24
Percent Similarity:       46.1%         Conservative:   17
Best Local Similarity:    27.0%         Mismatches:     27
Query Match:              16.9%         Indels:         21
DB:                       13            Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x ADQ96719 (1-10506)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::           |||:::   |||   ||||||
Db       2851 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 2898

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::                |||
Db       2899 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 2958

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||               |||   :::
Db       2959 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3003

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       3004 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3063

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       3064 TCTAAAGCAGAGGGTATTAAGAACTAT 3090

RESULT 66
ADT91343
ID   ADT91343 standard; DNA; 10506 BP.
XX
AC   ADT91343;
XX
DT   16-DEC-2004  (first entry)
XX
DE   N4 virion RNA polymerase (vRNAP) DNA #1.
XX
KW   N4 virion RNA polymerase; vRNAP; mini-vRNAP; gene; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .10503
FT                   /*tag=  a
FT                   /product= "N4 virion RNA polymerase (vRNAP)"
XX
PN   US2004191812-A1.
XX
PD   30-SEP-2004.
XX
PF   23-DEC-2003; 2003US-00743975.
XX
PR   22-MAY-2001; 2001US-0292845P.
PR   22-MAY-2002; 2002US-00153219.
PR   23-DEC-2002; 2002US-0436062P.
XX
PA   (DAVY/) DAVYDOVA E K.
PA   (ROTH/) ROTHMAN-DENES L B.
PA   (DAHL/) DAHL G A.
PA   (GERD/) GERDES S Y.
PA   (JEND/) JENDRISAK J J.
XX
PI   Davydova EK,  Rothman-Denes LB,  Dahl GA,  Gerdes SY,  Jendrisak JJ;
XX
DR   WPI; 2004-698714/68.
DR   P-PSDB; ADT91344.
XX
PT   Detecting target nucleic acid sequence, by contacting target probe having
PT   single-stranded DNA molecule with target sequence to form probe-target
PT   complex, contacting complex with ligase, RNA polymerase, detecting
PT   transcription product.
XX
PS   Claim 35; SEQ ID NO 1; 147pp; English.
XX
CC   The invention relates to methods, compositions and kits that use N4
CC   virion RNA polymerase (vRNAP; designated as mini-vRNAP) deletion mutants
CC   to detect and quantify target nucleic acid sequence. The method is useful
CC   for detecting target nucleic acid sequences and analytes in a sample. The
CC   present sequence is a N4 vRNAP DNA.
XX
SQ   Sequence 10506 BP; 3035 A; 2070 C; 2390 G; 3011 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                92.3          Length:         10506
Score:                    75.50         Matches:        24
Percent Similarity:       46.1%         Conservative:   17
Best Local Similarity:    27.0%         Mismatches:     27
Query Match:              16.9%         Indels:         21
DB:                       13            Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x ADT91343 (1-10506)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::           |||:::   |||   ||||||
Db       2851 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 2898

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::                |||
Db       2899 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 2958

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||               |||   :::
Db       2959 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3003

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       3004 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3063

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       3064 TCTAAAGCAGAGGGTATTAAGAACTAT 3090

RESULT 67
AAD52982
ID   AAD52982 standard; DNA; 10617 BP.
XX
AC   AAD52982;
XX
DT   14-MAY-2003  (first entry)
XX
DE   Bacteriophage N4 full-length virion RNA polymerase (vRNAP) DNA.
XX
KW   Virion RNA polymerase; nuclear magnetic resonance; NMR; microinjection;
```

```
KW   vRNAP; enzyme; gene; ds.
XX
OS   Bacteriophage N4.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .10614
FT                   /*tag= a
FT                   /product= "Bacteriophage N4 vRNAP protein"
XX
PN   WO200295002-A2.
XX
PD   28-NOV-2002.
XX
PF   22-MAY-2002; 2002WO-US016295.
XX
PR   22-MAY-2001; 2001US-0292845P.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Kazmierczak KM,  Davydova EK,  Rothman-Denes LB;
XX
DR   WPI; 2003-140368/13.
DR   P-PSDB; AAE34692.
XX
PT   New nucleic acid encoding an N4 virion RNA polymerase for e.g.
PT   synthesizing RNAs of a desired sequence, RNAs for use as probes in
PT   hybridization studies or Southern or Northern blot analysis, and RNA:DNA
PT   hybrids.
XX
PS   Claim 2; Page 148-151; 165pp; English.
XX
CC   The invention relates to bacteriophage N4-coded virion RNA polymerase
CC   (vRNAP) and its nucleic acid. The nucleic acid is used to make an N4
CC   vRNAP which is useful: in the synthesis of RNAs of a desired sequence,
CC   RNAs for use as probes in hybridisation studies or Southern or Northern
CC   blot analysis, and RNA:DNA hybrids for nuclear magnetic resonance (NMR)
CC   structure determination; for in vitro studies of spliceosome assembly,
CC   splicing reactions and antisense experiments; for in vitro translation or
CC   microinjection; and for nucleic acid amplification. The present sequence
CC   is Bacteriophage N4 full-length vRNAP DNA
XX
SQ   Sequence 10617 BP; 3063 A; 2092 C; 2424 G; 3038 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                93.6          Length:          10617
Score:                    75.50         Matches:         24
Percent Similarity:       46.1%         Conservative:    17
Best Local Similarity:    27.0%         Mismatches:      27
Query Match:              16.9%         Indels:          21
DB:                       8             Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x AAD52982 (1-10617)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||     :::||||||:::            |||:::   |||   ||||||
Db       2962 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 3009

Qy         29 ----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::                  |||
Db       3010 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 3069

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||                |||   :::
Db       3070 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3114

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       3115 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3174

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                :::|||   |||:::   :::|||
Db       3175 TCTAAAGCAGAGGGTATTAAGAACTAT 3201
```

```
RESULT 68
ADP09737
ID   ADP09737 standard; DNA; 10617 BP.
XX
AC   ADP09737;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Bacteriophage N4 His-tagged full-length vRNAP DNA, SEQ:14.
XX
KW   Transcription product; target sequence; single-stranded promoter;
KW   hairpin RNA; short interfering RNA; siRNA; expression attenuation;
KW   nucleic acid library; RNA polymerase; detection; probe synthesis; RNAi;
KW   RNA interference; cell signalling; gene expression analysis;
KW   disease diagnosis; contaminant detection; replica plated microorganism;
KW   bacteriophage N4; RNA polymerase; vRNAP; gene; ds.
XX
OS   Bacteriophage N4.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .10614
FT                   /*tag= a
XX
PN   WO2004048594-A2.
XX
PD   10-JUN-2004.
XX
PF   21-NOV-2003; 2003WO-US037356.
XX
PR   21-NOV-2002; 2002US-0428013P.
XX
PA   (EPIC-) EPICENTRE TECHNOLOGIES.
XX
PI   Dahl GA,  Jendrisak JJ,  Davydova E,  Rothman-Denes L,  Gerdes S;
XX
DR   WPI; 2004-450394/42.
DR   P-PSDB; ADP09738.
XX
PT   Making a transcription product of a target nucleic acid sequence, for
PT   diagnosing diseases in plants or animals, comprises admixing RNA
PT   polymerase, single-stranded transcription substrate and nucleoside
PT   triphosphates.
XX
PS   Claim 5; SEQ ID NO 14; 265pp; English.
XX
CC   The invention relates to a method of making a transcription product
CC   corresponding to a target nucleic acid using an RNA polymerase that is
CC   able to transcribe RNA using a single-stranded promoter and using a
CC   single-stranded DNA template. The method involves: producing a single-
CC   stranded transcription substrate by operably joining a single-stranded
CC   target nucleic acid to a single-stranded promoter sequence that can be
CC   bound by the RNA polymerase; and incubating the single-stranded
CC   transcription substrate, RNA polymerase and nucleoside triphosphates
CC   (NTPs) to allow synthesis of the transcription product. In certain
CC   embodiments the single-stranded target nucleic acid corresponds to or is
CC   complementary to a target nucleic acid in a sample, and in other
CC   embodiments the single-stranded target nucleic acid encodes a hairpin RNA
CC   such as a short interfering RNA (siRNA). The RNA polymerase used in the
CC   invention is preferably the bacteriophage N4 virion RNA polymerase
CC   (vRNAP) or functional deletion mutants thereof (referred to as mini-
CC   vRNAPs). The invention also relates to a method for obtaining additional
CC   rounds of synthesis of transcription product by reverse transcribing the
CC   transcription product, joining the produced first-strand cDNA to a single
CC   stranded promoter sequence to form a second transcription substrate, and
CC   incubating the second transcription substrate with an RNA polymerase that
CC   can use a single-stranded promoter. The invention additionally relates to
CC   a method for producing hairpin RNAs using the method of the invention; a
CC   cell comprising such a hairpin RNA; a method for attenuating expression
CC   of a target gene in a cell; a method of cloning a target nucleic acid; a
CC   method of constructing a nucleic acid library comprising clones of
CC   substantially all nucleic acids or all mRNAs within a sample; a circular
CC   single-stranded DNA molecule made using a method of the invention; host
```

```
CC   cells comprising the circular single-stranded DNA;and a signalling system
CC   for detecting and/or quantifying analytes. The methods are useful in
CC   research, diagnostic and therapeutic applications, such as preparing cDNA
CC   corresponding to mRNA, making sense or anti-sense probes, detecting gene
CC   or organism-specific sequences, cloning, cell signalling, or making RNA
CC   for use in RNA interference (RNAi). The methods are useful for making
CC   transcription products (e.g., hairpin RNA) corresponding to a target
CC   nucleic acid sequence to detect target nucleic acids in living cells. The
CC   methods are also useful for diagnosing diseases in plants and animals
CC   (including humans) and for testing products such as food, blood and
CC   tissue cultures for contaminants. The methods are additionally useful for
CC   detecting cellular nucleic acids in whole cells from a specimen, such as
CC   a fixed or paraffin-embedded section, or from microorganisms immobilised
CC   on a solid support such as replica-plated bacteria or yeast. The present
CC   sequence represents a specifically claimed polynucleotide encoding a full
CC   -length bacteriophage N4 virion RNA polymerase (vRNAP) which contains an
CC   N-terminal His-tag.
XX
SQ   Sequence 10617 BP; 3063 A; 2092 C; 2424 G; 3038 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                93.6          Length:         10617
Score:                    75.50         Matches:        24
Percent Similarity:       46.1%         Conservative:   17
Best Local Similarity:    27.0%         Mismatches:     27
Query Match:              16.9%         Indels:         21
DB:                       12            Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x ADP09737 (1-10617)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::           |||:::   |||   ||||||
Db       2962 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 3009

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::                |||
Db       3010 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 3069

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||               |||   :::
Db       3070 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3114

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                  :::
Db       3115 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3174

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       3175 TCTAAAGCAGAGGGTATTAAGAACTAT 3201

RESULT 69
ADQ96732
ID   ADQ96732 standard; DNA; 10617 BP.
XX
AC   ADQ96732;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Bacteriophage N4 virion RNA polymerase (vRNAP) full-length DNA + His tag.
XX
KW   environmental screening; genetic screening; pathogen detection;
KW   gene detection; SNP detection; food; blood; tissue culture contamination;
KW   ds; gene; virion RNA polymerase; vRNAP; His tag.
XX
OS   Bacteriophage N4.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .10614
FT                   /*tag=  a
FT                   /product= "Bacteriophage N4 virion RNA polymerase (vRNAP)
FT                   full-length protein + His tag"
XX
PN   WO2004059289-A2.
XX
PD   15-JUL-2004.
XX
PF   23-DEC-2003; 2003WO-US041046.
XX
PR   23-DEC-2002; 2002US-0436062P.
XX
PA   (EPIC-) EPICENTRE TECHNOLOGIES.
PA   (DAHL/) DAHL G A.
PA   (JEND/) JENDRISAK J J.
PA   (DAVY/) DAVYDOVA E.
PA   (ROTH/) ROTHMAN-DENES L.
PA   (GERD/) GERDES S.
XX
DR   WPI; 2004-543544/52.
DR   P-PSDB; ADQ96733.
XX
PT   Detecting a target nucleic acid sequence, useful for diagnosing diseases
PT   in plants and animals, comprises contacting target probes with a target
PT   nucleic acid sequence and incubating under hybridization conditions.
XX
PS   Claim 35; SEQ ID NO 14; 296pp; English.
XX
CC   The invention relates to a novel method for detecting a target nucleic
CC   acid sequence which comprises contacting a linear single-stranded DNA
CC   target probe with a target nucleic acid sequence and incubating under
CC   hybridisation conditions such that the target-complementary sequences
CC   anneal adjacently to the target nucleic acid sequence to form a target
CC   probe-target complex and contacting the target probe-target complex with
CC   a ligase to form a transcription substrate. The method or kit of the
CC   invention may be useful for research or environmental and genetic
CC   screening, as well as for detection and quantification of a pathogen, an
CC   analyte, a gene or allele or a single nucleotide polymorphism (SNP) or
CC   other type of gene mutation or variant. The method and kit may also be
CC   utilised for diagnosing diseases in plants and animals, including humans
CC   and for testing products, such as food, blood, and tissue cultures for
CC   contaminants. The current sequence is that of the Bacteriophage N4 virion
CC   RNA polymerase (vRNAP) full-length DNA with a purification-related His
CC   tag of the invention.
XX
SQ   Sequence 10617 BP; 3063 A; 2092 C; 2424 G; 3038 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                93.6          Length:         10617
Score:                    75.50         Matches:        24
Percent Similarity:       46.1%         Conservative:   17
Best Local Similarity:    27.0%         Mismatches:     27
Query Match:              16.9%         Indels:         21
DB:                       13            Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x ADQ96732 (1-10617)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::           |||:::   |||   ||||||
Db       2962 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 3009

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||      |||:::||||||||||||:::                |||
Db       3010 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 3069

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                 :::      :::|||   :::   :::|||               |||   :::
Db       3070 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3114

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                  :::
Db       3115 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3174

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       3175 TCTAAAGCAGAGGGTATTAAGAACTAT 3201
```

```
RESULT 70
ADT91356
ID   ADT91356 standard; DNA; 10617 BP.
XX
AC   ADT91356;
XX
DT   16-DEC-2004  (first entry)
XX
DE   N4 virion RNA polymerase (vRNAP) DNA #5.
XX
KW   N4 virion RNA polymerase; vRNAP; mini-vRNAP; gene; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .10614
FT                   /*tag=  a
FT                   /product= "N4 virion RNA polymerase (vRNAP)"
XX
PN   US2004191812-A1.
XX
PD   30-SEP-2004.
XX
PF   23-DEC-2003; 2003US-00743975.
XX
PR   22-MAY-2001; 2001US-0292845P.
PR   22-MAY-2002; 2002US-00153219.
PR   23-DEC-2002; 2002US-0436062P.
XX
PA   (DAVY/) DAVYDOVA E K.
PA   (ROTH/) ROTHMAN-DENES L B.
PA   (DAHL/) DAHL G A.
PA   (GERD/) GERDES S Y.
PA   (JEND/) JENDRISAK J J.
XX
PI   Davydova EK,  Rothman-Denes LB,  Dahl GA,  Gerdes SY,  Jendrisak JJ;
XX
DR   WPI; 2004-698714/68.
DR   P-PSDB; ADT91357.
XX
PT   Detecting target nucleic acid sequence, by contacting target probe having
PT   single-stranded DNA molecule with target sequence to form probe-target
PT   complex, contacting complex with ligase, RNA polymerase, detecting
PT   transcription product.
XX
PS   Claim 35; SEQ ID NO 14; 147pp; English.
XX
CC   The invention relates to methods, compositions and kits that use N4
CC   virion RNA polymerase (vRNAP; designated as mini-vRNAP) deletion mutants
CC   to detect and quantify target nucleic acid sequence. The method is useful
CC   for detecting target nucleic acid sequences and analytes in a sample. The
CC   present sequence is a N4 vRNAP DNA.
XX
SQ   Sequence 10617 BP; 3063 A; 2092 C; 2424 G; 3038 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                93.6          Length:          10617
Score:                    75.50         Matches:         24
Percent Similarity:       46.1%         Conservative:    17
Best Local Similarity:    27.0%         Mismatches:      27
Query Match:              16.9%         Indels:          21
DB:                       13            Gaps:            5

US-10-757-745-2_COPY_54_140 (1-87) x ADT91356 (1-10617)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::           |||:::   |||   ||||||
Db       2962 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 3009

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                       |||       |||:::|||||||||||||:::              |||
Db       3010 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 3069

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                    :::       :::|||   :::   :::|||                |||   :::
Db       3070 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3114

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||       :::   :::|||:::|||                     :::
Db       3115 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3174

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       3175 TCTAAAGCAGAGGGTATTAAGAACTAT 3201

RESULT 71
ADQ08623
ID   ADQ08623 standard; DNA; 2601 BP.
XX
AC   ADQ08623;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Ciona intestinalis nervous system associated gene SeqID25.
XX
KW   gene cluster; nervous system; sea-squirt tailbud; embryo; larva;
KW   nervous system disease; gene; ds.
XX
OS   Ciona intestinalis.
XX
PN   JP2004057127-A.
XX
PD   26-FEB-2004.
XX
PF   31-JUL-2002; 2002JP-00222532.
XX
PR   31-JUL-2002; 2002JP-00222532.
XX
PA   (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR   WPI; 2004-208712/20.
DR   P-PSDB; ADQ08624.
XX
PT   Novel genes derived from Ciona intestinalis (sea squirt), expressed in
PT   nervous system in the tailbud embryo or larva, useful for studying the
PT   development of nervous system.
XX
PS   Claim 2; SEQ ID NO 25; 897pp; Japanese.
XX
CC   This invention relates to a novel gene cluster, where the encoded
CC   proteins are expressed in the nervous system of sea-squirt tailbud embryo
CC   or larva. The invention is useful for studying the development of the
CC   nervous system of the sea-squirt and for research purposes. The genes may
CC   be used for determining the disease-development mechanisms in the nervous
CC   system. In addition, novel gene clusters expressed in nervous system of
CC   sea-squirt tailbud embryo or larva allows development of diagnostics and
CC   therapeutics related to nervous system diseases. The present sequence is
CC   that of a C intestinalis gene of the invention.
XX
SQ   Sequence 2601 BP; 815 A; 557 C; 508 G; 721 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                16.8          Length:          2601
Score:                    75.00         Matches:         30
Percent Similarity:       36.6%         Conservative:    11
Best Local Similarity:    26.8%         Mismatches:      31
Query Match:              16.7%         Indels:          40
DB:                       12            Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x ADQ08623 (1-2601)

Qy          3 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSer------------------ 16
              ||||||      :::|||       |||   :::||||||
Db        396 AGAGCAGAAGCTGCATATGATAACCCTAGCTTGGAAGAAGATAACCATTCAACTTATTCA 455

Qy         17 ---------------AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
```

```
                              ::::::         |||:::|||         ||||||
Db        456 AGAGCAACAACAATTACAATGGATAACTCCTCGGAAAGTATTCCACATAGAAAAACCGGC 515

Qy         32 ValAspLeu--------------------------------------------------Thr 35
                 :::|||                                                 :::
Db        516 ACTGAATTAAAGATGGAAAAATTAGGGATGGGACAAAGTGCGGACTTTGATGTTGATTCT 575

Qy         36 AsnGluGluThrThrAspSerThrThrSerLysIleSerPro------------------ 49
                |||||   |||||||||             |||   |||
Db        576 GAAGAGGAAGAAACAGATTCAGGAATGGTTGCGATAGGTCCTCGACCACAAACAGCAAAA 635

Qy         50 --------------SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                           |||||||||           ||||||      |||
Db        636 TCGAATTCTGCATCAAAAGAGGACACTGACGTTCGACCTGGCAGC------AGTGCATCG 689

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 76
                 :::||||||   |||:::||||||::::::|||
Db        690 AAATACAACATCCCCGGTATGGACTTGGACGACTTA 725

RESULT 72
AAF57111/c
ID   AAF57111 standard; cDNA; 2148 BP.
XX
AC   AAF57111;
XX
DT   29-MAY-2001  (first entry)
XX
DE   A. thaliana AtCNGC1 polypeptide encoding cDNA.
XX
KW   Calmodulin-binding polypeptide; CBP; calcium; transgenic plant; cation;
KW   phytoremediation; agriculture; CNGC1; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2148
FT                   /*tag=  a
FT                   /product= "CNGC1"
FT                   /note= "the stop codon is not indicated"
XX
PN   WO200116144-A2.
XX
PD   08-MAR-2001.
XX
PF   27-AUG-2000; 2000WO-IL000511.
XX
PR   29-AUG-1999;   99IL-00131633.
XX
PA   (YEDA ) YEDA RES & DEV CO LTD.
XX
PI   Fromm H,  Arazi T;
XX
DR   WPI; 2001-169004/17.
DR   P-PSDB; AAB62045.
XX
PT   New NtcCBP polypeptides, used to produce transgenic plants which are
PT   resistant to metal cations, or which can be used for phytoremediation of
PT   areas contaminated with metal cations.
XX
PS   Claim 26; Page 113-114; 129pp; English.
XX
CC   The invention provides plant derived calcium-dependent calmodulin-binding
CC   polypeptides (CBP) and polynucleotides encoding the CBPs. The CBPs can be
CC   used for producing transgenic plants capable of surviving in areas of
CC   high metal cation concentration, without taking up dangerous quantities
CC   of the cations, or capable of phytoremediation of an area contaminated
CC   with metal cations. The transgenic plants assist agriculture as they can
CC   be grown in areas contaminated with metal cations, without becoming
CC   toxic. The present sequence represents a cDNA encoding a A. thaliana
CC   AtCNGC1 polypeptide
XX
SQ   Sequence 2148 BP; 604 A; 460 C; 468 G; 616 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              17.8          Length:         2148
Score:                  74.00         Matches:        19
Percent Similarity:     46.7%         Conservative:   9
Best Local Similarity:  31.7%         Mismatches:     18
Query Match:            16.5%         Indels:         14
DB:                     4             Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x AAF57111 (1-2148)

Qy         30 ThrTyrValAspLeuThr---------AsnGluGluThrThrAspSerThrThrSerLys 46
                |||:::   ||||||:::         |||:::   |||:::         |||   :::
Db        546 ACGTTTAGCGATCTCTCGCTTGTCCTCAACAAGAACACCTCTCCCAAAAACACGAGACGA 487

Qy         47 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
                  |||            |||:::   ||||||
Db        486 AGGAGCGATAAAGCCAGTACGGAACTGGAAAAT-------------------------- 454

Qy         67 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                  |||      |||:::   |||:::||||||||||||:::      |||   |||
Db        453 ------GATGTGAAGGACATAAAAAACATCAGTGAAAGAGCGCAAAACGCTTGCTGTTAT 400

RESULT 73
AAS73107/c
ID   AAS73107 standard; cDNA; 2490 BP.
XX
AC   AAS73107;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8911.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08920.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 8911; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
```

```
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2490 BP; 504 A; 418 C; 512 G; 1056 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                21.7          Length:          2490
Score:                    74.00         Matches:         28
Percent Similarity:       44.6%         Conservative:    13
Best Local Similarity:    30.4%         Mismatches:      25
Query Match:              16.5%         Indels:          26
DB:                       5             Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x AAS73107 (1-2490)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGlu--- 19
              |||::::::   ||||||:::|||      |||   :::   :::      ||||||
Db       1994 ATGGATAAATTCCTCAACACATACACCCTCCCAAGACTAAACCAGGAAGAACTTGAATCT 1935

Qy         20 ---ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 38
                ||||||   |||   ||||||   :::|||   :::
Db       1934 CTGAATAGACCAATAACAGGCTCTGAAATTGAGACAATAATT------------------ 1893

Qy         39 ThrThrAspSerThrThrSerLysIleSerProSerGluAsp------------------ 52
              :::|||    :::|||  ||||||     |||
Db       1892 ------AATAGCTTACCAACCAAAAAAAGTCCAGGATCAGATGGATTCACAGCCGAATTC 1839

Qy         53 ------------------------ThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                                      |||::::::||||||                |||||:::
Db       1838 TACCAGAGGCCTGGCAGAGACACAACAAAAAAAGAGAATTTTAGACCAATATCCTTGATG 1779

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeu 76
              |||||||||      ||||||   :::
Db       1778 ------AACATTGATGCAAAAATCCTCAATAAAATA 1749

RESULT 74
AAF58469/c
ID   AAF58469 standard; DNA; 2547 BP.
XX
AC   AAF58469;
XX
DT   24-APR-2001  (first entry)
XX
DE   AtCNGC1/DND2 coding sequence.
XX
KW   AtCNGC1/DND2; cyclic nucleotide-gated ion channel; defence no death;
KW   disease resistance; chromosome 5; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200107596-A1.
XX
PD   01-FEB-2001.
XX
PF   24-JUL-2000; 2000WO-US020216.
XX
PR   23-JUL-1999;   99US-0145310P.
XX
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Bent AF,  Yu I,  Clough SJ,  Fengler KA,  Smith RK;
XX
DR   WPI; 2001-159718/16.
DR   P-PSDB; AAB68591.
XX
PT   Improving disease resistance, controlling cell death or reducing
PT   hypersensitive response to pathogen attack in plants, involves
PT   downregulating or inactivating cyclic nucleotide gated ion channel gene
PT   or gene product.
XX
PS   Claim 11; Fig 14; 91pp; English.
XX
CC   The present sequence is the coding sequence for AtCNGC1/DND2 from
CC   Arabidopsis thaliana. AtCNGC1 is "cyclic nucleotide-gated ion channel",
CC   and DND is "defence, no death". Manipulation of AtCNGC1/DND2 gene (i.e.
CC   downregulating, mutating or inactivating) allows generation of plants
CC   with improved disease resistance. AtCNGC1/DND2 can be used for reducing
CC   hypersensitive response (HR) to a pathogen attack in a plant, controlling
CC   cell death and improving resistance in a plant against a disease caused
CC   by virus such as nepovirus including tobacco ringspot virus, a bacteria
CC   such as a gram-negative bacterium e.g. bacteria of the genus Pseudomonas
CC   or Xanthorionas, including P. syringae pv.tomato and Xanthomonas
CC   campestris pv.campestris and a fungi such as an ascomycete fungus
CC   including fungi of genus Erysiphe e.g. E. orontii. The DND2 gene maps to
CC   the lower arm of Arabidopsis chromosome 5 (close to markers nga129 and
CC   LFY3)
XX
SQ   Sequence 2547 BP; 724 A; 527 C; 526 G; 770 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                22.3          Length:          2547
Score:                    74.00         Matches:         19
Percent Similarity:       46.7%         Conservative:    9
Best Local Similarity:    31.7%         Mismatches:      18
Query Match:              16.5%         Indels:          14
DB:                       4             Gaps:            2

US-10-757-745-2_COPY_54_140 (1-87) x AAF58469 (1-2547)

Qy         30 ThrTyrValAspLeuThr---------AsnGluGluThrThrAspSerThrThrSerLys 46
              |||:::   ||||||:::          |||:::   |||:::      |||   :::
Db       850 ACGTTTAGCGATCTCTCGCTTGTCCTCAACAAGAACACCTCTCCCAAAAACACGAGACGA 791

Qy         47 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
                 |||          |||:::   ||||||
Db       790 AGGAGCGATAAAGCCAGTACGGAACTGGAAAAT-------------------------- 758

Qy         67 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                   |||      |||:::   |||:::|||||||||||:::      |||   |||
Db       757 ------GATGTGAAGGACATAAAAAACATCAGTGAAAGAGCGCAAAACGCTTGCTGTTAT 704

RESULT 75
AAF58468/c
ID   AAF58468 standard; DNA; 5996 BP.
XX
AC   AAF58468;
XX
DT   24-APR-2001  (first entry)
XX
DE   AtCNGC1/DND2 gene.
XX
KW   AtCNGC1/DND2; cyclic nucleotide-gated ion channel; defence no death;
KW   disease resistance; chromosome 5; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200107596-A1.
XX
PD   01-FEB-2001.
XX
PF   24-JUL-2000; 2000WO-US020216.
XX
PR   23-JUL-1999;   99US-0145310P.
XX
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Bent AF,  Yu I,  Clough SJ,  Fengler KA,  Smith RK;
XX
DR   WPI; 2001-159718/16.
DR   P-PSDB; AAB68591.
```

XX
PT Improving disease resistance, controlling cell death or reducing
PT hypersensitive response to pathogen attack in plants, involves
PT downregulating or inactivating cyclic nucleotide gated ion channel gene
PT or gene product.
XX
PS Disclosure; Fig 13; 91pp; English.
XX
CC The present sequence is the genomic region containing AtCNGC1/DND2 gene
CC from Arabidopsis thaliana. AtCNGC1 is "cyclic nucleotide-gated ion
CC channel", and DND is "defence, no death". Manipulation of AtCNGC1/DND2
CC gene (i.e. downregulating, mutating or inactivating) allows generation of
CC plants with improved disease resistance. AtCNGC1/DND2 can be used for
CC reducing hypersensitive response (HR) to a pathogen attack in a plant,
CC controlling cell death and improving resistance in a plant against a
CC disease caused by virus such as nepovirus including tobacco ringspot
CC virus, a bacteria such as a gram-negative bacterium e.g. bacteria of the
CC genus Pseudomonas or Xanthorionas, including P. syringae pv.tomato and
CC Xanthomonas campestris pv.campestris and a fungi such as an ascomycete
CC fungus including fungi of genus Erysiphe e.g. E. orontii. The DND2 gene
CC maps to the lower arm of Arabidopsis chromosome 5 (close to markers
CC nga129 and LFY3)
XX
SQ Sequence 5996 BP; 1883 A; 1054 C; 1120 G; 1939 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	69.8	Length:	5996
Score:	74.00	Matches:	19
Percent Similarity:	46.7%	Conservative:	9
Best Local Similarity:	31.7%	Mismatches:	18
Query Match:	16.5%	Indels:	14
DB:	4	Gaps:	2

US-10-757-745-2_COPY_54_140 (1-87) x AAF58468 (1-5996)

```
Qy         30 ThrTyrValAspLeuThr---------AsnGluGluThrThrAspSerThrThrSerLys 46
              |||:::   ||||||:::          |||:::   |||:::          |||   :::
Db       3502 ACGTTTAGCGATCTCTCGCTTGTCCTCAACAAGAACACCTCTCCCAAAAACACGAGACGA 3443

Qy         47 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
                 |||           |||:::   ||||||
Db       3442 AGGAGCGATAAAGCCAGTACGGAACTGGAAAAT-------------------------- 3410

Qy         67 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                 |||      |||:::   |||:::|||||||||||||:::      |||   |||
Db       3409 ------GATGTGAAGGACATAAAAAACATCAGTGAAAGAGCGCAAAACGCTTGCTGTTAT 3356
```

Search completed: March 10, 2006, 15:34:11
Job time : 380.652 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:19:50 ; Search time 1808.48 Seconds
(without alignments)
1345.232 Million cell updates/sec

Title: US-10-757-745-6_COPY_2_45
Perfect score: 229
Sequence: 1 GDPGSEIIESVPPAGPEASE.....ENEDDIQFVSEGPSRPVLEY 44

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170333_4165/app_query.fasta_1
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=rnpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss02h -USER=US10757745 @CGN_1_1_20983_@runat_09032006_170333_4165
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUSB_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUSC_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US075_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US076_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US077_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US078_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US079_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US101_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US104_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US105_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US112_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US117_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US600_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US601_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US603_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US605_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US606_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US607_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	229	100.0	250	17	US-08-779-004-1142	Sequence 1142, Ap
2	229	100.0	250	26	US-09-540-233D-14754	Sequence 14754, A
3	229	100.0	250	74	US-60-009-671-1142	Sequence 1142, Ap
4	229	100.0	272	21	US-09-105-427-308	Sequence 308, App
5	229	100.0	272	26	US-09-540-208-8004	Sequence 8004, Ap
6	229	100.0	274	61	US-10-793-479-1651	Sequence 1651, Ap
7	229	100.0	284	26	US-09-540-764-58131	Sequence 58131, A
8	229	100.0	284	49	US-10-349-781-58131	Sequence 58131, A

9	229	100.0	312	24	US-09-489-036-12694	Sequence 12694, A
10	229	100.0	312	37	US-09-943-143-12694	Sequence 12694, A
11	229	100.0	366	22	US-09-205-070-10288	Sequence 10288, A
12	229	100.0	366	22	US-09-271-122-13443	Sequence 13443, A
13	229	100.0	366	23	US-09-340-623-10288	Sequence 10288, A
14	229	100.0	366	31	US-09-713-841-13443	Sequence 13443, A
15	229	100.0	366	34	US-09-898-888-10288	Sequence 10288, A
16	229	100.0	366	34	US-09-898-888A-10288	Sequence 10288, A
17	229	100.0	366	36	US-09-927-875-13443	Sequence 13443, A
18	229	100.0	366	36	US-09-927-875A-13443	Sequence 13443, A
19	229	100.0	366	48	US-10-302-689A-10288	Sequence 10288, A
20	229	100.0	373	24	US-09-489-036-20573	Sequence 20573, A
21	229	100.0	373	37	US-09-943-143-20573	Sequence 20573, A
22	229	100.0	418	24	US-09-431-517-5512	Sequence 5512, Ap
23	229	100.0	420	24	US-09-431-517-5511	Sequence 5511, Ap
24	229	100.0	632	1	PCT-US02-30474-2111	Sequence 2111, Ap
25	229	100.0	632	25	US-09-515-126-12696	Sequence 12696, A
26	229	100.0	632	27	US-09-577-409-12696	Sequence 12696, A
27	229	100.0	632	43	US-10-220-366A-12576	Sequence 12576, A
28	229	100.0	632	48	US-10-302-689A-106285	Sequence 106285,
29	229	100.0	632	51	US-10-489-448-2111	Sequence 2111, Ap
30	229	100.0	632	78	US-60-324-631-2095	Sequence 2095, Ap
31	229	100.0	1399	75	US-60-172-377-1376	Sequence 1376, Ap
32	229	100.0	1399	77	US-60-278-561-12578	Sequence 12578, A
33	229	100.0	1536	61	US-10-757-745-5	Sequence 5, Appli
34	229	100.0	1571	23	US-09-397-206A-5312	Sequence 5312, Ap
35	229	100.0	1697	31	US-09-758-465-402	Sequence 402, App
36	229	100.0	1697	43	US-10-222-927-402	Sequence 402, App
37	229	100.0	2789	29	US-09-644-867-6750	Sequence 6750, Ap
38	229	100.0	2789	29	US-09-652-913-10806	Sequence 10806, A
39	229	100.0	2789	30	US-09-699-997-10714	Sequence 10714, A
40	229	100.0	2789	31	US-09-717-350-4792	Sequence 4792, Ap
41	225	98.3	457	25	US-09-528-409-3837	Sequence 3837, Ap
42	225	98.3	457	36	US-09-933-524-3837	Sequence 3837, Ap
43	225	98.3	457	36	US-09-933-524A-3837	Sequence 3837, Ap
44	225	98.3	493	26	US-09-539-802A-2680	Sequence 2680, Ap
45	224	97.8	239	20	US-09-075-782-3266	Sequence 3266, Ap
46	224	97.8	239	27	US-09-540-766-48565	Sequence 48565, A
47	224	97.8	239	74	US-60-046-624-3266	Sequence 3266, Ap
48	223	97.4	351	22	US-09-289-768-14321	Sequence 14321, A
49	223	97.4	351	37	US-09-939-397-14321	Sequence 14321, A
50	223	97.4	410	22	US-09-287-618-7514	Sequence 7514, Ap
51	223	97.4	430	25	US-09-528-409-9804	Sequence 9804, Ap
52	223	97.4	430	36	US-09-933-524-9804	Sequence 9804, Ap
53	223	97.4	430	36	US-09-933-524A-9804	Sequence 9804, Ap
54	223	97.4	449	31	US-09-758-472-3977	Sequence 3977, Ap
55	223	97.4	449	43	US-10-235-926-3977	Sequence 3977, Ap
56	223	97.4	495	12	US-08-220-691-11835	Sequence 11835, A
57	223	97.4	495	12	US-08-220-691B-11835	Sequence 11835, A
58	223	97.4	495	33	US-09-813-155-11835	Sequence 11835, A
59	223	97.4	495	35	US-09-912-293-209365	Sequence 209365,
60	223	97.4	799	23	US-09-399-932-4403	Sequence 4403, Ap
61	223	97.4	1858	42	US-10-170-235-24020	Sequence 24020, A
62	223	97.4	1858	79	US-60-452-680-2346	Sequence 2346, Ap
63	223	97.4	3042	3	PCT-US04-17686-664	Sequence 664, App
64	223	97.4	3042	50	US-10-362-463-5	Sequence 5, Appli
65	223	97.4	3042	50	US-10-362-463-6	Sequence 6, Appli
66	223	97.4	3183	50	US-10-362-463-9	Sequence 9, Appli
67	223	97.4	3183	50	US-10-362-463-10	Sequence 10, Appl
68	223	97.4	4961	61	US-10-775-169-139	Sequence 139, App
69	223	97.4	4961	79	US-60-446-133-139	Sequence 139, App
70	223	97.4	4999	3	PCT-US04-24424-2360	Sequence 2360, Ap
71	223	97.4	4999	50	US-10-362-463-1	Sequence 1, Appli
72	223	97.4	4999	50	US-10-362-463-2	Sequence 2, Appli
73	223	97.4	4999	80	US-60-490-890-2360	Sequence 2360, Ap
74	223	97.4	4999	82	US-60-671-501-143	Sequence 143, App
75	223	97.4	5144	50	US-10-362-463-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-779-004-1142

; Sequence 1142, Application US/08779004
; GENERAL INFORMATION:
; APPLICANT: Stuart, Susan G.
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Lane, John C.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Snable, James I.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN TONSIL
; NUMBER OF SEQUENCES: 1459
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,004
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/60/009,671
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: LUTHER, BARBARA J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PD 0100P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 852-0195
; INFORMATION FOR SEQ ID NO: 1142:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 737520
US-08-779-004-1142

Alignment Scores:

Pred. No.:	1.78e-22	Length:	250
Score:	229.00	Matches:	44
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	17	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-08-779-004-1142 (1-250)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 75

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 135

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        136 GTTCTTGAATAC 147
```

RESULT 2
US-09-540-233D-14754
; Sequence 14754, Application US/09540233D
; GENERAL INFORMATION:

```
;  APPLICANT: Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION: POLYNUCLEOTIDES OF HEMIC AND IMMUNE SYSTEM TISSUE
;  FILE REFERENCE: PD-1030 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,233D
;  CURRENT FILING DATE:  2000-03-31
;  PRIOR APPLICATION NUMBER: 09/008,119
;  PRIOR FILING DATE: 1998-01-16
;  PRIOR APPLICATION NUMBER: 8/100,523
;  PRIOR FILING DATE: 1993-08-03
;  PRIOR APPLICATION NUMBER: 07/977,780
;  PRIOR FILING DATE: 1992-11-19
;  PRIOR APPLICATION NUMBER: 07/916,491
;  PRIOR FILING DATE: 1992-07-17
;  PRIOR APPLICATION NUMBER: 08/438,571
;  PRIOR FILING DATE: 1995-05-10
;  PRIOR APPLICATION NUMBER: 08/282,991
;  PRIOR FILING DATE: 1994-07-28
;  PRIOR APPLICATION NUMBER: 08/196,364
;  PRIOR FILING DATE: 1994-02-14
;  PRIOR APPLICATION NUMBER: 08/504,732
;  PRIOR FILING DATE: 1995-07-20
;  PRIOR APPLICATION NUMBER: 08/179,873
;  PRIOR FILING DATE: 1994-01-11
;  PRIOR APPLICATION NUMBER: 08/435,761
;  PRIOR FILING DATE: 1995-05-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 129340
;  SOFTWARE: PERL Program
; SEQ ID NO 14754
;   LENGTH: 250
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00808005
US-09-540-233D-14754

Alignment Scores:
Pred. No.:                1.78e-22       Length:          250
Score:                    229.00         Matches:         44
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       26             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-540-233D-14754 (1-250)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 75

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 135

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        136 GTTCTTGAATAC 147

RESULT 3
US-60-009-671-1142
; Sequence 1142, Application US/60009671
;  GENERAL INFORMATION:
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Lane, John C.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Snable, James I.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  HUMAN TONSIL
;    NUMBER OF SEQUENCES:  1459
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/60/009,671
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LUTHER, BARBARA J.
;      REGISTRATION NUMBER:  33954
;      REFERENCE/DOCKET NUMBER:  PD 0100P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 852-0195
;  INFORMATION FOR SEQ ID NO:  1142:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  250 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  737520
US-60-009-671-1142

Alignment Scores:
Pred. No.:                1.78e-22       Length:          250
Score:                    229.00         Matches:         44
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       74             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-009-671-1142 (1-250)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 75

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 135

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        136 GTTCTTGAATAC 147

RESULT 4
US-09-105-427-308
; Sequence 308, Application US/09105427
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Naughton, Rebecca E.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  HUMAN PROSTATE WITH BENIGN PROSTATIC HYPERPLASIA
;    NUMBER OF SEQUENCES:  3842
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
```

```
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/105,427
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PD-0390P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 845-4166
;  INFORMATION FOR SEQ ID NO:  308:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  272 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  3361508H1
US-09-105-427-308

Alignment Scores:
Pred. No.:                2e-22          Length:          272
Score:                    229.00         Matches:         44
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       21             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-105-427-308 (1-272)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 119

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 179

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        180 GTTCTTGAATAC 191

RESULT 5
US-09-540-208-8004
; Sequence 8004, Application US/09540208
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF MALE REPRODUCTIVE TISSUE
;  FILE REFERENCE:  PD-1029 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,208
;  CURRENT FILING DATE:  2000-03-31
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS:  70811
;  SOFTWARE:  PERL Program
; SEQ ID NO 8004
;   LENGTH: 272
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00434261
US-09-540-208-8004

Alignment Scores:
Pred. No.:                2e-22          Length:          272
Score:                    229.00         Matches:         44
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       26             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-540-208-8004 (1-272)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 119

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 179

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        180 GTTCTTGAATAC 191

RESULT 6
US-10-793-479-1651
; Sequence 1651, Application US/10793479
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/10/793,479
;  CURRENT FILING DATE:  2004-03-03
;  PRIOR APPLICATION NUMBER: US/09/513,999
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 1651
;   LENGTH: 274
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 108..272
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 18
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 43
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 45
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 52
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 54
;   OTHER INFORMATION: k=g or t
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: 60
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 62
;   OTHER INFORMATION: y=c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 128
;   OTHER INFORMATION: r=a or g
US-10-793-479-1651

Alignment Scores:
Pred. No.:                2.02e-22       Length:         274
Score:                    229.00         Matches:        44
Percent Similarity:       100.0%         Conservative:   0
Best Local Similarity:    100.0%         Mismatches:     0
Query Match:              100.0%         Indels:         0
DB:                       61             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-793-479-1651 (1-274)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GGAGACCCGGGGTCGGARATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 170

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 230

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        231 GTTCTTGAATAC 242

RESULT 7
US-09-540-764-58131
; Sequence 58131, Application US/09540764
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF CELLS AND TISSUES OF THE FEMALE REPRODUCTIVE
;  FILE REFERENCE:  PD-1028 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,764
;  CURRENT FILING DATE:  2000-03-30
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS:  61458
;  SOFTWARE:  PERL Program
; SEQ ID NO 58131
;   LENGTH: 284
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu01215592
US-09-540-764-58131

Alignment Scores:
Pred. No.:                2.12e-22       Length:         284
Score:                    229.00         Matches:        44
Percent Similarity:       100.0%         Conservative:   0
Best Local Similarity:    100.0%         Mismatches:     0
Query Match:              100.0%         Indels:         0
DB:                       26             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-540-764-58131 (1-284)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 101

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 161

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        162 GTTCTTGAATAC 173

RESULT 8
US-10-349-781-58131
; Sequence 58131, Application US/10349781
; GENERAL INFORMATION:
;  APPLICANT: Seilhamer, Jeffrey J.; Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.; Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.; Naughton, Rebecca E.
;  TITLE OF INVENTION: POLYNUCLEOTIDES OF CELLS AND TISSUES OF THE FEMALE REPRODUCTIVE S
;  FILE REFERENCE: PD-1028-2 CON
;  CURRENT APPLICATION NUMBER: US/10/349,781
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: 09/540,764
;  PRIOR FILING DATE: 2000-03-30
;  PRIOR APPLICATION NUMBER: 08/992,868
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/032,838
;  PRIOR FILING DATE: 1996-12-13
;  PRIOR APPLICATION NUMBER: 08/734,050
;  PRIOR FILING DATE: 1996-10-18
;  PRIOR APPLICATION NUMBER: 60/006,111
;  PRIOR FILING DATE: 1995-10-24
;  PRIOR APPLICATION NUMBER: 08/733,814
;  PRIOR FILING DATE: 1996-10-18
;  PRIOR APPLICATION NUMBER: 60/005,864
;  PRIOR FILING DATE: 1995-10-26
;  PRIOR APPLICATION NUMBER: 08/763,920
;  PRIOR FILING DATE: 1996-12-11
;  PRIOR APPLICATION NUMBER: 60/008,794
;  PRIOR FILING DATE: 1995-12-14
;  PRIOR APPLICATION NUMBER: 08/772,783
;  PRIOR FILING DATE: 1996-12-23
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 61458
;  SOFTWARE: PERL Program
; SEQ ID NO 58131
;   LENGTH: 284
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu01215592
US-10-349-781-58131

Alignment Scores:
Pred. No.:                2.12e-22       Length:         284
Score:                    229.00         Matches:        44
Percent Similarity:       100.0%         Conservative:   0
Best Local Similarity:    100.0%         Mismatches:     0
Query Match:              100.0%         Indels:         0
DB:                       49             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-349-781-58131 (1-284)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 101

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 161

Qy         41 ValLeuGluTyr 44
              ||||||||||||
```

```
Db        162 GTTCTTGAATAC 173

RESULT 9
US-09-489-036-12694
; Sequence 12694, Application US/09489036
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION: From Various Libraries
;  FILE REFERENCE: 783
;  CURRENT APPLICATION NUMBER: US/09/489,036
;  CURRENT FILING DATE: 2000-01-19
;  NUMBER OF SEQ ID NOS: 35324
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 12694
;   LENGTH: 312
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(312)
;   OTHER INFORMATION: n = A,T,C or G
US-09-489-036-12694

Alignment Scores:
Pred. No.:              2.42e-22        Length:         312
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     24              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-489-036-12694 (1-312)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 112

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 172

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        173 GTTCTTGAATAC 184

RESULT 10
US-09-943-143-12694
; Sequence 12694, Application US/09943143
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 783
;  CURRENT APPLICATION NUMBER: US/09/943,143
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/489,036
;  PRIOR FILING DATE: 2000-01-19
;  NUMBER OF SEQ ID NOS: 35324
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 12694
;   LENGTH: 312
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(312)
;   OTHER INFORMATION: n = A,T,C or G
US-09-943-143-12694

Alignment Scores:
Pred. No.:              2.42e-22        Length:         312
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     37              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-943-143-12694 (1-312)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 112

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 172

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        173 GTTCTTGAATAC 184

RESULT 11
US-09-205-070-10288
; Sequence 10288, Application US/09205070
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
;  TITLE OF INVENTION: LIBRARIES
;  FILE REFERENCE: 20411-748
;  CURRENT APPLICATION NUMBER: US/09/205,070
;  CURRENT FILING DATE: 1998-12-03
;  NUMBER OF SEQ ID NOS: 45207
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10288
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(366)
;   OTHER INFORMATION: n = A,T,C or G
US-09-205-070-10288

Alignment Scores:
Pred. No.:              3.03e-22        Length:         366
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     22              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-205-070-10288 (1-366)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 102

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 162

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        163 GTTCTTGAATAC 174

RESULT 12
US-09-271-122-13443
; Sequence 13443, Application US/09271122
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION: FROM A cDNA LIBRARY OF HUMAN BONE MARROW AND HUMAN FETAL
;  TITLE OF INVENTION: BRAIN
;  FILE REFERENCE: 20411-731CON1
;  CURRENT APPLICATION NUMBER: US/09/271,122
```

```
;  CURRENT FILING DATE: 1999-03-17
;  EARLIER APPLICATION NUMBER: US 09/045,400
;  EARLIER FILING DATE: 1998-03-20
;  NUMBER OF SEQ ID NOS: 13540
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13443
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-271-122-13443

Alignment Scores:
Pred. No.:                3.03e-22        Length:         366
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       22              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-271-122-13443 (1-366)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 102

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 162

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        163 GTTCTTGAATAC 174

RESULT 13
US-09-340-623-10288
; Sequence 10288, Application US/09340623
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
;  TITLE OF INVENTION: LIBRARIES
;  FILE REFERENCE: 20411-748CON1
;  CURRENT APPLICATION NUMBER: US/09/340,623
;  CURRENT FILING DATE: 1999-06-28
;  EARLIER APPLICATION NUMBER: US 09/205,070
;  EARLIER FILING DATE: 1998-12-03
;  NUMBER OF SEQ ID NOS: 45207
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10288
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(366)
;   OTHER INFORMATION: n = A,T,C or G
US-09-340-623-10288

Alignment Scores:
Pred. No.:                3.03e-22        Length:         366
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       23              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-340-623-10288 (1-366)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 102

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 162

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        163 GTTCTTGAATAC 174

RESULT 14
US-09-713-841-13443
; Sequence 13443, Application US/09713841
; GENERAL INFORMATION:
;  APPLICANT:  Hyseq, Inc.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM A cDNA LIBRARY OF HUMAN BONE MARROW AND HUMAN FETAL
;  TITLE OF INVENTION:  BRAIN
;  FILE REFERENCE:  20411-731CON1
;  CURRENT APPLICATION NUMBER: US/09/713,841
;  CURRENT FILING DATE:  2000-11-15
;  PRIOR APPLICATION NUMBER:  US 09/045,400
;  PRIOR FILING DATE:  1998-03-20
;  NUMBER OF SEQ ID NOS:  13540
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 13443
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-713-841-13443

Alignment Scores:
Pred. No.:                3.03e-22        Length:         366
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       31              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-713-841-13443 (1-366)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 102

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 162

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        163 GTTCTTGAATAC 174

RESULT 15
US-09-898-888-10288
; Sequence 10288, Application US/09898888
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
;  TITLE OF INVENTION:  LIBRARIES
;  FILE REFERENCE: 20411-748CON1
;  CURRENT APPLICATION NUMBER: US/09/898,888
;  CURRENT FILING DATE:  2001-07-03
;  PRIOR APPLICATION NUMBER: 09/340,623
;  PRIOR FILING DATE: 1999-06-28
;  NUMBER OF SEQ ID NOS: 45207
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10288
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(366)
;   OTHER INFORMATION: n = A,T,C or G
US-09-898-888-10288
```

```
Alignment Scores:
Pred. No.:                3.03e-22        Length:         366
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       34              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-898-888-10288 (1-366)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 102

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 162

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        163 GTTCTTGAATAC 174

RESULT 16
US-09-898-888A-10288
; Sequence 10288, Application US/09898888A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
;  TITLE OF INVENTION:  LIBRARIES
;  FILE REFERENCE: 20411-748CON1
;  CURRENT APPLICATION NUMBER: US/09/898,888A
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/340,623
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: US 09/205,070
;  PRIOR FILING DATE: 1998-12-03
;  NUMBER OF SEQ ID NOS: 45207
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10288
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(366)
;   OTHER INFORMATION: n = A,T,C or G
US-09-898-888A-10288

Alignment Scores:
Pred. No.:                3.03e-22        Length:         366
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       34              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-898-888A-10288 (1-366)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 102

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 162

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        163 GTTCTTGAATAC 174

RESULT 17
US-09-927-875-13443
; Sequence 13443, Application US/09927875
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM A cDNA LIBRARY OF HUMAN BONE MARROW AND HUMAN FETAL
;  TITLE OF INVENTION:  BRAIN
;  FILE REFERENCE: 20411-731CON1
;  CURRENT APPLICATION NUMBER: US/09/927,875
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 09/713,841
;  PRIOR FILING DATE: 2000-11-15
;  NUMBER OF SEQ ID NOS: 13540
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13443
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-927-875-13443

Alignment Scores:
Pred. No.:                3.03e-22        Length:         366
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       36              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-927-875-13443 (1-366)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 102

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 162

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        163 GTTCTTGAATAC 174

RESULT 18
US-09-927-875A-13443
; Sequence 13443, Application US/09927875A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM A cDNA LIBRARY OF HUMAN BONE MARROW AND HUMAN FETAL
;  TITLE OF INVENTION:  BRAIN
;  FILE REFERENCE: 20411-731CON1
;  CURRENT APPLICATION NUMBER: US/09/927,875A
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 09/713,841
;  PRIOR FILING DATE: 2000-11-15
;  NUMBER OF SEQ ID NOS: 13540
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13443
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-927-875A-13443

Alignment Scores:
Pred. No.:                3.03e-22        Length:         366
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       36              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-927-875A-13443 (1-366)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          43 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 102

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         103 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 162

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         163 GTTCTTGAATAC 174
```

RESULT 19
US-10-302-689A-10288
; Sequence 10288, Application US/10302689A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Asundi, Vinod
; APPLICANT: Ballinger, Dennis
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; APPLICANT: Loeb, Deborah
; APPLICANT: Montgomery, Julia, R.
; APPLICANT: Pace, Ann M.
; APPLICANT: Sheridan, James P.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 502CIP
; CURRENT APPLICATION NUMBER: US/10/302,689A
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 10/084,643
; PRIOR FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: 09/989,660
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 10/014,487
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 09/952,981
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/905,059
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 09/898,888
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 09/919,002
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 158931
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 10288
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(366)
; OTHER INFORMATION: n = a,t,c or g
US-10-302-689A-10288

Alignment Scores:

Pred. No.:	3.03e-22	Length:	366
Score:	229.00	Matches:	44
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	48	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-302-689A-10288 (1-366)

```
Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          43 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 102

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         103 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 162

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         163 GTTCTTGAATAC 174
```

RESULT 20
US-09-489-036-20573
; Sequence 20573, Application US/09489036
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 783
; CURRENT APPLICATION NUMBER: US/09/489,036
; CURRENT FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 35324
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 20573
; LENGTH: 373
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-036-20573

Alignment Scores:

Pred. No.:	3.12e-22	Length:	373
Score:	229.00	Matches:	44
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	24	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-489-036-20573 (1-373)

```
Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          16 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 75

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 135

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         136 GTTCTTGAATAC 147
```

RESULT 21
US-09-943-143-20573
; Sequence 20573, Application US/09943143
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 783
; CURRENT APPLICATION NUMBER: US/09/943,143
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/489,036
; PRIOR FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 35324
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 20573
; LENGTH: 373
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-943-143-20573

Alignment Scores:

Pred. No.:	3.12e-22	Length:	373

```
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     37              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-943-143-20573 (1-373)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 75

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 135

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        136 GTTCTTGAATAC 147

RESULT 22
US-09-431-517-5512
; Sequence 5512, Application US/09431517
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-751CON1
;  CURRENT APPLICATION NUMBER: US/09/431,517
;  CURRENT FILING DATE: 1999-11-01
;  EARLIER APPLICATION NUMBER: US 09/170,294
;  EARLIER FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 31760
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5512
;   LENGTH: 418
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(418)
;   OTHER INFORMATION: n = A,T,C or G
US-09-431-517-5512

Alignment Scores:
Pred. No.:              3.65e-22        Length:         418
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     24              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-431-517-5512 (1-418)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         45 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 104

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 164

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        165 GTTCTTGAATAC 176

RESULT 23
US-09-431-517-5511
; Sequence 5511, Application US/09431517
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-751CON1
;  CURRENT APPLICATION NUMBER: US/09/431,517
;  CURRENT FILING DATE: 1999-11-01
;  EARLIER APPLICATION NUMBER: US 09/170,294
;  EARLIER FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 31760
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5511
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-431-517-5511

Alignment Scores:
Pred. No.:              3.68e-22        Length:         420
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     24              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-431-517-5511 (1-420)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         45 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 104

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 164

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        165 GTTCTTGAATAC 176

RESULT 24
PCT-US02-30474-2111
; Sequence 2111, Application PC/TUS0230474
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 810CIP PCT
;  CURRENT APPLICATION NUMBER: PCT/US02/30474
;  CURRENT FILING DATE:  2002-09-24
;  PRIOR APPLICATION NUMBER: US 60/324,631
;  PRIOR FILING DATE: 2001-09-24
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: PCT/US00/35017
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: PCT/US01/02623
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: US 09/496,914
```

```
;  PRIOR FILING DATE: 2000-02-03
;  PRIOR APPLICATION NUMBER: US 09/560,875
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: PCT/US01/03800
;  PRIOR FILING DATE: 2001-02-05
;  PRIOR APPLICATION NUMBER: US 09/515,126
;  PRIOR FILING DATE: 2000-02-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 3476
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 2111
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US02-30474-2111

Alignment Scores:
Pred. No.:                6.53e-22        Length:         632
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       1               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x PCT-US02-30474-2111 (1-632)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 278

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 338

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        339 GTTCTTGAATAC 350

RESULT 25
US-09-515-126-12696
; Sequence 12696, Application US/09515126
; GENERAL INFORMATION:
;  APPLICANT: Tang, Yuanhua T.
;  APPLICANT: Tillinghast, John
;  APPLICANT: Sinku, Ankura
;  APPLICANT: Liu, Chenghua
;  APPLICANT: Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Contigs Obtained
;  TITLE OF INVENTION: From Various Libraries
;  FILE REFERENCE: 788
;  CURRENT APPLICATION NUMBER: US/09/515,126
;  CURRENT FILING DATE: 2000-02-28
;  NUMBER OF SEQ ID NOS: 14074
;  SOFTWARE: pt_CT_genes Version 1.2
; SEQ ID NO 12696
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (174)...(403)
;   OTHER INFORMATION: similar to gi3043676 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-515-126-12696

Alignment Scores:
Pred. No.:                6.53e-22        Length:         632
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       25              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-515-126-12696 (1-632)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 278

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 338

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        339 GTTCTTGAATAC 350

RESULT 26
US-09-577-409-12696
; Sequence 12696, Application US/09577409
; GENERAL INFORMATION:
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Sinku, Ankura
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION:  Novel Nucleic Acids
;  TITLE OF INVENTION:  and Polypeptides
;  FILE REFERENCE:  788CIP
;  CURRENT APPLICATION NUMBER: US/09/577,409
;  CURRENT FILING DATE:  2000-05-18
;  PRIOR APPLICATION NUMBER:  09/515,126
;  PRIOR FILING DATE:  2000-02-28
;  NUMBER OF SEQ ID NOS:  14074
;  SOFTWARE:  pt_CT_genes Version 1.2
; SEQ ID NO 12696
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (174)...(403)
;   OTHER INFORMATION: similar to gi3043676 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-577-409-12696

Alignment Scores:
Pred. No.:                6.53e-22        Length:         632
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       27              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-577-409-12696 (1-632)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 278

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 338

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        339 GTTCTTGAATAC 350

RESULT 27
US-10-220-366A-12576
; Sequence 12576, Application US/10220366A
; GENERAL INFORMATION:
;  APPLICANT: HYSEQ, INC
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-042
;  CURRENT APPLICATION NUMBER: US/10/220,366A
```

```
;  CURRENT FILING DATE:  2002-08-28
;  PRIOR APPLICATION NUMBER: 09/577,409
;  PRIOR FILING DATE: 2000-05-18
;  PRIOR APPLICATION NUMBER: 09/515,126
;  PRIOR FILING DATE: 2001-02-23
;  NUMBER OF SEQ ID NOS: 27802
;  SOFTWARE: Custom
; SEQ ID NO 12576
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-220-366A-12576

Alignment Scores:
Pred. No.:              6.53e-22        Length:         632
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     43              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-220-366A-12576 (1-632)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 278

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 338

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        339 GTTCTTGAATAC 350

RESULT 28
US-10-302-689A-106285
; Sequence 106285, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ballinger, Dennis
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Liu, Jin
;  APPLICANT:  Loeb, Deborah
;  APPLICANT:  Montgomery, Julia, R.
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 106285
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-302-689A-106285

Alignment Scores:
Pred. No.:              6.53e-22        Length:         632
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     48              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-302-689A-106285 (1-632)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 278

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 338

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        339 GTTCTTGAATAC 350

RESULT 29
US-10-489-448-2111
; Sequence 2111, Application US/10489448
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 810CIP PCT
;  CURRENT APPLICATION NUMBER: US/10/489,448
;  CURRENT FILING DATE:  1004-03-10
;  PRIOR APPLICATION NUMBER: US 60/324,631
;  PRIOR FILING DATE: 2001-09-24
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: PCT/US00/35017
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: PCT/US01/02623
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: US 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  PRIOR APPLICATION NUMBER: US 09/560,875
;  PRIOR FILING DATE: 2000-04-27
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/03800
;  PRIOR FILING DATE: 2001-02-05
;  PRIOR APPLICATION NUMBER: US 09/515,126
;  PRIOR FILING DATE: 2000-02-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 3476
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 2111
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-489-448-2111

Alignment Scores:
Pred. No.:                6.53e-22        Length:         632
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       51              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-489-448-2111 (1-632)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 278

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 338

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        339 GTTCTTGAATAC 350

RESULT 30
US-60-324-631-2095
; Sequence 2095, Application US/60324631
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 810
;  CURRENT APPLICATION NUMBER: US/60/324,631
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: PCT/US00/35017
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: PCT/US01/02623
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: US 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  PRIOR APPLICATION NUMBER: US 09/560,875
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: PCT/US01/03800
;  PRIOR FILING DATE: 2001-02-05
;  PRIOR APPLICATION NUMBER: US 09/515,126
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: US 09/577,409
;  PRIOR FILING DATE: 2000-05-18
;  PRIOR APPLICATION NUMBER: PCT/US01/04927
;  PRIOR FILING DATE: 2001-02-26
;  PRIOR APPLICATION NUMBER: US 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 09/574,454
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: PCT/US01/04941
;  PRIOR FILING DATE: 2001-03-05
;  PRIOR APPLICATION NUMBER: US 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: US 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: US 09/552,929
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: US 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  PRIOR APPLICATION NUMBER: PCT/US01/08656
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: US 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  PRIOR APPLICATION NUMBER: PCT/US01/14827
;  PRIOR FILING DATE: 2001-05-16
;  NUMBER OF SEQ ID NOS: 3334
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 2095
;   LENGTH: 632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-324-631-2095

Alignment Scores:
Pred. No.:                6.53e-22        Length:         632
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       78              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-324-631-2095 (1-632)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 278

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 338

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        339 GTTCTTGAATAC 350

RESULT 31
US-60-172-377-1376
; Sequence 1376, Application US/60172377
; GENERAL INFORMATION:
;  APPLICANT: Morris, MacDonald
;  APPLICANT: Lal, Preeti
;  APPLICANT: Diep, Dinh
;  TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
;  TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymorp
;  FILE REFERENCE: GX-0008 P
;  CURRENT APPLICATION NUMBER: US/60/172,377
;  CURRENT FILING DATE: 1999-12-16
```

```
;  NUMBER OF SEQ ID NOS: 13,749
;  SOFTWARE: PERL Program
; SEQ ID NO 1376
;   LENGTH: 1399
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 445036.1
US-60-172-377-1376

Alignment Scores:
Pred. No.:                1.99e-21        Length:         1399
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       75              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-172-377-1376 (1-1399)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 140

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 200

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        201 GTTCTTGAATAC 212

RESULT 32
US-60-278-561-12578
; Sequence 12578, Application US/60278561
; GENERAL INFORMATION:
;  APPLICANT: Morris, MacDonald
;  APPLICANT:  Lal, Preeti
;  APPLICANT:  Diep, Dinh
;  TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
;  TITLE OF INVENTION:  Polynucleotide Sequence Databases, and Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms Identified Thereby
;  FILE REFERENCE: GX-0012-1 P
;  CURRENT APPLICATION NUMBER: US/60/278,561
;  CURRENT FILING DATE:  2001-03-23
;  NUMBER OF SEQ ID NOS: 15598
;  SOFTWARE: PERL Program
; SEQ ID NO 12578
;   LENGTH: 1399
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 445036.1
US-60-278-561-12578

Alignment Scores:
Pred. No.:                1.99e-21        Length:         1399
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       77              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-278-561-12578 (1-1399)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 140

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 200

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        201 GTTCTTGAATAC 212

RESULT 33
US-10-757-745-5
; Sequence 5, Application US/10757745
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 1536
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1530)..(1531)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (209)..(1534)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1392)..(1392)
;   OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-5

Alignment Scores:
Pred. No.:                2.27e-21        Length:         1536
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       61              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-757-745-5 (1-1536)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 271

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 331

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        332 GTTCTTGAATAC 343

RESULT 34
US-09-397-206A-5312
; Sequence 5312, Application US/09397206A
; GENERAL INFORMATION:
;  APPLICANT: Gearing, David P.
;  APPLICANT:  Holtzman, Douglas A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
;  TITLE OF INVENTION:  HUMAN FETAL SPLEEN LIBRARY
;  FILE REFERENCE: MLN98-41pM
;  CURRENT APPLICATION NUMBER: US/09/397,206A
```

```
;  CURRENT FILING DATE:  1999-09-17
;  PRIOR APPLICATION NUMBER: 60/101,133
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/107,254
;  PRIOR FILING DATE: 1998-11-05
;  PRIOR APPLICATION NUMBER: 60/126,842
;  PRIOR FILING DATE: 1999-03-30
;  NUMBER OF SEQ ID NOS: 6473
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5312
;   LENGTH: 1571
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1571)
;   OTHER INFORMATION: n = A,T,C or G
US-09-397-206A-5312

Alignment Scores:
Pred. No.:              2.34e-21      Length:        1571
Score:                  229.00        Matches:       44
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            100.0%        Indels:        0
DB:                     23            Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-397-206A-5312 (1-1571)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 129

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 189

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        190 GTTCTTGAATAC 201

RESULT 35
US-09-758-465-402
; Sequence 402, Application US/09758465
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PM006
;  CURRENT APPLICATION NUMBER: US/09/758,465
;  CURRENT FILING DATE:  2001-01-16
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 1276
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 402
;   LENGTH: 1697
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (25)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (1519)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (1555)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (1569)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (1615)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (1673)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (1685)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-758-465-402

Alignment Scores:
Pred. No.:              2.61e-21      Length:        1697
Score:                  229.00        Matches:       44
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            100.0%        Indels:        0
DB:                     31            Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-758-465-402 (1-1697)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 134

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 194

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        195 GTTCTTGAATAC 206

RESULT 36
US-10-222-927-402
; Sequence 402, Application US/10222927
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PM006C1N
;  CURRENT APPLICATION NUMBER: US/10/222,927
;  CURRENT FILING DATE:  2002-08-19
;  PRIOR APPLICATION NUMBER: 09/758,465
;  PRIOR FILING DATE: 2001-01-11
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 1276
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 402
;   LENGTH: 1697
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (25)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1519)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1555)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1569)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1615)
```

```
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1673)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1685)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-222-927-402

Alignment Scores:
Pred. No.:                2.61e-21        Length:         1697
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       43              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-222-927-402 (1-1697)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 134

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 194

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        195 GTTCTTGAATAC 206

RESULT 37
US-09-644-867-6750
; Sequence 6750, Application US/09644867
; GENERAL INFORMATION:
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Donovan, Michael J.
;  APPLICANT:  Leiby, Kevin R.
;  APPLICANT:  Culpepper, Janice A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1173-001
;  CURRENT APPLICATION NUMBER: US/09/644,867
;  CURRENT FILING DATE:  2000-08-28
;  PRIOR APPLICATION NUMBER:  60/151,061
;  PRIOR FILING DATE:  1999-08-27
;  NUMBER OF SEQ ID NOS:  8090
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6750
;   LENGTH: 2789
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2789)
;   OTHER INFORMATION: n = A,T,C or G
US-09-644-867-6750

Alignment Scores:
Pred. No.:                5.24e-21        Length:         2789
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       29              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-644-867-6750 (1-2789)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 129

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 189

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        190 GTTCTTGAATAC 201

RESULT 38
US-09-652-913-10806
; Sequence 10806, Application US/09652913
; GENERAL INFORMATION:
;  APPLICANT:  Falb, Dean R.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1174-001
;  CURRENT APPLICATION NUMBER: US/09/652,913
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER:  60/152,107
;  PRIOR FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  10833
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 10806
;   LENGTH: 2789
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2789)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-913-10806

Alignment Scores:
Pred. No.:                5.24e-21        Length:         2789
Score:                    229.00          Matches:        44
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       29              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-652-913-10806 (1-2789)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 129

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 189

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        190 GTTCTTGAATAC 201

RESULT 39
US-09-699-997-10714
; Sequence 10714, Application US/09699997
; GENERAL INFORMATION:
;  APPLICANT:  Curtis, Rory A.J.
;  APPLICANT:  Silos-Santiago, Inmaculada
;  APPLICANT:  DiStefano, Peter
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2020-001
;  CURRENT APPLICATION NUMBER: US/09/699,997
;  CURRENT FILING DATE:  2000-10-30
;  PRIOR APPLICATION NUMBER:  60/162,359
;  PRIOR FILING DATE:  1999-10-29
;  NUMBER OF SEQ ID NOS:  12714
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 10714
;   LENGTH: 2789
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2789)
;   OTHER INFORMATION: n = A,T,C or G
US-09-699-997-10714

Alignment Scores:
Pred. No.:              5.24e-21        Length:         2789
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     30              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-699-997-10714 (1-2789)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 129

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 189

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        190 GTTCTTGAATAC 201

RESULT 40
US-09-717-350-4792
; Sequence 4792, Application US/09717350
; GENERAL INFORMATION:
;  APPLICANT:  Gearing, David P.
;  APPLICANT:  Villeval, Jean-Luc
;  APPLICANT:  Fraser, Christopher C.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2034-001
;  CURRENT APPLICATION NUMBER: US/09/717,350
;  CURRENT FILING DATE:  2000-11-21
;  PRIOR APPLICATION NUMBER:  60/167,885
;  PRIOR FILING DATE:  1999-11-24
;  NUMBER OF SEQ ID NOS:  5955
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 4792
;   LENGTH: 2789
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2789)
;   OTHER INFORMATION: n = A,T,C or G
US-09-717-350-4792

Alignment Scores:
Pred. No.:              5.24e-21        Length:         2789
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-717-350-4792 (1-2789)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 129

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 189

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        190 GTTCTTGAATAC 201

RESULT 41
US-09-528-409-3837
; Sequence 3837, Application US/09528409
; GENERAL INFORMATION:
;  APPLICANT:  Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION:  Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE:  774
;  CURRENT APPLICATION NUMBER: US/09/528,409
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER:  60/125,453
;  PRIOR FILING DATE:  1999-03-19
;  NUMBER OF SEQ ID NOS:  116231
;  SOFTWARE:  Hy-patent.pl Version 3.1
; SEQ ID NO 3837
;   LENGTH: 457
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(457)
;   OTHER INFORMATION: n = A,T,C or G
US-09-528-409-3837

Alignment Scores:
Pred. No.:              1.57e-21        Length:         457
Score:                  225.00          Matches:        43
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  97.7%           Mismatches:     0
Query Match:            98.3%           Indels:         0
DB:                     25              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-528-409-3837 (1-457)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 290

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db        291 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTAAAGGACCATCGAGACCT 350

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        351 GTTCTTGAATAC 362

RESULT 42
US-09-933-524-3837
; Sequence 3837, Application US/09933524
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
```

```
; SEQ ID NO 3837
;   LENGTH: 457
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(457)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524-3837

Alignment Scores:
Pred. No.:              1.57e-21        Length:         457
Score:                  225.00          Matches:        43
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  97.7%           Mismatches:     0
Query Match:            98.3%           Indels:         0
DB:                     36              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-933-524-3837 (1-457)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 290

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        291 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTAAAGGACCATCGAGACCT 350

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        351 GTTCTTGAATAC 362

RESULT 43
US-09-933-524A-3837
; Sequence 3837, Application US/09933524A
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524A
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 3837
;   LENGTH: 457
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(457)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524A-3837

Alignment Scores:
Pred. No.:              1.57e-21        Length:         457
Score:                  225.00          Matches:        43
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  97.7%           Mismatches:     0
Query Match:            98.3%           Indels:         0
DB:                     36              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-933-524A-3837 (1-457)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 290

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        291 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTAAAGGACCATCGAGACCT 350

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        351 GTTCTTGAATAC 362

RESULT 44
US-09-539-802A-2680
; Sequence 2680, Application US/09539802A
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF MIXED TISSUES
;  FILE REFERENCE:  PD-1038 CIP
;  CURRENT APPLICATION NUMBER: US/09/539,802A
;  CURRENT FILING DATE:  2000-03-30
;  NUMBER OF SEQ ID NOS:  3017
;  SOFTWARE:  PERL Program
; SEQ ID NO 2680
;   LENGTH: 493
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu01150010
US-09-539-802A-2680

Alignment Scores:
Pred. No.:              1.74e-21        Length:         493
Score:                  225.00          Matches:        43
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  97.7%           Mismatches:     0
Query Match:            98.3%           Indels:         0
DB:                     26              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-539-802A-2680 (1-493)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         45 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 104

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        105 TCAACAACGGATGAAAATGAAGACGAAATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 164

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        165 GTTCTTGAATAC 176

RESULT 45
US-09-075-782-3266
; Sequence 3266, Application US/09075782
;  GENERAL INFORMATION:
;     APPLICANT:  Gooding, Douglas H.
;     APPLICANT:  Stuve, Laura L.
;     APPLICANT:  Stuart, Susan G.
;     APPLICANT:  Ito, Laura Y.
;     APPLICANT:  Akerblom, Ingrid E.
;     APPLICANT:  Delegeane, Angelo M.
;     APPLICANT:  Naughton, Rebecca E.
;     APPLICANT:  Klingler, Tod M.
;     TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;     TITLE OF INVENTION:  CHRONIC ULCERATIVE COLITIS
;     NUMBER OF SEQUENCES:  3537
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
```

```
;       STREET:  3174 PORTER DRIVE
;       CITY:  PALO ALTO
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/075,782
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/046,624
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CERRONE, MICHAEL C.
;       REGISTRATION NUMBER:  39,132
;       REFERENCE/DOCKET NUMBER:  PD-0371P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 855-0555
;       TELEFAX:  (415) 845-4166
;   INFORMATION FOR SEQ ID NO:  3266:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  239 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     IMMEDIATE SOURCE:
;       CLONE:  3236565H2
US-09-075-782-3266

Alignment Scores:
Pred. No.:                8.79e-22        Length:         239
Score:                    224.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.8%           Indels:         0
DB:                       20              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-075-782-3266 (1-239)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db         74 GGAGACCCGGGGTCGGAAATAATAGAATCTGNCCCTCCAGCTGGCCCTGAGGCATCTGAG 133

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 193

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        194 GTTCTTGAATAC 205

RESULT 46
US-09-540-766-48565
; Sequence 48565, Application US/09540766
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE
;  FILE REFERENCE:  PD-1024 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,766
;  CURRENT FILING DATE:  2000-03-30
;  PRIOR APPLICATION DATA removed - refer tp PALM or File Wrapper
;  NUMBER OF SEQ ID NOS:  77960
;  SOFTWARE:  PERL Program
; SEQ ID NO 48565
;   LENGTH: 239
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00108066
;   NAME/KEY: unsure
;   LOCATION: 105
;   OTHER INFORMATION: a, t, c, g, or other
US-09-540-766-48565

Alignment Scores:
Pred. No.:                8.79e-22        Length:         239
Score:                    224.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.8%           Indels:         0
DB:                       27              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-540-766-48565 (1-239)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db         74 GGAGACCCGGGGTCGGAAATAATAGAATCTGNCCCTCCAGCTGGCCCTGAGGCATCTGAG 133

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 193

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        194 GTTCTTGAATAC 205

RESULT 47
US-60-046-624-3266
; Sequence 3266, Application US/60046624
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuve, Laura L.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Naughton, Rebecca E.
;    APPLICANT:  Klingler, Tod M.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  CHRONIC ULCERATIVE COLITIS
;    NUMBER OF SEQUENCES:  3537
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/60/046,624
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PD-0371P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 845-4166
```

```
;  INFORMATION FOR SEQ ID NO:  3266:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  239 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  3236565H2
US-60-046-624-3266

Alignment Scores:
Pred. No.:                8.79e-22        Length:         239
Score:                    224.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.8%           Indels:         0
DB:                       74              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-046-624-3266 (1-239)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db         74 GGAGACCCGGGGTCGGAAATAATAGAATCTGNCCCTCCAGCTGGCCCTGAGGCATCTGAG 133

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 193

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        194 GTTCTTGAATAC 205

RESULT 48
US-09-289-768-14321
; Sequence 14321, Application US/09289768
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-765
;  CURRENT APPLICATION NUMBER: US/09/289,768
;  CURRENT FILING DATE: 1999-04-08
;  NUMBER OF SEQ ID NOS: 39996
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14321
;   LENGTH: 351
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-289-768-14321

Alignment Scores:
Pred. No.:                2.1e-21         Length:         351
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       22              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-289-768-14321 (1-351)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 94

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         95 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 154

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        155 GTTCTTGAATAC 166

RESULT 49
US-09-939-397-14321
; Sequence 14321, Application US/09939397
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-765
;  CURRENT APPLICATION NUMBER: US/09/939,397
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: 09/289,768
;  PRIOR FILING DATE: 1999-04-08
;  NUMBER OF SEQ ID NOS: 39996
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14321
;   LENGTH: 351
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-939-397-14321

Alignment Scores:
Pred. No.:                2.1e-21         Length:         351
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       37              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-939-397-14321 (1-351)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 94

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         95 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 154

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        155 GTTCTTGAATAC 166

RESULT 50
US-09-287-618-7514
; Sequence 7514, Application US/09287618
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-768
;  CURRENT APPLICATION NUMBER: US/09/287,618
;  CURRENT FILING DATE: 1999-04-02
;  NUMBER OF SEQ ID NOS: 35865
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7514
;   LENGTH: 410
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(410)
;   OTHER INFORMATION: n = A,T,C or G
US-09-287-618-7514

Alignment Scores:
Pred. No.:                2.61e-21        Length:         410
Score:                    223.00          Matches:        43
Percent Similarity:       97.7%           Conservative:   0
Best Local Similarity:    97.7%           Mismatches:     1
Query Match:              97.4%           Indels:         0
DB:                       22              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-287-618-7514 (1-410)
```

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 123

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        124 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 183

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        184 GTTCTTGAATAC 195
```

RESULT 51
US-09-528-409-9804
; Sequence 9804, Application US/09528409
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/528,409
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,453
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 9804
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-528-409-9804

Alignment Scores:

Pred. No.:	2.79e-21	Length:	430
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	25	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-528-409-9804 (1-430)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 137

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        138 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 197

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        198 GTTCTTGAATAC 209
```

RESULT 52
US-09-933-524-9804
; Sequence 9804, Application US/09933524
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 9804
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-933-524-9804

Alignment Scores:

Pred. No.:	2.79e-21	Length:	430
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	36	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-933-524-9804 (1-430)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 137

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        138 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 197

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        198 GTTCTTGAATAC 209
```

RESULT 53
US-09-933-524A-9804
; Sequence 9804, Application US/09933524A
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524A
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 9804
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-933-524A-9804

Alignment Scores:

Pred. No.:	2.79e-21	Length:	430
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	36	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-933-524A-9804 (1-430)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 137

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        138 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 197
```

```
Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         198 GTTCTTGAATAC 209

RESULT 54
US-09-758-472-3977
; Sequence 3977, Application US/09758472
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PH001
;  CURRENT APPLICATION NUMBER: US/09/758,472
;  CURRENT FILING DATE:  2001-01-11
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 9632
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3977
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-758-472-3977

Alignment Scores:
Pred. No.:              2.97e-21        Length:         449
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-758-472-3977 (1-449)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          55 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 114

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db         115 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 174

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         175 GTTCTTGAATAC 186

RESULT 55
US-10-235-926-3977
; Sequence 3977, Application US/10235926
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PH001C1N
;  CURRENT APPLICATION NUMBER: US/10/235,926
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: 09/758,472
;  PRIOR FILING DATE: 2001-01-11
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  NUMBER OF SEQ ID NOS: 9632
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3977
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-235-926-3977

Alignment Scores:
Pred. No.:              2.97e-21        Length:         449
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     43              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-235-926-3977 (1-449)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          55 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 114

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db         115 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 174

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         175 GTTCTTGAATAC 186

RESULT 56
US-08-220-691-11835
; Sequence 11835, Application US/08220691
;  GENERAL INFORMATION:
;    APPLICANT:  Haseltine, William, et. al.
;    TITLE OF INVENTION:  Human Genes, Sequences, and
;    TITLE OF INVENTION:  ExpressionProducts
;    NUMBER OF SEQUENCES:  12049
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi, Stewart, & Olstein
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 5.0
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/220,691
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Olstein, Elliot
;      REGISTRATION NUMBER:  24,025
;      REFERENCE/DOCKET NUMBER:  325800-71
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 994-1700
;      TELEFAX:  (201) 994-1744
;  INFORMATION FOR SEQ ID NO:  11835:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  495 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-220-691-11835

Alignment Scores:
Pred. No.:              3.4e-21         Length:         495
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     12              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-08-220-691-11835 (1-495)
```

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 114

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 174

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        175 GTTCTTGAATAC 186
```

RESULT 57
US-08-220-691B-11835
; Sequence 11835, Application US/08220691B
; GENERAL INFORMATION:
; APPLICANT: Haseltine, William A.
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; TITLE OF INVENTION: Human Genes, Sequences and Expression Products
; NUMBER OF SEQUENCES: 12051
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,691B
; FILING DATE: March 31, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PO-12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 11835:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-220-691B-11835

Alignment Scores:

Pred. No.:	3.4e-21	Length:	495
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-08-220-691B-11835 (1-495)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 114

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 174

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        175 GTTCTTGAATAC 186
```

RESULT 58
US-09-813-155-11835
; Sequence 11835, Application US/09813155
; GENERAL INFORMATION:
; APPLICANT: Haseltine, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
; FILE REFERENCE: PO-12C2
; CURRENT APPLICATION NUMBER: US/09/813,155
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/741,830
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 08/220,691
; PRIOR FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 12051
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11835
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (376)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (422)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (440)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (444)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (459)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (461)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (488)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-813-155-11835

Alignment Scores:

Pred. No.:	3.4e-21	Length:	495
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	33	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-813-155-11835 (1-495)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 114

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 174

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        175 GTTCTTGAATAC 186
```

RESULT 59
US-09-912-293-209365

```
; Sequence 209365, Application US/09912293
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 100
;  FILE REFERENCE: PO-100
;  CURRENT APPLICATION NUMBER: US/09/912,293
;  CURRENT FILING DATE:  2001-07-26
;  PRIOR APPLICATION NUMBER: 08/103,744
;  PRIOR FILING DATE: 1993-08-09
;  PRIOR APPLICATION NUMBER: 09/249,651
;  PRIOR FILING DATE: 1999-02-12
;  PRIOR APPLICATION NUMBER: 08/104,507
;  PRIOR FILING DATE: 1993-08-09
;  PRIOR APPLICATION NUMBER: 08/196,363
;  PRIOR FILING DATE: 1994-02-15
;  PRIOR APPLICATION NUMBER: 09/859,490
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: 08/196,362
;  PRIOR FILING DATE: 1994-02-15
;  PRIOR APPLICATION NUMBER: 08/221,623
;  PRIOR FILING DATE: 1994-03-31
;  PRIOR APPLICATION NUMBER: 08/220,691
;  PRIOR FILING DATE: 1994-03-31
;  PRIOR APPLICATION NUMBER: 09/741,830
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 09/813,155
;  PRIOR FILING DATE: 2001-03-21
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 244538
; SEQ ID NO 209365
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (376)..(376)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (422)..(422)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (440)..(440)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (444)..(444)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (459)..(459)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (461)..(461)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (488)..(488)
;   OTHER INFORMATION: n is equal to a,t,g, or c
US-09-912-293-209365

Alignment Scores:
Pred. No.:                3.4e-21        Length:          495
Score:                    223.00         Matches:         43
Percent Similarity:       97.7%          Conservative:    0
Best Local Similarity:    97.7%          Mismatches:      1
Query Match:              97.4%          Indels:          0
DB:                       35             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-912-293-209365 (1-495)
```

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 114

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 174

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        175 GTTCTTGAATAC 186

RESULT 60
US-09-399-932-4403
; Sequence 4403, Application US/09399932
; GENERAL INFORMATION:
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Gearing, David P.
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES DERIVED FROM A
;  TITLE OF INVENTION:  HUMAN BURKITT'S LYMPHOMA LIBRARY
;  FILE REFERENCE:  MLN98-47pM
;  CURRENT APPLICATION NUMBER: US/09/399,932
;  CURRENT FILING DATE:  1999-09-15
;  PRIOR APPLICATION NUMBER:  60/100,464
;  PRIOR FILING DATE:  1998-09-15
;  PRIOR APPLICATION NUMBER:  60/101,670
;  PRIOR FILING DATE:  1998-09-24
;  PRIOR APPLICATION NUMBER:  60/106,456
;  PRIOR FILING DATE:  1998-10-30
;  PRIOR APPLICATION NUMBER:  60/126,902
;  PRIOR FILING DATE:  1999-03-30
;  PRIOR APPLICATION NUMBER:  60/132,067
;  PRIOR FILING DATE:  1999-04-30
;  PRIOR APPLICATION NUMBER:  60/144,447
;  PRIOR FILING DATE:  1999-07-16
;  NUMBER OF SEQ ID NOS:  5743
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 4403
;   LENGTH: 799
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-399-932-4403

Alignment Scores:
Pred. No.:                6.66e-21       Length:          799
Score:                    223.00         Matches:         43
Percent Similarity:       97.7%          Conservative:    0
Best Local Similarity:    97.7%          Mismatches:      1
Query Match:              97.4%          Indels:          0
DB:                       23             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-09-399-932-4403 (1-799)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 252

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 312

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        313 GTTCTTGAATAC 324

RESULT 61
US-10-170-235-24020
; Sequence 24020, Application US/10170235
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig
;  TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
;  TITLE OF INVENTION:  TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
;  FILE REFERENCE: CL001380
```

```
;  CURRENT APPLICATION NUMBER: US/10/170,235
;  CURRENT FILING DATE:  2003-03-17
;  NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 24020
;   LENGTH: 1858
;   TYPE: DNA
;   ORGANISM: HUMAN
US-10-170-235-24020

Alignment Scores:
Pred. No.:              2.18e-20        Length:         1858
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     42              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-170-235-24020 (1-1858)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 241

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        242 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 301

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        302 GTTCTTGAATAC 313

RESULT 62
US-60-452-680-2346
; Sequence 2346, Application US/60452680
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  GRUPE, Andrew
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001450
;  CURRENT APPLICATION NUMBER: US/60/452,680
;  CURRENT FILING DATE:  2003-03-07
;  NUMBER OF SEQ ID NOS: 116213
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2346
;   LENGTH: 1858
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-452-680-2346

Alignment Scores:
Pred. No.:              2.18e-20        Length:         1858
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     79              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-452-680-2346 (1-1858)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 241

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        242 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 301

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        302 GTTCTTGAATAC 313

RESULT 63
PCT-US04-17686-664
; Sequence 664, Application PC/TUS0417686
; GENERAL INFORMATION:
;  APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY
;  APPLICANT:  THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES, CENTERS
;  APPLICANT:  FOR DISEASE CONTROL AND PREVENTION
;  TITLE OF INVENTION: PNI MICROARRAY AND USES
;  FILE REFERENCE: 14114.0375P1
;  CURRENT APPLICATION NUMBER: PCT/US04/17686
;  CURRENT FILING DATE:  2004-06-21
;  PRIOR APPLICATION NUMBER: 60/475,915
;  PRIOR FILING DATE: 2003-06-04
;  NUMBER OF SEQ ID NOS: 3085
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 664
;   LENGTH: 3042
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US04-17686-664

Alignment Scores:
Pred. No.:              4.35e-20        Length:         3042
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x PCT-US04-17686-664 (1-3042)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135

RESULT 64
US-10-362-463-5
; Sequence 5, Application US/10362463
; GENERAL INFORMATION:
;  APPLICANT: Akzo Nobel N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 3042
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-5

Alignment Scores:
Pred. No.:              4.35e-20        Length:         3042
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     50              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-5 (1-3042)
```

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135
```

RESULT 65
US-10-362-463-6
; Sequence 6, Application US/10362463
; GENERAL INFORMATION:
; APPLICANT: Akzo Nobel N.V.
; TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
; FILE REFERENCE: 2000,569WO
; CURRENT APPLICATION NUMBER: US/10/362,463
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: EP202905.6
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: EP1201771.1
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-362-463-6

Alignment Scores:

Pred. No.:	4.35e-20	Length:	3042
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	50	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-6 (1-3042)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135
```

RESULT 66
US-10-362-463-9
; Sequence 9, Application US/10362463
; GENERAL INFORMATION:
; APPLICANT: Akzo Nobel N.V.
; TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
; FILE REFERENCE: 2000,569WO
; CURRENT APPLICATION NUMBER: US/10/362,463
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: EP202905.6
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: EP1201771.1
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3183
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-362-463-9

Alignment Scores:

Pred. No.:	4.64e-20	Length:	3183
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	50	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-9 (1-3183)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135
```

RESULT 67
US-10-362-463-10
; Sequence 10, Application US/10362463
; GENERAL INFORMATION:
; APPLICANT: Akzo Nobel N.V.
; TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
; FILE REFERENCE: 2000,569WO
; CURRENT APPLICATION NUMBER: US/10/362,463
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: EP202905.6
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: EP1201771.1
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3183
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-362-463-10

Alignment Scores:

Pred. No.:	4.64e-20	Length:	3183
Score:	223.00	Matches:	43
Percent Similarity:	97.7%	Conservative:	0
Best Local Similarity:	97.7%	Mismatches:	1
Query Match:	97.4%	Indels:	0
DB:	50	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-10 (1-3183)

```
Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         64 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 123

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        124 GTTCTTGAATAC 135
```

RESULT 68
US-10-775-169-139
; Sequence 139, Application US/10775169
; GENERAL INFORMATION:
; APPLICANT: Wyeth

```
;  APPLICANT:  Burczynski, Michael
;  APPLICANT:  Twine, Natalie
;  APPLICANT:  Dorner, Andrew
;  APPLICANT:  Trepicchio, William
;  TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
;  FILE REFERENCE: AM101080 (031896-013000)
;  CURRENT APPLICATION NUMBER: US/10/775,169
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 5278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 139
;   LENGTH: 4961
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-775-169-139

Alignment Scores:
Pred. No.:              8.64e-20        Length:         4961
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     61              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-775-169-139 (1-4961)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 106

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 166

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        167 GTTCTTGAATAC 178

RESULT 69
US-60-446-133-139
; Sequence 139, Application US/60446133
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Burczynski, Michael E
;  APPLICANT:  Twine, Natalie E
;  APPLICANT:  Dorner, Andrew E
;  APPLICANT:  Trepicchio, William E
;  TITLE OF INVENTION: Method for Monitoring Anti-Cancer Drug Activities
;  FILE REFERENCE: AM101055-2
;  CURRENT APPLICATION NUMBER: US/60/446,133
;  CURRENT FILING DATE:  2003-02-11
;  NUMBER OF SEQ ID NOS: 386
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 139
;   LENGTH: 4961
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-446-133-139

Alignment Scores:
Pred. No.:              8.64e-20        Length:         4961
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     79              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-446-133-139 (1-4961)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 106

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 166

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        167 GTTCTTGAATAC 178

RESULT 70
PCT-US04-24424-2360
; Sequence 2360, Application PC/TUS0424424
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PCT
;  CURRENT APPLICATION NUMBER: PCT/US04/24424
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US 60/490,890
;  PRIOR FILING DATE: 2003-07-29
;  NUMBER OF SEQ ID NOS: 2786
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2360
;   LENGTH: 4999
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US04-24424-2360

Alignment Scores:
Pred. No.:              8.74e-20        Length:         4999
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x PCT-US04-24424-2360 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 71
US-10-362-463-1
; Sequence 1, Application US/10362463
; GENERAL INFORMATION:
;  APPLICANT: Akzo Nobel N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;. PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4999
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-1
```

```
Alignment Scores:
Pred. No.:              8.74e-20        Length:         4999
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     50              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-1 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 72
US-10-362-463-2
; Sequence 2, Application US/10362463
; GENERAL INFORMATION:
;  APPLICANT: Akzo Nobel N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4999
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-2

Alignment Scores:
Pred. No.:              8.74e-20        Length:         4999
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     50              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-2 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 73
US-60-490-890-2360
; Sequence 2360, Application US/60490890
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PSP
;  CURRENT APPLICATION NUMBER: US/60/490,890
;  CURRENT FILING DATE:  2003-07-29
;  NUMBER OF SEQ ID NOS: 2779
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2360
;   LENGTH: 4999
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-490-890-2360

Alignment Scores:
Pred. No.:              8.74e-20        Length:         4999
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     80              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-490-890-2360 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 74
US-60-671-501-143
; Sequence 143, Application US/60671501
; GENERAL INFORMATION:
;  APPLICANT: Wim Van Criekinge
;  APPLICANT:  Josef Straub
;  TITLE OF INVENTION: METHYLATION MARKERS FOR DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF CANCERS
;  FILE REFERENCE: 006791.00001
;  CURRENT APPLICATION NUMBER: US/60/671,501
;  CURRENT FILING DATE:  2005-04-15
;  NUMBER OF SEQ ID NOS: 420
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 143
;   LENGTH: 4999
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-671-501-143

Alignment Scores:
Pred. No.:              8.74e-20        Length:         4999
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     82              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-671-501-143 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348
```

```
Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         349 GTTCTTGAATAC 360

RESULT 75
US-10-362-463-3
; Sequence 3, Application US/10362463
; GENERAL INFORMATION:
;  APPLICANT: Akzo Nobel N.V.
;  TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
;  FILE REFERENCE: 2000,569WO
;  CURRENT APPLICATION NUMBER: US/10/362,463
;  CURRENT FILING DATE:  2003-02-21
;  PRIOR APPLICATION NUMBER: EP202905.6
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: EP1201771.1
;  PRIOR FILING DATE: 2001-05-14
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 5144
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-362-463-3

Alignment Scores:
Pred. No.:                9.09e-20        Length:          5144
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       50              Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-362-463-3 (1-5144)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               |||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db         289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         349 GTTCTTGAATAC 360


Search completed: March 11, 2006, 12:08:39
Job time : 1818.32 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:22:00 ; Search time 531.106 Seconds
(without alignments)
1013.401 Million cell updates/sec

Title: US-10-757-745-2_COPY_121_347
Perfect score: 1192
Sequence: 1 IDGLDLNNLSERARGVCSYL.....HIIPRSLDLLGLEKLDCGRF 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5007642 seqs, 1185514400 residues

Total number of hits satisfying chosen parameters: 10015284

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170335_4211/app_query.fasta_1
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_240_@runat_09032006_170335_4211 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq1:*
9: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1192	100.0	1288	9	US-11-266-748A-226184	Sequence 226184,
2	1192	100.0	1898	6	US-10-917-503B-13209	Sequence 13209, A
3	1192	100.0	1940	6	US-10-567-867-921	Sequence 921, App
4	1192	100.0	1940	11	US-60-742-219-2071	Sequence 2071, Ap
5	642.5	53.9	849	9	US-11-266-748A-166646	Sequence 166646,
6	622.5	52.2	1417	6	US-10-276-817B-5136	Sequence 5136, Ap
7	568	47.7	752	6	US-10-917-503B-4908	Sequence 4908, Ap
8	520	43.6	1179	9	US-11-266-748A-19660	Sequence 19660, A
9	395	33.1	1205	9	US-11-266-748A-184613	Sequence 184613,
10	395	33.1	1205	9	US-11-266-748A-192113	Sequence 192113,
11	218	18.3	1000	9	US-11-266-748A-117422	Sequence 117422,
c 12	218	18.3	1000	9	US-11-266-748A-159586	Sequence 159586,
13	218	18.3	1000	9	US-11-266-748A-289180	Sequence 289180,
c 14	218	18.3	1000	9	US-11-266-748A-340609	Sequence 340609,
15	201	16.9	1643	6	US-10-953-349-4575	Sequence 4575, Ap
16	158.5	13.3	525	9	US-11-292-078-14543	Sequence 14543, A
17	97.5	8.2	2928	6	US-10-540-898-785	Sequence 785, App
18	97.5	8.2	47040	6	US-10-540-898-784	Sequence 784, App
c 19	97	8.1	6198	6	US-10-868-184C-10580	Sequence 10580, A
c 20	97	8.1	6198	6	US-10-868-184D-10580	Sequence 10580, A
21	94.5	7.9	1767	6	US-10-998-197A-33	Sequence 33, Appl
22	94.5	7.9	1769	6	US-10-540-898-788	Sequence 788, App
23	94.5	7.9	49355	6	US-10-540-898-787	Sequence 787, App
24	94	7.9	1731	9	US-11-266-748A-184394	Sequence 184394,
25	94	7.9	1731	9	US-11-266-748A-191899	Sequence 191899,
26	93.5	7.8	1000	9	US-11-266-748A-283521	Sequence 283521,
c 27	93.5	7.8	1000	9	US-11-266-748A-310161	Sequence 310161,
c 28	93.5	7.8	228835	6	US-10-540-898-398	Sequence 398, App
29	92.5	7.8	2300	11	US-60-762-056-19596	Sequence 19596, A
c 30	92	7.7	6198	6	US-10-868-184C-10303	Sequence 10303, A
c 31	92	7.7	6198	6	US-10-868-184D-10303	Sequence 10303, A
c 32	91	7.6	6075	6	US-10-868-184C-10659	Sequence 10659, A
c 33	91	7.6	6075	6	US-10-868-184D-10659	Sequence 10659, A
c 34	90.5	7.6	5310	6	US-10-868-184C-10289	Sequence 10289, A
c 35	90.5	7.6	5310	6	US-10-868-184D-10289	Sequence 10289, A
c 36	89.5	7.5	1414	6	US-10-953-349-37498	Sequence 37498, A
c 37	89.5	7.5	6063	6	US-10-868-184C-10371	Sequence 10371, A
c 38	89.5	7.5	6063	6	US-10-868-184D-10371	Sequence 10371, A
c 39	88.5	7.4	5381	6	US-10-868-184C-10470	Sequence 10470, A
c 40	88.5	7.4	5381	6	US-10-868-184C-10513	Sequence 10513, A
c 41	88.5	7.4	5381	6	US-10-868-184D-10470	Sequence 10470, A
c 42	88.5	7.4	5381	6	US-10-868-184D-10513	Sequence 10513, A
c 43	88.5	7.4	5671	6	US-10-868-184C-10537	Sequence 10537, A
c 44	88.5	7.4	5671	6	US-10-868-184D-10537	Sequence 10537, A
c 45	88.5	7.4	5870	6	US-10-868-184C-10398	Sequence 10398, A
c 46	88.5	7.4	5870	6	US-10-868-184D-10398	Sequence 10398, A
47	88.5	7.4	70708	6	US-10-540-898-582	Sequence 582, App
48	88	7.4	3958	11	US-60-742-219-1685	Sequence 1685, Ap
c 49	88	7.4	6093	6	US-10-868-184C-10476	Sequence 10476, A
c 50	88	7.4	6093	6	US-10-868-184D-10476	Sequence 10476, A
c 51	87	7.3	5691	6	US-10-868-184C-10429	Sequence 10429, A
c 52	87	7.3	5691	6	US-10-868-184D-10429	Sequence 10429, A
c 53	87	7.3	10367	6	US-10-868-184C-10437	Sequence 10437, A
c 54	87	7.3	10367	6	US-10-868-184D-10437	Sequence 10437, A
c 55	86.5	7.3	5647	6	US-10-868-184C-10373	Sequence 10373, A
c 56	86.5	7.3	5647	6	US-10-868-184D-10373	Sequence 10373, A
c 57	86.5	7.3	36783	6	US-10-868-184C-10392	Sequence 10392, A
c 58	86.5	7.3	36783	6	US-10-868-184D-10392	Sequence 10392, A
c 59	86.5	7.3	49375	6	US-10-868-184C-10277	Sequence 10277, A
c 60	86.5	7.3	49375	6	US-10-868-184D-10277	Sequence 10277, A
c 61	86.5	7.3	2944528	9	US-11-045-004-1	Sequence 1, Appli
62	86	7.2	2300	11	US-60-762-056-9276	Sequence 9276, Ap
c 63	86	7.2	6044	6	US-10-868-184C-10381	Sequence 10381, A
c 64	86	7.2	6044	6	US-10-868-184D-10381	Sequence 10381, A
c 65	86	7.2	6065	6	US-10-868-184C-10280	Sequence 10280, A
c 66	86	7.2	6065	6	US-10-868-184D-10280	Sequence 10280, A
c 67	86	7.2	6495	6	US-10-868-184C-10386	Sequence 10386, A
c 68	86	7.2	6495	6	US-10-868-184D-10386	Sequence 10386, A
c 69	86	7.2	14411	6	US-10-868-184C-10688	Sequence 10688, A
c 70	86	7.2	14411	6	US-10-868-184D-10688	Sequence 10688, A
c 71	86	7.2	256164	9	US-11-266-748A-60910	Sequence 60910, A
72	85.5	7.2	1349	9	US-11-266-748A-184392	Sequence 184392,
c 73	85.5	7.2	1575	6	US-10-461-673-2142	Sequence 2142, Ap
c 74	85.5	7.2	6060	6	US-10-868-184C-10629	Sequence 10629, A
c 75	85.5	7.2	6060	6	US-10-868-184D-10629	Sequence 10629, A

ALIGNMENTS

RESULT 1

```
US-11-266-748A-226184
; Sequence 226184, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 226184
;   LENGTH: 1288
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-226184

Alignment Scores:
Pred. No.:                2.32e-153      Length:          1288
Score:                    1192.00        Matches:         227
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       9              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-226184 (1-1288)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 272

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 332

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 392

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 452

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 512

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 572

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 632

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 692

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 752

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 812

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        813 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 872

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db        873 AAACTGGACTGTGGTAGATTT 893

RESULT 2
US-10-917-503B-13209
; Sequence 13209, Application US/10917503B
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
;   LENGTH: 1898
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1086)
US-10-917-503B-13209

Alignment Scores:
Pred. No.:                4.31e-153      Length:          1898
Score:                    1192.00        Matches:         227
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       6              Gaps:            0
```

```
US-10-757-745-2_COPY_121_347 (1-227) x US-10-917-503B-13209 (1-1898)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTT 1041

RESULT 3
US-10-567-867-921
; Sequence 921, Application US/10567867
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PCT
;  CURRENT APPLICATION NUMBER: US/10/567,867
;  CURRENT FILING DATE:  2006-01-27
;  PRIOR APPLICATION NUMBER: US 60/490,890
;  PRIOR FILING DATE: 2003-07-29
;  NUMBER OF SEQ ID NOS: 2786
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-567-867-921

Alignment Scores:
Pred. No.:              4.47e-153       Length:         1940
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-567-867-921 (1-1940)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066

RESULT 4
US-60-742-219-2071
; Sequence 2071, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
```

```
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2071
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-2071

Alignment Scores:
Pred. No.:              4.47e-153       Length:         1940
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-742-219-2071 (1-1940)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066

RESULT 5
```

```
US-11-266-748A-166646
; Sequence 166646, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 166646
;   LENGTH: 849
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-166646

Alignment Scores:
Pred. No.:              7.27e-78        Length:         849
Score:                  642.50          Matches:        133
Percent Similarity:     90.1%           Conservative:   3
Best Local Similarity:  88.1%           Mismatches:     10
Query Match:            53.9%           Indels:         5
DB:                     9               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-166646 (1-849)

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal---SerGlyAsnGluLeu 100
              ||||||   :::     |||   |||||||||    :::       |||||||||
Db          8 ACCAAAAAGATAGAGGAACTTGGATGTGTGCATGAGCAACTGTGTCAGGACAATGAGCTA 67

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 127

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 187

Qy        141 GlyAspThrAsnLeuArgAspArg-GluValThr-ArgCysGlyGlyLeuProAsnAsnI 160
              |||||||||||||||||||||| :::|||||    |||||||||||||||||||||
Db        188 GGAGATACAAATCTAAGGGATCGAGCAGGTTACCAGCATGTGGTGGTTTACCCAACAACA 247

Qy        160 leValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp-AspThr 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db        248 TTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGCATACA 307

Qy        180 GlnMet-AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePh 199
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CAAATGCAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATT 367

Qy        199 ePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLe 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 TTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATT 427
```

```
Qy        219 uGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||||||||
Db        428 AGAAAAACTGGACTGTGGTAGATTT 452

RESULT 6
US-10-276-817B-5136
; Sequence 5136, Application US/10276817B
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-104
;  CURRENT APPLICATION NUMBER: US/10/276,817B
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: PCT/US01/14827
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 16102
;  SOFTWARE: Custom
; SEQ ID NO 5136
;   LENGTH: 1417
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (48)..(422)
;   OTHER INFORMATION: 79% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=494.
US-10-276-817B-5136

Alignment Scores:
Pred. No.:              9.46e-75      Length:          1417
Score:                  622.50        Matches:         130
Percent Similarity:     92.3%         Conservative:    1
Best Local Similarity:  91.5%         Mismatches:      6
Query Match:            52.2%         Indels:          5
DB:                     6             Gaps:            1

US-10-757-745-2_COPY_121_347 (1-227) x US-10-276-817B-5136 (1-1417)

Qy         91 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 110
              :::  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 CTGCAGGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACC 107

Qy        111 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 AGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAG 167

Qy        131 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluVal 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTT 227

Qy        151 ThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 ACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCT 287

Qy        171 LysHisCysGln-TyrThr-TrpAspThrGlnMetAsn---SerAsnLeuGlyIleThrA 189
              ||||||||    |||||| ||||||||||||          ||||||||||||||
Db        288 AAACATTGCCACGTATACATTGGGATACACAAACTGAACTCCTAATCTTGGGAATAACTG 347

Qy        189 la-AlaCysLysLeuArgPheAspArgIlePhePheArg-AlaAlaAlaGluGluGlyHi 208
              || ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        348 CCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAAGCAGCAGCAGAAGAGGGACA 407

Qy        208 sIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 CATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 465

RESULT 7
US-10-917-503B-4908
; Sequence 4908, Application US/10917503B
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4908
;   LENGTH: 752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (685)..(685)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (694)..(694)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (740)..(740)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-10-917-503B-4908

Alignment Scores:
Pred. No.:              1.13e-67      Length:          752
Score:                  568.00        Matches:         119
Percent Similarity:     90.9%         Conservative:    1
Best Local Similarity:  90.2%         Mismatches:      9
Query Match:            47.7%         Indels:          3
DB:                     6             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-917-503B-4908 (1-752)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540
```

```
Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy          81 SerThrLysMetMetArg-AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
               |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         601 AGTACCAAAATGATGAGAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 660

Qy         100 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 120
               |||||||   ||||||   |||    :::|||   ||||||      ||||||||||||
Db         661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy         120 lnLeuLysMetValLeu-LysLysMetGln 129
               ||   ||||||   |||    |||||||||
Db         721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750

RESULT 8
US-11-266-748A-19660
; Sequence 19660, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19660
;   LENGTH: 1179
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-19660

Alignment Scores:
Pred. No.:              9.65e-61        Length:         1179
Score:                  520.00          Matches:        96
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            43.6%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-19660 (1-1179)

Qy         132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 151
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACC 60

Qy         152 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 171
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 AGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAA 120

Qy         172 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 191
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGT 180

Qy         192 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 211
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 240

Qy         212 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
               ||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 288

RESULT 9
US-11-266-748A-184613
; Sequence 184613, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 184613
;   LENGTH: 1205
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (64)..(81)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-184613

Alignment Scores:
Pred. No.:              1.74e-43        Length:         1205
Score:                  395.00          Matches:        75
Percent Similarity:     98.7%           Conservative:   3
Best Local Similarity:  94.9%           Mismatches:     1
Query Match:            33.1%           Indels:         1
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-184613 (1-1205)

Qy         149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
               :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         134 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 193

Qy         169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 188
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         194 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 253

Qy         189 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 208
               |||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db         254 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAAAGCAGCAGCAGAAGAGGGACAC 313

Qy         209 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
               ||||||||||||||||||   ||||||||||||:::||||||||||||||||||||||
```

```
Db        314 ATTATTCCCCGAAGTTTGGC-CTTCTTGGATTAAAAAAACTGGACTGTGGTAGATTT 369

RESULT 10
US-11-266-748A-192113
; Sequence 192113, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 192113
;   LENGTH: 1205
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (64)..(81)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-192113

Alignment Scores:
Pred. No.:              1.74e-43      Length:          1205
Score:                  395.00        Matches:         75
Percent Similarity:     98.7%         Conservative:    3
Best Local Similarity:  94.9%         Mismatches:      1
Query Match:            33.1%         Indels:          1
DB:                     9             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-192113 (1-1205)

Qy        149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 193

Qy        169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 253

Qy        189 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 208
              ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db        254 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAAAGCAGCAGCAGAAGAGGGACAC 313

Qy        209 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||    ||||||||||:::|||||||||||||||||||||||
Db        314 ATTATTCCCCGAAGTTTGGC-CTTCTTGGATTAAAAAAACTGGACTGTGGTAGATTT 369

RESULT 11
US-11-266-748A-117422
; Sequence 117422, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 117422
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-117422

Alignment Scores:
Pred. No.:              3.32e-19      Length:          1000
Score:                  218.00        Matches:         42
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            18.3%         Indels:          0
DB:                     9             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-117422 (1-1000)

Qy        186 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 60

Qy        206 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGly 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 120

Qy        226 ArgPhe 227
              ||||||
Db        121 AGATTT 126

RESULT 12
US-11-266-748A-159586/c
; Sequence 159586, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
```

```
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 159586
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-159586

Alignment Scores:
Pred. No.:                3.32e-19        Length:         1000
Score:                    218.00          Matches:        42
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              18.3%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-159586 (1-1000)

Qy         186 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 205
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1000 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 941

Qy         206 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGly 225
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         940 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 881

Qy         226 ArgPhe 227
               ||||||
Db         880 AGATTT 875

RESULT 13
US-11-266-748A-289180
; Sequence 289180, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 289180
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-289180

Alignment Scores:
Pred. No.:                3.32e-19        Length:         1000
Score:                    218.00          Matches:        42
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              18.3%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-289180 (1-1000)

Qy         186 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 205
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 60

Qy         206 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGly 225
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 120

Qy         226 ArgPhe 227
               ||||||
Db         121 AGATTT 126

RESULT 14
US-11-266-748A-340609/c
; Sequence 340609, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 340609
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-340609

Alignment Scores:
Pred. No.:                3.32e-19        Length:         1000
Score:                    218.00          Matches:        42
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              18.3%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-340609 (1-1000)

Qy         186 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 205
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1000 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 941
```

```
Qy        206 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGly 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 881

Qy        226 ArgPhe 227
              ||||||
Db        880 AGATTT 875

RESULT 15
US-10-953-349-4575
; Sequence 4575, Application US/10953349
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCONDED THERBY
;  FILE REFERENCE: 2750-1579PUS2
;  CURRENT APPLICATION NUMBER: US/10/953,349
;  CURRENT FILING DATE:  2004-09-30
;  NUMBER OF SEQ ID NOS: 40252
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4575
;   LENGTH: 1643
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-10-953-349-4575

Alignment Scores:
Pred. No.:                1.63e-16        Length:          1643
Score:                    201.00          Matches:         69
Percent Similarity:       46.3%           Conservative:    32
Best Local Similarity:    31.7%           Mismatches:      85
Query Match:              16.9%           Indels:          32
DB:                       6               Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-953-349-4575 (1-1643)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
              :::   |||:::||||||             |||   |||   :::       :::
Db        577 GAAGACCTTGAGTTGAAT---------CTAAGGATGAGAGCGATTGGTCATCTTATTCAG 627

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
              |||:::||||||   :::|||       |||||||||   |||       |||
Db        628 CTTCATTCTCCTCATCTCATATGTTTTCAGGAAGTCACTCCTGAGATATATGATATTTTC 687

Qy         42 LysLysArg-----------------SerSerAsnTyrGluIleIleThrGlyHisGlu 55
              :::|||                   |||   :::       :::
Db        688 CGCAAGTCGAATTGGTGGAAAGCGTATTCTTGTTCTGTCTCAGTTGATGTGGCAGTTTCG 747

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||:::      ::::::          |||::::::   :::||||||
Db        748 AGAGGCTACTACTGCATGCTGTTG---------AGCAAGTTGGGAGTGAAATCTTTCAGC 798

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
                 |||   ::::::   |||   |||   |||           |||   |||
Db        799 AGCAAAAGCTTTGGGAACTCGATAATGGGAAGAGAACTATCTATCGCAGAAGTTGAAGTT 858

Qy         96 SerGlyAsnGlu---LeuCysLeuMetThrSerHisLeuGluSer--------------- 109
              |||   :::   |||           ||||||||||||||||||
Db        859 CCAGGGAGAAAGCCACTTGTTTTTGCAACTAGCCATCTTGAGAGTCCCTGTCCAGGTCCT 918

Qy        110 ---------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
                      :::       :::   ||||||:::   |||   |||   ::::::
Db        919 CCAAAATGGGACCAAATGTTCAGTCGAGAACGCGTTGAACAAGCCAAGGAGGCAATTGAG 978

Qy        127 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 146
              ::::::           :::|||   |||||||||   ||||||   |||
Db        979 ATTCTCAGA---------CCTAATGCCAATGTGATATTCGGGGGAGACATGAACTGGTGC 1029

Qy        147 AspArgGluValThrArgCysGlyGly---------LeuProAsnAsnIleValAspVal 163
              |||           :::       |||         ||||||:::       |||||||||
Db       1030 GAC-----------AAGCTAGATGGGAAGTTTCCTTTGCCAGACAAATGGGTGGATGTT 1077

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
                 ||||||   |||   ||||||           :::|||:::||||||:::   |||
Db       1078 TGGGAGGTTTTG---AAACCAGGTGATTTGGGTTTTACATATGACACAAAAGCAAACCCA 1134

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
              |||                 |||   :::   |||   |||||||||       |||
Db       1135 ATGTTGTCCGGTAACCGGGCATTGCAGAAACGGTTAGACCGGATTTTGTGTCGT 1188

RESULT 16
US-11-292-078-14543
; Sequence 14543, Application US/11292078
; GENERAL INFORMATION:
;  APPLICANT: Deikman, Jill
;  APPLICANT:  Feng, Paul C.C.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Ziegler, Todd E.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.346 - 38-21(52274)C
;  CURRENT APPLICATION NUMBER: US/11/292,078
;  CURRENT FILING DATE:  2005-12-02
;  PRIOR APPLICATION NUMBER: US 10/021,323
;  PRIOR FILING DATE: 2001-12-12
;  PRIOR APPLICATION NUMBER: US 60/255,619
;  PRIOR FILING DATE: 2000-12-14
;  NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 14543
;   LENGTH: 525
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3829-034-Q1-K6-G12
US-11-292-078-14543

Alignment Scores:
Pred. No.:                1.9e-11         Length:          525
Score:                    158.50          Matches:         53
Percent Similarity:       46.1%           Conservative:    29
Best Local Similarity:    29.8%           Mismatches:      63
Query Match:              13.3%           Indels:          33
DB:                       9               Gaps:            7

US-10-757-745-2_COPY_121_347 (1-227) x US-11-292-078-14543 (1-525)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              :::|||   :::   :::         :::   |||:::|||||||||:::|||
Db         19 AAGAGGATGAAAGCTATTGGTGACCTTATACAACTTCATTCCCCTGATATCATTTGTTTT 78

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||   |||       |||             |||   ||||||:::
Db         79 CAAGAGGTTACTCCGATGATTTATGACATTTTT------CGTGGGTCTAATTGG------ 126

Qy         51 IleThrGlyHisGluGluGlyTyrPheThrAlaIle---------------------- 62
                        :::||||||       :::|||
Db        127 -----------TGGAAAGGGTACCGGTGCTCTATCTCGGACAACACGGCGAGTTTGAGA 174

Qy         63 ---------MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 79
                      ::::::   |||::::::   :::|||||             ||||||
Db        175 GCTTACTTTTGTATACAGTTAAGCAAATTGCCTGTGAAATCCTTTCGTCGTGAACCTTTT 234

Qy         80 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 99
              ::::::   |||   |||   |||           |||   |||||||||   |||
Db        235 GATAATTCGATAATGGGGCGAGAACTTTGCATGACTGAGGTTGAAGTTTCTGGGGGTGAA 294

Qy        100 ------LeuCysLeuMetThrSerHisLeuGluSer----------------------- 109
                    |||   :::   |||:::||||||||||||
Db        295 ACAACGCTTGTTGTTGCTACTACCCATCTCGAGAGTCCGTGTCCTGCTCCGCCGAAATGG 354

Qy        110 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 129
              :::   ::::::   ||||||:::   |||   |||:::   :::       ::::::
Db        355 GATCAAATGTATAGTAAGGAACGTGTAGAACAAGCGAAAGTAGCGATCAATATTCTCGAG 414

Qy        130 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 149
                    |||             |||||||||   ||||||   |||       |||:::
```

```
Db        415 AATAACCCG---------AATGTTATTTTTGGGGGTGATATGAATTGGGATGACAAATTA 465

Qy        150 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
Db        466 GATGGTCGATTT---CCTTTAACTGATGGATGGATTGATGCTTGGAACGAGTTA 516

RESULT 17
US-10-540-898-785
; Sequence 785, Application US/10540898
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: CHIR0056-101 (PP023367.0003)
;  CURRENT APPLICATION NUMBER: US/10/540,898
;  CURRENT FILING DATE:  2005-06-27
;  PRIOR APPLICATION NUMBER: US 10/330,773
;  PRIOR FILING DATE: 2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785
;   LENGTH: 2928
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2928)
;   OTHER INFORMATION: n = A,T,C or G
US-10-540-898-785

Alignment Scores:
Pred. No.:              0.0772          Length:         2928
Score:                  97.50           Matches:        51
Percent Similarity:     37.1%           Conservative:   40
Best Local Similarity:  20.8%           Mismatches:     111
Query Match:            8.2%            Indels:         43
DB:                     6               Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-540-898-785 (1-2928)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
Db       1041 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 1100

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
Db       1101 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 1148

Qy         51 IleThrGlyHisGluGluGlyTyrPhe---------------------------------- 59
Db       1149 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 1205

Qy         60 ----------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
Db       1206 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 1265

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
Db       1266 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 1325

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
Db       1326 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 1385

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
Db       1386 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 1445

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
Db       1446 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 1505

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
Db       1506 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 1565

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
Db       1566 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 1613

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
Db       1614 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 1667

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuGluGlyLeuGluLys 221
Db       1668 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 1718

Qy        222 LeuAspCysGlyArg 226
Db       1719 ATTGGGCCCAACCGG 1733

RESULT 18
US-10-540-898-784
; Sequence 784, Application US/10540898
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: CHIR0056-101 (PP023367.0003)
;  CURRENT APPLICATION NUMBER: US/10/540,898
;  CURRENT FILING DATE:  2005-06-27
;  PRIOR APPLICATION NUMBER: US 10/330,773
;  PRIOR FILING DATE: 2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 784
;   LENGTH: 47040
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(47040)
;   OTHER INFORMATION: n = A,T,C or G
US-10-540-898-784

Alignment Scores:
Pred. No.:              6.58            Length:         47040
Score:                  97.50           Matches:        51
Percent Similarity:     37.1%           Conservative:   40
Best Local Similarity:  20.8%           Mismatches:     111
Query Match:            8.2%            Indels:         43
DB:                     6               Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-540-898-784 (1-47040)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
Db     35153 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 35212

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
Db     35213 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 35260

Qy         51 IleThrGlyHisGluGluGlyTyrPhe---------------------------------- 59
Db     35261 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 35317

Qy         60 ----------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
Db     35318 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 35377

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
Db     35378 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 35437
```

```
Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                       ||||||   :::   |||:::      :::||||||:::
Db      35438 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 35497

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              :::     |||         |||     |||         :::|||:::   :::
Db      35498 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 35557

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
              |||         :::   |||   :::     |||       ||||||   |||
Db      35558 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 35617

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                 |||                                 |||:::       :::
Db      35618 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 35677

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
              :::   |||         |||:::   :::         ||||||
Db      35678 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 35725

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                    :::     :::     |||:::       |||   |||::::::
Db      35726 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 35779

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                  |||               :::     :::|||||||||||     |||:::
Db      35780 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 35830

Qy        222 LeuAspCysGlyArg 226
              :::     |||
Db      35831 ATTGGGCCCAACCGG 35845

RESULT 19
US-10-868-184C-10580/c
; Sequence 10580, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10580
;   LENGTH: 6198
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10580

Alignment Scores:
Pred. No.:              0.301          Length:        6198
Score:                  97.00          Matches:       51
Percent Similarity:     47.2%          Conservative:  42
Best Local Similarity:  25.9%          Mismatches:    68
Query Match:            8.1%           Indels:        36
DB:                     6              Gaps:          9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10580 (1-6198)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ::::::||||||          |||   |||   :::          :::   :::::::::
Db       4045 GTAAATGGACTA---------AATGCTTCAATTAAAAGACACAGACTGGCAAATTGGATA 3995

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                        |||   |||       :::||||||                  ::::::
Db       3994 AAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCAT 3953

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||     |||   ::::::         |||   |||       :::   |||
Db       3952 CTCATGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCA 3893

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                      ||||||:::   |||   :::   :::   :::          |||
Db       3892 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGCCTCTGATAAAACAGATTTTAAA 3833

Qy         81 SerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnValSer 96
                 ||||||:::   |||            :::   :::   |||   |||::::::
Db       3832 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGTATCAATTCAA 3773

Qy         97 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
              ||||||   ::::::   :::          :::|||||          |||
Db       3772 CAAGAAGAATTAACTATCCTAAATACATGTGCACCCAATACAAGA---------GCATCC 3722

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||         :::|||   ||||||      :::|||          :::|||   |||
Db       3721 AGATTC------ATAAAGCAAGTCCTTAGTGACCTACAAAGAGACTTAGACTCCCACACA 3668

Qy        137 ValIlePheAlaGlyAsp------------------------ThrAsnLeuArgAspAr 148
              :::|||     ||||||                       |||:::   ||||||||
Db       3667 ATAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAG 3611

Qy        148 gGluValThrArgCys-GlyGlyLeu--------------ProAsn-AsnIleValAsp 162
              |:::::||||||     |||:::               |||::: :::::::|||
Db       3610 AAAGTTAACAAGGATACCCAGGAATTCAACTCAGCTCTGCACCAAGTGGACCTAATAGAC 3551

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::::::      |||      ||||||      :::|||||::
Db       3550 ATCTACAGAATTCTC---CACCCCAAATCAACAGAATATACATTT 3509

RESULT 20
US-10-868-184D-10580/c
; Sequence 10580, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
```

```
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10580
;   LENGTH: 6198
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10580

Alignment Scores:
Pred. No.:              0.301          Length:         6198
Score:                  97.00          Matches:        51
Percent Similarity:     47.2%          Conservative:   42
Best Local Similarity:  25.9%          Mismatches:     68
Query Match:            8.1%           Indels:         36
DB:                     6              Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10580 (1-6198)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ::::::|||||||          |||   |||   :::          :::   :::::::::
Db       4045 GTAAATGGACTA---------AATGCTTCAATTAAAAGACACAGACTGGCAAATTGGATA 3995

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                     |||   |||          :::||||||                  ::::::
Db       3994 AAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCAT 3953

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||     |||   ::::::              |||   |||       :::   |||
Db       3952 CTCATGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCA 3893

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                    ||||||:::   |||   :::   :::   :::            |||
Db       3892 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGCCTCTGATAAAACAGATTTTAAA 3833

Qy         81 SerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnValSer 96
              ||||||:::   |||             :::   :::   |||   |||::::::
Db       3832 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGTATCAATTCAA 3773

Qy         97 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                    ||||||   ::::::   :::          :::|||||||           |||
Db       3772 CAAGAAGAATTAACTATCCTAAATACATGTGCACCCAATACAAGA---------GCATCC 3722

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||          :::|||   ||||||      :::|||          :::|||   |||
Db       3721 AGATTC------ATAAAGCAAGTCCTTAGTGACCTACAAAGAGACTTAGACTCCCACACA 3668

Qy        137 ValIlePheAlaGlyAsp------------------------ThrAsnLeuArgAspAr 148
              :::|||     ||||||                       |||:::   ||||||||
Db       3667 ATAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAG 3611

Qy        148 gGluValThrArgCys-GlyGlyLeu---------------ProAsn-AsnIleValAsp 162
              |:::::||||||       |||:::                |||::: :::::::|||
Db       3610 AAAGTTAACAAGGATACCCAGGAATTCAACTCAGCTCTGCACCAAGTGGACCTAATAGAC 3551

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::::::     |||       ||||||    :::|||||:::
Db       3550 ATCTACAGAATTCTC---CACCCCAAATCAACAGAATATACATTT 3509

RESULT 21
US-10-998-197A-33
; Sequence 33, Application US/10998197A
; GENERAL INFORMATION:
;  APPLICANT: Dohrmann, Cord
;  APPLICANT:  Austen, Matthias
;  TITLE OF INVENTION: Pancreas-Specific Proteins
;  FILE REFERENCE: 2923-669
;  CURRENT APPLICATION NUMBER: US/10/998,197A
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: PCT/EP03/05700
;  PRIOR FILING DATE: 2003-05-30
;  PRIOR APPLICATION NUMBER: EP 02020829.4
;  PRIOR FILING DATE: 2002-09-17
;  PRIOR APPLICATION NUMBER: EP 02011963.2
;  PRIOR FILING DATE: 2002-05-29
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
;   LENGTH: 1767
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-998-197A-33

Alignment Scores:
Pred. No.:              0.0891         Length:         1767
Score:                  94.50          Matches:        50
Percent Similarity:     37.1%          Conservative:   41
Best Local Similarity:  20.4%          Mismatches:     111
Query Match:            7.9%           Indels:         43
DB:                     6              Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-998-197A-33 (1-1767)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              ||||||          :::          :::      |||   ||||||::::::   |||
Db        713 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 772

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              ||||||||           |||:::      :::          |||
Db        773 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 820

Qy         51 IleThrGlyHisGluGluGlyTyrPhe-------------------------------- 59
                    |||::::::      :::|||
Db        821 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 877

Qy         60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                  |||:::      :::::::|||   ||||||
Db        878 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 937

Qy         73 SerGlnGluIleIleProProProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
              :::          |||          :::      :::       |||   :::   :::
Db        938 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 997

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                               ||||||   :::   |||:::       :::|||||||:::
Db        998 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 1057

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              :::          |||          |||          :::|||:::   :::
Db       1058 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 1117

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
              |||          :::   |||   :::      |||      ||||||   |||
Db       1118 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 1177

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                   |||                                 |||:::          :::
Db       1178 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 1237

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
              :::   |||         |||:::   :::              ||||||
Db       1238 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 1285

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                        :::      :::      |||:::          |||   |||::::::
Db       1286 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 1339

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                   |||              :::      :::|||||||||||||          |||:::
Db       1340 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 1390

Qy        222 LeuAspCysGlyArg 226
              :::         |||
```

```
Db       1391 ATTGGACCCAACAGG 1405

RESULT 22
US-10-540-898-788
; Sequence 788, Application US/10540898
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: CHIR0056-101 (PP023367.0003)
;  CURRENT APPLICATION NUMBER: US/10/540,898
;  CURRENT FILING DATE:  2005-06-27
;  PRIOR APPLICATION NUMBER: US 10/330,773
;  PRIOR FILING DATE: 2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 788
;   LENGTH: 1769
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-540-898-788

Alignment Scores:
Pred. No.:              0.0893      Length:          1769
Score:                  94.50       Matches:         50
Percent Similarity:     37.1%       Conservative:    41
Best Local Similarity:  20.4%       Mismatches:      111
Query Match:            7.9%        Indels:          43
DB:                     6           Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-540-898-788 (1-1769)

Qy          11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
               ||||||         :::        :::      |||   ||||||::::::   |||
Db         713 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 772

Qy          31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
               |||||||||         |||:::        :::        |||
Db         773 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 820

Qy          51 IleThrGlyHisGluGluGlyTyrPhe---------------------------------- 59
                  |||::::::      :::|||
Db         821 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 877

Qy          60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                   |||:::        :::::|||   ||||||
Db         878 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 937

Qy          73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
               :::        |||       :::      :::      |||   :::   :::
Db         938 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 997

Qy          92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                        ||||||   :::   |||:::      :::||||||:::
Db         998 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 1057

Qy         109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
               :::      |||       |||      |||        :::|||:::   :::
Db        1058 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 1117

Qy         129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
               |||        :::   |||   :::      |||      ||||||   |||
Db        1118 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 1177

Qy         147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                  |||                                  |||:::        :::
Db        1178 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 1237

Qy         165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
               :::   |||         |||:::   :::              ||||||
Db        1238 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 1285

Qy         182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                        :::      :::      |||:::         |||   |||::::::
Db        1286 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 1339

Qy         202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                     |||                 :::      :::||||||||||||      |||:::
Db        1340 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 1390

Qy         222 LeuAspCysGlyArg 226
               :::         |||
Db        1391 ATTGGACCCAACAGG 1405

RESULT 23
US-10-540-898-787
; Sequence 787, Application US/10540898
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: CHIR0056-101 (PP023367.0003)
;  CURRENT APPLICATION NUMBER: US/10/540,898
;  CURRENT FILING DATE:  2005-06-27
;  PRIOR APPLICATION NUMBER: US 10/330,773
;  PRIOR FILING DATE: 2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 787
;   LENGTH: 49355
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(49355)
;   OTHER INFORMATION: n = A,T,C or G
US-10-540-898-787

Alignment Scores:
Pred. No.:              18.4        Length:          49355
Score:                  94.50       Matches:         50
Percent Similarity:     37.1%       Conservative:    41
Best Local Similarity:  20.4%       Mismatches:      111
Query Match:            7.9%        Indels:          43
DB:                     6           Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-540-898-787 (1-49355)

Qy          11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
               ||||||         :::        :::      |||   ||||||::::::   |||
Db      38299 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 38358

Qy          31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
               |||||||||         |||:::        :::        |||
Db      38359 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 38406

Qy          51 IleThrGlyHisGluGluGlyTyrPhe---------------------------------- 59
                  |||::::::      :::|||
Db      38407 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 38463

Qy          60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                   |||:::        :::::|||   ||||||
Db      38464 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 38523

Qy          73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
               :::        |||       :::      :::      |||   :::   :::
Db      38524 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 38583

Qy          92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                        ||||||   :::   |||:::      :::||||||:::
Db      38584 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 38643

Qy         109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
               :::      |||       |||      |||        :::|||:::   :::
Db      38644 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 38703
```

```
Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
              |||         :::   |||   :::      |||      ||||||   |||
Db      38704 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 38763

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              |||                                 |||:::          :::
Db      38764 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 38823

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
              :::   |||         |||:::   :::             ||||||
Db      38824 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 38871

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                    :::     :::      |||:::         |||   |||::::::
Db      38872 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 38925

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
              |||           :::     :::|||||||||||||        |||:::
Db      38926 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 38976

Qy        222 LeuAspCysGlyArg 226
              :::      |||
Db      38977 ATTGGACCCAACAGG 38991
```

RESULT 24
US-11-266-748A-184394
; Sequence 184394, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 184394
; LENGTH: 1731
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1568)..(1578)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-184394

Alignment Scores:

Pred. No.:	0.101	Length:	1731
Score:	94.00	Matches:	63
Percent Similarity:	37.7%	Conservative:	37
Best Local Similarity:	23.8%	Mismatches:	96
Query Match:	7.9%	Indels:	70
DB:	9	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-184394 (1-1731)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||      |||  :::|||                  :::||||||
Db        167 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 226

Qy         33 ValIleProProTyr---------------------------------------TyrSer 39
              :::      |||                                          |||
Db        227 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 286

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::      |||          :::||||||      ||||||       ||
Db        287 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 331

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
              ||||       :::   :::|||||||||          :::|||   |||
Db        332 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 391

Qy         76 ------------------------------------------------IleIleProPhePr 80
                                                             |||:::      ||
Db        392 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 451

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::     ||||||         ::::::|||:::  :::  ::
Db        452 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 505

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::||||||          |||       ||
Db        506 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 564

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |      :::|||   :::|||   :::|||                 ::::::   ||
Db        565 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 618

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||                  ||||||:::||||||      |||           ||
Db        619 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 660

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |   ::::::   :::   |||          |||   :::|||      |||
Db        661 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 720

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
              :::|||       :::
Db        721 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 780

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||          ::::::||||||   :::            |||
Db        781 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 834

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                    |||   ::::::   |||   |||:::                |||
Db        835 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 894

Qy        221 sLeuAspCysGly 225
               |||   ||||||
Db        895 TCTAATTTGTGGT 907
```

RESULT 25
US-11-266-748A-191899
; Sequence 191899, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03

```
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 191899
;   LENGTH: 1731
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1568)..(1578)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-191899

Alignment Scores:
Pred. No.:                0.101          Length:          1731
Score:                    94.00          Matches:         63
Percent Similarity:       37.7%          Conservative:    37
Best Local Similarity:    23.8%          Mismatches:      96
Query Match:              7.9%           Indels:          70
DB:                       9              Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-191899 (1-1731)

Qy          13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
               |||:::|||      |||   :::|||                :::||||||
Db         167 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 226

Qy          33 ValIleProProTyr-------------------------------------TyrSer 39
               :::      |||                                         |||
Db         227 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 286

Qy          40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
               |||:::     |||              :::||||||     ||||||       ||
Db         287 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 331

Qy          59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
               ||||      :::   :::|||||||||         :::|||   |||
Db         332 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 391

Qy          76 ----------------------------------------------IleIleProPhePr 80
                                                            |||:::      ||
Db         392 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 451

Qy          80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
               |:::|||      :::      ||||||         ::::::|||:::   :::   ::
Db         452 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 505

Qy         100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
               :||||||:::      ||||||:::||||||              |||          ||
Db         506 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 564

Qy         120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
               |         :::|||   :::|||   :::|||                 ::::::   ||
Db         565 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 618

Qy         140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
               ||||              ||||||:::||||||        |||              ||
Db         619 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 660

Qy         160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
               |   ::::::   :::   |||            |||   :::|||      |||
Db         661 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 720

Qy         178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
               :::|||            :::
Db         721 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 780

Qy         187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||         ::::::||||||   :::                |||
Db         781 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 834

Qy         204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                    |||   ::::::   |||   |||:::                  |||
Db         835 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 894

Qy         221 sLeuAspCysGly 225
               |||   ||||||
Db         895 TCTAATTTGTGGT 907

RESULT 26
US-11-266-748A-283521
; Sequence 283521, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 283521
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-283521

Alignment Scores:
Pred. No.:                0.0492         Length:          1000
Score:                    93.50          Matches:         50
Percent Similarity:       37.1%          Conservative:    41
Best Local Similarity:    20.4%          Mismatches:      111
Query Match:              7.8%           Indels:          43
DB:                       9              Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-283521 (1-1000)

Qy          11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
               ||||||          :::      :::      |||   ||||||::::::   |||
Db         173 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 232

Qy          31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
               |||||||||          |||:::      :::          |||
Db         233 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 280

Qy          51 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 59
                  |||::::::      :::|||
```

```
Db        281 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 337

Qy         60 ---------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys  72

Db        338 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 397

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal---  91

Db        398 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGACCATTGCA 457

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108

Db        458 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 517

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128

Db        518 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 577

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146

Db        578 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 637

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164

Db        638 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 697

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181

Db        698 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 745

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201

Db        746 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 799

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221

Db        800 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 850

Qy        222 LeuAspCysGlyArg 226

Db        851 ATTGGACCCAACAGG 865
```

RESULT 27
US-11-266-748A-310161/c
; Sequence 310161, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 310161
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-310161

Alignment Scores:

Pred. No.:	0.0492	Length:	1000
Score:	93.50	Matches:	50
Percent Similarity:	37.1%	Conservative:	41
Best Local Similarity:	20.4%	Mismatches:	111
Query Match:	7.8%	Indels:	43
DB:	9	Gaps:	9

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-310161 (1-1000)

```
Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu  30

Db        828 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 769

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle  50

Db        768 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 721

Qy         51 IleThrGlyHisGluGluGlyTyrPhe-------------------------------- 59

Db        720 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 664

Qy         60 ---------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys  72

Db        663 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 604

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal---  91

Db        603 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGACCATTGCA 544

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108

Db        543 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 484

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128

Db        483 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 424

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146

Db        423 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 364

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164

Db        363 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 304

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181

Db        303 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 256

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201

Db        255 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 202

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221

Db        201 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 151

Qy        222 LeuAspCysGlyArg 226

Db        150 ATTGGACCCAACAGG 136
```

RESULT 28
US-10-540-898-398/c
; Sequence 398, Application US/10540898
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro

```
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: CHIR0056-101 (PP023367.0003)
;  CURRENT APPLICATION NUMBER: US/10/540,898
;  CURRENT FILING DATE:  2005-06-27
;  PRIOR APPLICATION NUMBER: US 10/330,773
;  PRIOR FILING DATE: 2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 398
;   LENGTH: 228835
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-540-898-398

Alignment Scores:
Pred. No.:                295             Length:         228835
Score:                    93.50           Matches:        45
Percent Similarity:       41.5%           Conservative:   43
Best Local Similarity:    21.2%           Mismatches:     85
Query Match:              7.8%            Indels:         39
DB:                       6               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-540-898-398 (1-228835)

Qy          8 AsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspVal 27
              |||          ::::::   |||:::   :::::::::           |||   |||
Db      91221 AATGCCCCAGTTAAAAGTCACGGACTGGCAAATTGGATAAAGAGTCAAGACCCATGGGTG 91162

Qy         28 IlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsn 47
                 :::||||||                    ::::::         :::   ::::::
Db      91161 TGCTGTATTCAGGAG-----------------ACCCATTTCATGTGCAAAGACGCACAT 91120

Qy         48 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 67
                 |||   |||      :::   |||   |||             ||||||:::
Db      91119 AGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAACAAATGGAAAGCAAAAAAAAGCA 91060

Qy         68 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 87
                 |||   :::          :::         |||      ||||||:::   |||:::
Db      91059 GGGGTTACAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGAC 91000

Qy         88 LeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMet 103
                 |||:::   |||   |||:::           ||||||   ::::::
Db      90999 AAAGAAGGCCACCACATAATGGTAAGGGGATCAATTCACCAAGAAGAGCTAACTATCCTA 90940

Qy        104 ThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMet 123
                 :::      :::|||             |||   |||          :::|||
Db      90939 AATATATATGCACCCAATACA---------GGAGCAACCAGATTC------ATAAAGCAA 90895

Qy        124 ValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly------ 141
                 ||||||:::   :::|||         :::|||   |||:::|||
Db      90894 GTCCTTAGAGATCTACAAAGAGACTTAGACTCCCACACAATAATAATGGGAGACTTTAAA 90835

Qy        142 -----------------AspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
                              |||   ::::::|||   :::                  |||
Db      90834 ACCCCACTGTCAGTATTAGACAGATCAATGAGACAGAAGGTTAACAAGGATATCGAGGGC 90775

Qy        156 LeuProAsn--------------AsnIleValAspValTrpGluPheLeuGlyLysPro 170
                 |||   :::          :::::::::|||::::::       |||      |||
Db      90774 TTGAACTCAGCTCTGCACCAAGTAGACCTAATAGACATCTACAGAACTCTC---CACCCC 90718

Qy        171 Lys----------------------------HisCysGlnTyrThrTrpAspThrGln 180
                 |||                         ||||||   |||:::
Db      90717 AAATCAACAGAATATACATTCTTCTCAGCACCACATTGCACTTATTCTAAAATTGACCAC 90658

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
                 :::      :::      :::   :::   ||||||
Db      90657 TTAGTTGGAAGTAAAGCACTCCTCAGCAAATGTAAA 90622

RESULT 29
US-60-762-056-19596
; Sequence 19596, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 19596
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-19596

Alignment Scores:
Pred. No.:                0.257           Length:         2300
Score:                    92.50           Matches:        49
Percent Similarity:       45.4%           Conservative:   40
Best Local Similarity:    25.0%           Mismatches:     71
Query Match:              7.8%            Indels:         36
DB:                       11              Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-60-762-056-19596 (1-2300)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::          :::         :::   :::::::::
Db       1243 TTAAATGTAAATGAACTAAATGCTCCAGTTAAAAAACACAGACTGGCAAATTGGATAAAG 1302

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                   |||   |||      :::||||||                        :::::::|||
Db       1303 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------ACCCATCTC 1344

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                   |||   ::::::          |||   |||       :::   |||      |||
Db       1345 ACATGCAGAGACACACACAGACTCAAAATAAAGGGATGGAGGAAGATATAC---TAAGCA 1401

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                   :::|||   |||      |||:::|||      :::|||   |||
Db       1402 AATGGAAAACAAAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATTGACTTTAAACCA 1461

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
                 ||||||:::   |||:::          :::::::   |||   |||:::
Db       1462 ACAAAGATCAAAAGAGACAAAGAAGGGCATTACATAATGGTAAAGGGATCAATTCAACAA 1521

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                    ||||||   |||:::      :::      :::|||
Db       1522 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACAGGAGCA-------------- 1566

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                 :::      :::|||   ||||||:::   :::|||          :::|||   |||:::
Db       1567 CTCAGATTCATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCTACACAATA 1626

Qy        138 IlePheAlaGlyAspThrAsnLeu-Arg----------------------AspArgGl 149
                 |||      ||||||   |||                          ||||||::
Db       1627 ATAATG---GGAGATTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAA 1683

Qy        149 uValThrArgCysGlyGlyLeu--------------------ProAsnAsnIleValAs 162
                 ::::||||||                           |||   ::::::::||
Db       1684 GTTAACAAGGATATCCAGGACTTGAACTCAGCTCTGCACCCACCCAATGGACCTAATAGA 1743

Qy        162 pValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
                 |::::::      |||      ||||||      :::||||||:::
Db       1744 CATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1786

RESULT 30
US-10-868-184C-10303/c
; Sequence 10303, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
```

```
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10303
;   LENGTH: 6198
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10303

Alignment Scores:
Pred. No.:                1.47          Length:         6198
Score:                    92.00         Matches:        50
Percent Similarity:       46.7%         Conservative:   42
Best Local Similarity:    25.4%         Mismatches:     69
Query Match:              7.7%          Indels:         36
DB:                       6             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10303 (1-6198)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              :::::||||||          |||   |||   :::          :::   :::::::::
Db       4045 GTAAATGGACTA---------AATGCTTCAATTAAAAGACACAGACTGGCAAATTGGATA 3995

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                        |||   |||       :::||||||                    ::::::
Db       3994 AAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCAT 3953

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||     |||   ::::::          |||   |||      :::   |||
Db       3952 CTCATGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCA 3893

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                      ||||||:::   |||   :::   :::   :::          |||
Db       3892 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGCCTCTGATAAAACAGATTTTAAA 3833

Qy         81 SerThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnValSer 96
                 ||||||:::   |||            :::   :::   |||   |||::::::
Db       3832 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGTATCAATTCAA 3773

Qy         97 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                ||||||   ::::::   :::         :::|||          |||
Db       3772 CAAGAAGAATTAACTATCCTAAATACATGTGCACCCAATACA---------GGAGCATCC 3722

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||         :::|||  ||||||   :::|||        :::|||   |||
Db       3721 AGATTC------ATAAAGCAAGTCCTTAGTGACCTACAAAGAGACTTAGACTCCCACACA 3668

Qy        137 ValIlePheAlaGlyAsp------------------------ThrAsnLeuArgAspAr 148
              :::|||      ||||||                      |||:::   ||||||||
Db       3667 ATAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAG 3611

Qy        148 gGluValThrArgCys-GlyGlyLeu--------------ProAsn-AsnIleValAsp 162
```

```
              |:::::||||||       |||:::              |||::: :::::::::|||
Db       3610 AAAGTTAACAAGGATACCCAGGAATTCAACTCAGCTCTGCACCAAGTGGACCTAATAGAC 3551

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::::::      |||    ||||||     :::||||||:::
Db       3550 ATCTACAGAATTCTC---CACCCCAAATCAACAGAATATACATTT 3509

RESULT 31
US-10-868-184D-10303/c
; Sequence 10303, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10303
;   LENGTH: 6198
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10303

Alignment Scores:
Pred. No.:                1.47          Length:         6198
Score:                    92.00         Matches:        50
Percent Similarity:       46.7%         Conservative:   42
Best Local Similarity:    25.4%         Mismatches:     69
Query Match:              7.7%          Indels:         36
DB:                       6             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10303 (1-6198)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              :::::||||||          |||   |||   :::          :::   :::::::::
Db       4045 GTAAATGGACTA---------AATGCTTCAATTAAAAGACACAGACTGGCAAATTGGATA 3995

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                        |||   |||       :::||||||                    ::::::
Db       3994 AAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCAT 3953

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||     |||   ::::::          |||   |||      :::   |||
Db       3952 CTCATGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCA 3893

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                      ||||||:::   |||   :::   :::   :::          |||
Db       3892 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGCCTCTGATAAAACAGATTTTAAA 3833

Qy         81 SerThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnValSer 96
                 ||||||:::   |||            :::   :::   |||   |||::::::
```

```
Db       3832 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGTATCAATTCAA 3773

Qy         97 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
Db       3772 CAAGAAGAATTAACTATCCTAAATACATGTGCACCCAATACA---------GGAGCATCC 3722

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
Db       3721 AGATTC------ATAAAGCAAGTCCTTAGTGACCTACAAAGAGACTTAGACTCCCACACA 3668

Qy        137 ValIlePheAlaGlyAsp----------------------ThrAsnLeuArgAspAr 148
Db       3667 ATAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAG 3611

Qy        148 gGluValThrArgCys-GlyGlyLeu---------------ProAsn-AsnIleValAsp 162
Db       3610 AAAGTTAACAAGGATACCCAGGAATTCAACTCAGCTCTGCACCAAGTGGACCTAATAGAC 3551

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
Db       3550 ATCTACAGAATTCTC---CACCCCAAATCAACAGAATATACATTT 3509

RESULT 32
US-10-868-184C-10659/c
; Sequence 10659, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10659
;   LENGTH: 6075
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10659

Alignment Scores:
Pred. No.:                1.96          Length:          6075
Score:                    91.00         Matches:         51
Percent Similarity:       44.8%         Conservative:    36
Best Local Similarity:    26.3%         Mismatches:      68
Query Match:              7.6%          Indels:          39
DB:                       6             Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10659 (1-6075)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
Db       4039 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 3983

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
Db       3982 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 3941

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db       3940 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 3881

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db       3880 AAGGCAGGGGTTGCAATCCTAGTCTCTGACAAAACAGACTTTAAACCAACAAAGATCAAA 3821

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------------GluLeuCys 101
Db       3820 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 3761

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db       3760 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 3716

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db       3715 AAGCAAGTCCTCAGTCACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 3659

Qy        142 Asp-------------------------ThrAsnLeuArgAspArgGluValThrArgCy 153
Db       3658 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGG-- 3601

Qy        153 sGlyGlyLeuPro--------------------------AsnAsnIleValAspValTr 164
Db       3600 -------ATACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGAACCTAATAGACATCTA 3548

Qy        164 pGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
Db       3547 CAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3511

RESULT 33
US-10-868-184D-10659/c
; Sequence 10659, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10659
;   LENGTH: 6075
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-10-868-184D-10659

Alignment Scores:

Pred. No.:	1.96	Length:	6075
Score:	91.00	Matches:	51
Percent Similarity:	44.8%	Conservative:	36
Best Local Similarity:	26.3%	Mismatches:	68
Query Match:	7.6%	Indels:	39
DB:	6	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10659 (1-6075)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::            |||
Db       4039 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 3983

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
              |||      :::||||||                 :::::::|||      |||
Db       3982 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 3941

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||      :::   |||                  |||
Db       3940 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 3881

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db       3880 AAGGCAGGGGTTGCAATCCTAGTCTCTGACAAAACAGACTTTAAACCAACAAAGATCAAA 3821

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::         ::::::   |||   |||:::             ||||||
Db       3820 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 3761

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||           |||   |||         :::
Db       3760 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 3716

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db       3715 AAGCAAGTCCTCAGTCACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 3659

Qy        142 Asp-------------------------ThrAsnLeuArgAspArgGluValThrArgCy 153
              |||                        |||:::  |||||||||:::   ||||||
Db       3658 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGG-- 3601

Qy        153 sGlyGlyLeuPro---------------------------AsnAsnIleValAspValTr 164
                  :::|||                           |||::::::|||:::::
Db       3600 -------ATACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGAACCTAATAGACATCTA 3548

Qy        164 pGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :      |||      ||||||      :::|||||||:::
Db       3547 CAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3511
```

RESULT 34
US-10-868-184C-10289/c
; Sequence 10289, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10289
; LENGTH: 5310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-868-184C-10289

Alignment Scores:

Pred. No.:	1.85	Length:	5310
Score:	90.50	Matches:	46
Percent Similarity:	43.8%	Conservative:	39
Best Local Similarity:	23.7%	Mismatches:	76
Query Match:	7.6%	Indels:	33
DB:	6	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10289 (1-5310)

```
Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db       4275 TTAAATGTAAATGAACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 4216

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                    |||   |||       :::||||||                    :::::::|||
Db       4215 ACTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------ACCCATCTG 4174

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                  |||   ::::::         |||   |||   ::::::   |||
Db       4173 ACGTGCAGAGACACACACAGGCTCAAAATAAAGGGATGGCAGAAGATCTACCAAGCAAAT 4114

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                    ||||||:::   |||   :::         :::          |||
Db       4113 GGAAAACAAAAAAAAGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 4054

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
                ||:::   |||:::      |||   ::::::   |||   |||:::
Db       4053 AGAAAGATCAAAAGAGACAAAGAATGTCGTTACATAATGGTAAAGGGATCAATTCAACAA 3994

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                  ||||||   ::::::      :::         :::|||              |||   |||
Db       3993 GAAGAACTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3943

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                    :::|||   ||||||      :::|||          :::|||   |||:::
Db       3942 TTC------ATAAAGCAAGTCCTTAGTGACATACAAAGAGACTTAGACTCCCATACAATA 3889

Qy        138 IlePheAlaGlyAspThrAsn------------LeuArgAspArgGluValThrArg--- 152
              |||      ||||||   |||                  ||||||      :::
Db       3888 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACAAGACAGAAA 3832

Qy        153 ------------------------CysGlyGlyLeuProAsn---AsnIleValAspVal 163
                                     ||||||      :::::::::|||:::
Db       3831 GTTAATAAGGATAACCAGGAATTTAACTCAGCTCTGCCCCAAGCAGACCTAATAGACATC 3772

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::|||||:::
Db       3771 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 3733
```

RESULT 35
US-10-868-184D-10289/c
; Sequence 10289, Application US/10868184D
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins

```
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10289
;    LENGTH: 5310
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-868-184D-10289

Alignment Scores:
Pred. No.:              1.85          Length:         5310
Score:                  90.50         Matches:        46
Percent Similarity:     43.8%         Conservative:   39
Best Local Similarity:  23.7%         Mismatches:     76
Query Match:            7.6%          Indels:         33
DB:                     6             Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10289 (1-5310)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::        :::      :::   :::::::::
Db       4275 TTAAATGTAAATGAACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 4216

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                   |||   |||      :::||||||                   ::::::|||
Db       4215 ACTCAAGACCCATCAGTGTGCTGTATTCAGGAG----------------ACCCATCTG 4174

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::         |||   |||   ::::::   |||
Db       4173 ACGTGCAGAGACACACACAGGCTCAAAATAAAGGGATGGCAGAAGATCTACCAAGCAAAT 4114

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                   ||||||:::  |||  :::      :::          |||
Db       4113 GGAAAACAAAAAAAAGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 4054

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
               |||:::   |||:::      |||   ::::::   |||   |||:::
Db       4053 AGAAAGATCAAAAGAGACAAAGAATGTCGTTACATAATGGTAAAGGGATCAATTCAACAA 3994

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                  ||||||   ::::::      :::      :::|||              |||   |||
Db       3993 GAAGAACTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3943

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   ||||||      :::|||      :::|||   |||:::
Db       3942 TTC------ATAAAGCAAGTCCTTAGTGACATACAAAGAGACTTAGACTCCCATACAATA 3889

Qy        138 IlePheAlaGlyAspThrAsn-----------LeuArgAspArgGluValThrArg--- 152
               |||      ||||||   |||                  ||||||         :::
Db       3888 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACAAGACAGAAA 3832

Qy        153 ----------------------CysGlyGlyLeuProAsn---AsnIleValAspVal 163
                                       ||||||      :::::::::|||:::
Db       3831 GTTAATAAGGATAACCAGGAATTTAACTCAGCTCTGCCCCAAGCAGACCTAATAGACATC 3772

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
               :::      |||      ||||||      :::||||||:::
Db       3771 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 3733

RESULT 36
US-10-953-349-37498/c
; Sequence 37498, Application US/10953349
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCONDED THERBY
;  FILE REFERENCE: 2750-1579PUS2
;  CURRENT APPLICATION NUMBER: US/10/953,349
;  CURRENT FILING DATE:  2004-09-30
;  NUMBER OF SEQ ID NOS: 40252
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 37498
;    LENGTH: 1414
;    TYPE: DNA
;    ORGANISM: Zea mays subsp. mays
US-10-953-349-37498

Alignment Scores:
Pred. No.:              0.305         Length:         1414
Score:                  89.50         Matches:        47
Percent Similarity:     33.6%         Conservative:   27
Best Local Similarity:  21.4%         Mismatches:     51
Query Match:            7.5%          Indels:         95
DB:                     6             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-953-349-37498 (1-1414)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVal--------------- 16
              |||   ||||||   :::   |||:::||||||      ||||||
Db        600 GATTAGCTGGATCAGGATAAGCTTGCTGAGAGGCCCTAGGGCGTTGTTCTGTAGGTAATG 541

Qy         17 --------------------------------CysSerTyrLeuAlaLeuTyr---Ser 24
                                                ||||:::   ||||||::::::
Db        540 ATGGATTTGTGAAATCTCTAGATCGGTCGTTGTTGTAATGGACTTGCAGTTCACTGAGAT 481

Qy         25 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 44
                 ::::::   :::|||
Db        480 TTGAATATTGAATATCTA-----------------------------------GCT 460

Qy         45 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 64
                 ||||||:::   |||   :::   |||   :::                  :::|||
Db        459 GGCAGCAACCATCATATATGTTCCTTCCATGCAAAGAACCAAATGCTTCCACAGCTTCTT 400

Qy         65 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 84
                 |||      :::   :::   :::   :::|||      ||||||
Db        399 GGCGCGAGTGAAACTGAATAAGAGCTTGAAAACCTGATCAGAATTCCTTCC--------- 349

Qy         85 MetArgAsnLeuLeuCys----------------------ValHisValAsnValSer 96
              ||||||   |||:::|||                     :::||||||||:::
Db        348 ATGCGAGCTCTGGTCTGTGGTAAGCTCTTGGTGAGATGAGAACTGCATGTAAACATTCCT 289

Qy         97 Gly------------------------------------------------------- 97

Db        288 TCCCCTTATGCTAGGTTGAACGGAAGTGTAATATTGCAGCGCACTCACGGAAGCATGTAT 229

Qy         98 --------------AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
                             |||   |||   |||   |||                  |||
Db        228 GTCTTCCATCTGCAGAAGAGTTTGGTTCTTGGCACGCAACAT------GACGATCTTGGA 175

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
                 |||||||||:::   |||||||||   :::
Db        174 GACCACGCCGAACGGCTGCAGCAGCTGAAGGAGATC----------------------- 139
```

```
Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
                                                                        |||
Db        138 ------------------------------------------------------------AGA 136

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              |||   |||      ::::::  ||||||  ::::::||||||:::  :::|||
Db        135 CTCGGCAATCTCGTGCCCGACGTTGCGGATGTG---GATTACCTTGGAAGGCTCCGACAT 79

RESULT 37
US-10-868-184C-10371/c
; Sequence 10371, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10371
;   LENGTH: 6063
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10371

Alignment Scores:
Pred. No.:              3.14          Length:          6063
Score:                  89.50         Matches:         45
Percent Similarity:     42.3%         Conservative:    37
Best Local Similarity:  23.2%         Mismatches:      79
Query Match:            7.5%          Indels:          33
DB:                     6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10371 (1-6063)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::      :::   :::   |||   :::
Db       4038 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAGTTCGATAAAG 3979

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||   ::::::||||||                  ::::::|||
Db       3978 AGTCAAGACCCATCAGTGTGCTATATTCAGGAA----------------ACCCATCTT 3937

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::          |||   |||   :::   |||
Db       3936 ACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 3877

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||   :::         :::         |||
Db       3876 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAGCAGACTTTAAACCA 3817

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
```

```
                 ||||||:::   |||:::             ::::::   |||   |||:::
Db       3816 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 3757

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                  ||||||   ::::::       |||       :::|||              |||   |||
Db       3756 GAGGAGCTAACTATCCTAAATATACATGCACCCAATACA---------GGAGCACCCAGA 3706

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   ||||||      :::|||          :::|||   |||:::
Db       3705 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 3652

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
               |||       ||||||  |||
Db       3651 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAA 3595

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||  :::                                 :::::::::|||:::
Db       3594 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATC 3535

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
               :::      |||      ||||||      :::|||||:::
Db       3534 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3496

RESULT 38
US-10-868-184D-10371/c
; Sequence 10371, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10371
;   LENGTH: 6063
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10371

Alignment Scores:
Pred. No.:              3.14          Length:          6063
Score:                  89.50         Matches:         45
Percent Similarity:     42.3%         Conservative:    37
Best Local Similarity:  23.2%         Mismatches:      79
Query Match:            7.5%          Indels:          33
DB:                     6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10371 (1-6063)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::      :::   :::   |||   :::
```

```
Db       4038 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAGTTCGATAAAG 3979

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41

Db       3978 AGTCAAGACCCATCAGTGTGCTATATTCAGGAA----------------ACCCATCTT 3937

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61

Db       3936 ACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 3877

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81

Db       3876 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAGCAGACTTTAAACCA 3817

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98

Db       3816 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 3757

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117

Db       3756 GAGGAGCTAACTATCCTAAATATACATGCACCCAATACA---------GGAGCACCCAGA 3706

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137

Db       3705 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 3652

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145

Db       3651 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAA 3595

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163

Db       3594 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATC 3535

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177

Db       3534 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3496
```

```
RESULT 39
US-10-868-184C-10470/c
; Sequence 10470, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10470
;   LENGTH: 5381
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10470

Alignment Scores:
Pred. No.:              3.56          Length:          5381
Score:                  88.50         Matches:         45
Percent Similarity:     42.8%         Conservative:    38
Best Local Similarity:  23.2%         Mismatches:      78
Query Match:            7.4%          Indels:          33
DB:                     6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10470 (1-5381)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21

Db       4030 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG 3971

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41

Db       3970 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 3929

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61

Db       3928 ACGTGCAGAGATACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAACAAAT 3869

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81

Db       3868 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3809

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98

Db       3808 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 3749

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117

Db       3748 GAGGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3698

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137

Db       3697 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 3644

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145

Db       3643 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 3587

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163

Db       3586 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGTACCAAGCAGACCTAATAGACATC 3527

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177

Db       3526 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3488
```

```
RESULT 40
US-10-868-184C-10513/c
; Sequence 10513, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
```

```
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10513
;   LENGTH: 5381
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10513

Alignment Scores:
Pred. No.:              3.56          Length:        5381
Score:                  88.50         Matches:       45
Percent Similarity:     42.8%         Conservative:  38
Best Local Similarity:  23.2%         Mismatches:    78
Query Match:            7.4%          Indels:        33
DB:                     6             Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10513 (1-5381)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::     :::         :::   |||::::::
Db       4030 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG 3971

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                    |||   |||       :::||||||                   ::::::|||
Db       3970 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 3929

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||      :::   |||   |||
Db       3928 ACGTGCAGAGATACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAACAAAT 3869

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||  :::        :::         |||
Db       3868 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3809

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::           ::::::   |||   |||:::
Db       3808 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 3749

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                |||||   ::::::      :::      :::|||          |||   |||
Db       3748 GAGGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3698

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   ||||||      :::|||          :::|||   |||:::
Db       3697 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 3644

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db       3643 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 3587

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                  :::|||  :::                    :::::::::|||:::
Db       3586 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGTACCAAGCAGACCTAATAGACATC 3527

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::||||||:::
Db       3526 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3488

RESULT 41
US-10-868-184D-10470/c
; Sequence 10470, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10470
;   LENGTH: 5381
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10470

Alignment Scores:
Pred. No.:              3.56          Length:        5381
Score:                  88.50         Matches:       45
Percent Similarity:     42.8%         Conservative:  38
Best Local Similarity:  23.2%         Mismatches:    78
Query Match:            7.4%          Indels:        33
DB:                     6             Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10470 (1-5381)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   |||::::::
Db       4030 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG 3971

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                    |||   |||       :::||||||                   ::::::|||
Db       3970 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 3929

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||      :::   |||   |||
Db       3928 ACGTGCAGAGATACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAACAAAT 3869

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||  :::        :::         |||
Db       3868 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3809

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::           ::::::   |||   |||:::
Db       3808 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 3749

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                |||||   ::::::      :::      :::|||          |||   |||
Db       3748 GAGGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3698

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   ||||||      :::|||          :::|||   |||:::
Db       3697 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 3644

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db       3643 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 3587

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                  :::|||  :::                    :::::::::|||:::
Db       3586 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGTACCAAGCAGACCTAATAGACATC 3527

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::||||||:::
Db       3526 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3488

RESULT 41
US-10-868-184D-10470/c
; Sequence 10470, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
```

```
Db       3643 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 3587

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                  :::|||   :::                                 :::::::::|||:::
Db       3586 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGTACCAAGCAGACCTAATAGACATC 3527

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::||||||:::
Db       3526 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3488

RESULT 42
US-10-868-184D-10513/c
; Sequence 10513, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10513
;   LENGTH: 5381
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10513

Alignment Scores:
Pred. No.:              3.56          Length:         5381
Score:                  88.50         Matches:        45
Percent Similarity:     42.8%         Conservative:   38
Best Local Similarity:  23.2%         Mismatches:     78
Query Match:            7.4%          Indels:         33
DB:                     6             Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10513 (1-5381)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::         :::         :::   |||::::::
Db       4030 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG 3971

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                   |||   |||      :::||||||                    :::::::|||
Db       3970 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 3929

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::         |||   |||       :::   |||   |||
Db       3928 ACGTGCAGAGATACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAACAAAT 3869

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                   ||||||:::   |||   :::         :::         |||
Db       3868 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3809

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::             ::::::   |||   |||:::
Db       3808 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 3749

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||          |||   |||
Db       3748 GAGGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3698

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   ||||||      :::|||          :::|||   |||:::
Db       3697 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 3644

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db       3643 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 3587

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                  :::|||   :::                                 :::::::::|||:::
Db       3586 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGTACCAAGCAGACCTAATAGACATC 3527

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::||||||:::
Db       3526 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3488

RESULT 43
US-10-868-184C-10537/c
; Sequence 10537, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10537
;   LENGTH: 5671
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10537

Alignment Scores:
Pred. No.:              3.88          Length:         5671
Score:                  88.50         Matches:        42
Percent Similarity:     44.8%         Conservative:   45
Best Local Similarity:  21.6%         Mismatches:     74
Query Match:            7.4%          Indels:         33
DB:                     6             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10537 (1-5671)
```

```
Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::         :::          :::   :::::::::
Db       4017 TTAAATGTAAATGGGCTGAATGCTCCAATTAAAAGACACAGACTGGCAAACTGGATGAAG 3958

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||  ::::::||||||                        ::::::|||
Db       3957 AGTCAAGACCCGTCAGTGTGCTATATTCAGGAA-----------------ACCCATCTC 3916

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||       :::   |||
Db       3915 ATGTGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 3856

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                |||::::::   |||   :::         :::         |||
Db       3855 GGAAAACAAAAAGAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3796

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::             ::::::   |||   |||:::
Db       3795 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 3736

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                ||||||   ::::::      |||       :::                |||   |||
Db       3735 GAAGAGCTAACTATCTTAAATATACATGCACCCAATATA---------GGAGCACCCAGA 3685

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                :::|||   ||||||:::   :::|||         :::|||   |||:::
Db       3684 TTC------ATAAAGCAGGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACAATA 3631

Qy        138 IlePheAlaGlyAspThrAsnLeuArg------------------------------- 146
              |||       ||||||   |||   |||
Db       3630 ATAATG---GGAGACTTTAACACCCGACTGTCAACATTAGACAGATCAATGAGACAGAAG 3574

Qy        147 ---AspArgGluValThrArgCysGlyGlyLeuProAsn------AsnIleValAspVal 163
                :::::::::::   :::             :::      :::::::::|||:::
Db       3573 GTTAACAAGGATATCCAGAAATTGAACTCAGCTCTGCACCAAGTGGACCTAATAGACATC 3514

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::   |||       ||||||      :::||||||:::
Db       3513 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 3475

RESULT 44
US-10-868-184D-10537/c
; Sequence 10537, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10537
;   LENGTH: 5671
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10537

Alignment Scores:
Pred. No.:                3.88          Length:          5671
Score:                    88.50         Matches:         42
Percent Similarity:       44.8%         Conservative:    45
Best Local Similarity:    21.6%         Mismatches:      74
Query Match:              7.4%          Indels:          33
DB:                       6             Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10537 (1-5671)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::         :::          :::   :::::::::
Db       4017 TTAAATGTAAATGGGCTGAATGCTCCAATTAAAAGACACAGACTGGCAAACTGGATGAAG 3958

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||  ::::::||||||                        ::::::|||
Db       3957 AGTCAAGACCCGTCAGTGTGCTATATTCAGGAA-----------------ACCCATCTC 3916

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||       :::   |||
Db       3915 ATGTGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 3856

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                |||::::::   |||   :::         :::         |||
Db       3855 GGAAAACAAAAAGAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3796

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::             ::::::   |||   |||:::
Db       3795 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 3736

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                ||||||   ::::::      |||       :::                |||   |||
Db       3735 GAAGAGCTAACTATCTTAAATATACATGCACCCAATATA---------GGAGCACCCAGA 3685

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                :::|||   ||||||:::   :::|||         :::|||   |||:::
Db       3684 TTC------ATAAAGCAGGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACAATA 3631

Qy        138 IlePheAlaGlyAspThrAsnLeuArg------------------------------- 146
              |||       ||||||   |||   |||
Db       3630 ATAATG---GGAGACTTTAACACCCGACTGTCAACATTAGACAGATCAATGAGACAGAAG 3574

Qy        147 ---AspArgGluValThrArgCysGlyGlyLeuProAsn------AsnIleValAspVal 163
                :::::::::::   :::             :::      :::::::::|||:::
Db       3573 GTTAACAAGGATATCCAGAAATTGAACTCAGCTCTGCACCAAGTGGACCTAATAGACATC 3514

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::   |||       ||||||      :::||||||:::
Db       3513 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 3475

RESULT 45
US-10-868-184C-10398/c
; Sequence 10398, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
```

```
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10398
;   LENGTH: 5870
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10398

Alignment Scores:
Pred. No.:                4.1          Length:          5870
Score:                    88.50        Matches:         44
Percent Similarity:       42.3%        Conservative:    38
Best Local Similarity:    22.7%        Mismatches:      79
Query Match:              7.4%         Indels:          33
DB:                       6            Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10398 (1-5870)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::         :::        :::   ::::::|||
Db       4064 TTAAATATAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGTTAAAG 4005

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::||||||                   ::::::|||
Db       4004 AGTCAACATCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 3963

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||      :::   |||
Db       3962 CCATGCAGAGAAACACATAGGCTCCAAATAAAAGGATGGAGGAAGATCTACCAAGCAGAT 3903

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                ||||||:::   |||   :::        :::        |||
Db       3902 TGAAAACAAAAAAAGGCAGAGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3843

Qy         82 ThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnValSerGly 97
              ||||||:::   |||           :::         |||   |||::::::
Db       3842 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACAGAATGGTAAAGGTATCAATTCAACAA 3783

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                ||||||   ::::::      :::      :::|||              |||   |||
Db       3782 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3732

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                :::|||   ||||||      :::|||         :::|||   |||:::
Db       3731 TTT------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 3678

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db       3677 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAATCAGAAAGAAA 3621

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::   :::                    :::::::::||:::
Db       3620 GTTAACAAGGATGCCCAGGAATCGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 3561

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::|||||:::
Db       3560 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3522

RESULT 46
US-10-868-184D-10398/c
; Sequence 10398, Application US/10868184D
```

```
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10398
;   LENGTH: 5870
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10398

Alignment Scores:
Pred. No.:                4.1          Length:          5870
Score:                    88.50        Matches:         44
Percent Similarity:       42.3%        Conservative:    38
Best Local Similarity:    22.7%        Mismatches:      79
Query Match:              7.4%         Indels:          33
DB:                       6            Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10398 (1-5870)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::         :::        :::   ::::::|||
Db       4064 TTAAATATAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGTTAAAG 4005

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::||||||                   ::::::|||
Db       4004 AGTCAACATCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 3963

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||      :::   |||
Db       3962 CCATGCAGAGAAACACATAGGCTCCAAATAAAAGGATGGAGGAAGATCTACCAAGCAGAT 3903

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                ||||||:::   |||   :::        :::        |||
Db       3902 TGAAAACAAAAAAAGGCAGAGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3843

Qy         82 ThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnValSerGly 97
              ||||||:::   |||           :::         |||   |||::::::
Db       3842 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACAGAATGGTAAAGGTATCAATTCAACAA 3783

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                ||||||   ::::::      :::      :::|||              |||   |||
Db       3782 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3732

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                :::|||   ||||||      :::|||         :::|||   |||:::
Db       3731 TTT------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 3678
```

```
Qy        138 IlePheAlaGlyAspThrAsnLeu----------------------------------- 145
              |||      ||||||   |||
Db       3677 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAATCAGAAAGAAA 3621

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::   :::                 :::::::::|||::: 
Db       3620 GTTAACAAGGATGCCCAGGAATCGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 3561

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||     :::||||||::: 
Db       3560 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3522

RESULT 47
US-10-540-898-582
; Sequence 582, Application US/10540898
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: CHIR0056-101 (PP023367.0003)
;  CURRENT APPLICATION NUMBER: US/10/540,898
;  CURRENT FILING DATE:  2005-06-27
;  PRIOR APPLICATION NUMBER: US 10/330,773
;  PRIOR FILING DATE: 2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 582
;   LENGTH: 70708
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-540-898-582

Alignment Scores:
Pred. No.:              220          Length:        70708
Score:                  88.50        Matches:       44
Percent Similarity:     43.3%        Conservative:  40
Best Local Similarity:  22.7%        Mismatches:    77
Query Match:            7.4%         Indels:        33
DB:                     6            Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-540-898-582 (1-70708)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::         :::       :::   :::::::::
Db      17945 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 18004

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::||||||                 ::::::|||
Db      18005 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACACATCTC 18046

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||:::::::::          |||   |||       :::   |||
Db      18047 ACGTGCAGAAACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 18106

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::||||||   :::       :::       |||
Db      18107 GGAAAACAAAAAAAGGCAAGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 18166

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::          :::::   |||   |||:::
Db      18167 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 18226

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||           |||   |||
Db      18227 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 18277

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   ||||||       :::|||       :::|||   |||:::
Db      18278 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 18331

Qy        138 IlePheAlaGlyAspThrAsnLeu----------------------------------- 145
              |||      ||||||   |||
Db      18332 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 18388

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                 :::::::::|||::: 
Db      18389 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGATCTAATAGACATC 18448

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||     :::||||||::: 
Db      18449 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 18487

RESULT 48
US-60-742-219-1685
; Sequence 1685, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1685
;   LENGTH: 3958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-1685

Alignment Scores:
Pred. No.:              2.55         Length:        3958
Score:                  88.00        Matches:       63
Percent Similarity:     36.5%        Conservative:  32
Best Local Similarity:  24.2%        Mismatches:    87
Query Match:            7.4%         Indels:        79
DB:                     11           Gaps:          15

US-10-757-745-2_COPY_121_347 (1-227) x US-60-742-219-1685 (1-3958)

Qy          9 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 28
              |||   ||||||          |||       ::      |||   |||   :::
Db        868 CTTGAGGAAAGGGAAAACTCAGCATGCTTGTTAGCATCTTCAGAGAGTGATGATTCCCTG 927

Qy         29 Phe-LeuGlnGluValIleProProTyrTyrSerTyrLeu---------LysLysArgSe 45
              ::: :::                |||   :::|||   |||         |||   :::
Db        928 TATGATGATAAATCACAAGCACCCACAATGGTCGATACTATTCCAAGTGAAGTGGAAACA 987

Qy         45 rSerAsnTyrGluIleIleThrGlyHisGluGluGly---TyrPheThrAlaIleMetLe 64
              |||:::|||:::   :::   |||                ::::|||      |||||::
Db        988 AAGTCATTACAAAACTCTAGCTGGTGCGAGACTTTCCCTTCACTTTCGGAAGATAATGAT 1047

Qy         64 uLysLys---SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLy 83
              :||||||   :::|||::::::|||   ||||||:::
Db        1048 GAAAAAGAGAACAAGAATAAGATTGATATCTCAGCAG---------------------- 1084

Qy         83 sMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMe 103
                                |||||                        |||||:::::
Db        1085 ---------------TTATGT--------------------------CTGTGCATTCT 1101

Qy        103 tThrSer------------HisLeuGluSerThrArgGly------HisAlaAlaGluAr 117
              :   |||            |||:::       :::         |||
Db        1102 GCAGTCAAAGAGGAAAGTGCATGTGTTAAAGATGAAAAGTTTGCTCCACCTTTGTCACCT 1161

Qy        117 gMetAsnGlnLeuLysMetValLeu---LysLysMetGlnGluAlaProGluSerAlaTh 136
                 :::|||:::   :::::::::   |||   ::::::|||       |||
```

```
Db       1162 TTGTCATCAGATATG-CTCATTATGCCCAAAGCTTTGAAAGAAGAGAGCGAG-------- 1212

Qy        136 rValIlePheAlaGlyAspThrAsnLeuArgAspArgGluVal---ThrArgCysGlyGl 155

Db       1213 ----------------GATACCTGCCTAGAGACCCTGGCTGTTCCCTCTGAATGCTCTGC 1256

Qy        155 yLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTy 175

Db       1257 TTTCTCAGAGAATATTGAGGATCCTGGAGAG------GGTCCCTCAAATCCATGCTTAGA 1310

Qy        175 rThrTrpAspThrGln------MetAsnSerAsnLeuGlyIleThrAlaAlaCys----- 191

Db       1311 TACCAGCCAGAATCAACCTTCCATGGAATCAGAGATGGGG---GCTGCAGCATGCCCTGG 1367

Qy        191 ------------------------------------------------------------ 191

Db       1368 GAGTTGTTCAAGGGAATGCGAGGTTTCATTTAGTGCCTCTAACCCTGTCTGGGATTATTC 1427

Qy        192 ----------------LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGl 206

Db       1428 ACATCTTATGAGTAGTGAAAGAAATTTTCAGAGACTGGATTTTGAAGAACTTGAGGAAGA 1487

Qy        206 uGlyHisIleIleProArgSerLeu------AspLeuLeuGlyLeuGluLysLeuAsp 223

Db       1488 AGGTCAAGCCTCTGACAAGTCATTGCTTCCAAGTCGCATTAATCTTTCTCTATTAGAT 1545
```

RESULT 49
US-10-868-184C-10476/c
; Sequence 10476, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10476
; LENGTH: 6093
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-868-184C-10476

Alignment Scores:

Pred. No.:	5.1	Length:	6093
Score:	88.00	Matches:	45
Percent Similarity:	43.2%	Conservative:	37
Best Local Similarity:	23.7%	Mismatches:	76
Query Match:	7.4%	Indels:	32
DB:	6	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10476 (1-6093)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25

Db       4031 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 3975

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45

Db       3974 TCAGTGTGCTGTATTCAGGAA----------------ACCCATCTCACGTGCAGAGAC 3933

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65

Db       3932 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 3873

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85

Db       3872 AAGGCAAGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 3813

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101

Db       3812 AAAGACAAAGAAGGCCATTACATAATGGTAAAAGGATCAATTCAACAAGAGGAGCTAACT 3753

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121

Db       3752 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 3708

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141

Db       3707 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 3651

Qy        142 AspThrAsnLeu----------------------------------------ArgAsp 147

Db       3650 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 3591

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167

Db       3590 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 3531

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177

Db       3530 ---CACCCCAAATCAACAGAATATACATTT 3504
```

RESULT 50
US-10-868-184D-10476/c
; Sequence 10476, Application US/10868184D
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184D
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 10476
;   LENGTH: 6093
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10476

Alignment Scores:
Pred. No.:                5.1          Length:        6093
Score:                    88.00        Matches:       45
Percent Similarity:       43.2%        Conservative:  37
Best Local Similarity:    23.7%        Mismatches:    76
Query Match:              7.4%         Indels:        32
DB:                       6            Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10476 (1-6093)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::     :::||| |||   :::   |||:::::              |||
Db       4031 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 3975

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||       :::||||||                 ::::::|||      |||
Db       3974 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 3933

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||      :::   |||                 |||
Db       3932 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 3873

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::||||||   :::      :::          |||        ||||||:::
Db       3872 AAGGCAAGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 3813

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              ::::::            ::::::   |||   |||:::            ||||||
Db       3812 AAAGACAAAGAAGGCCATTACATAATGGTAAAAGGATCAATTCAACAAGAGGAGCTAACT 3753

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||   |||          :::
Db       3752 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 3708

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||     :::|||        :::|||   |||:::|||          |||
Db       3707 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 3651

Qy        142 AspThrAsnLeu-------------------------------------------ArgAsp 147
              |||   |||                                            :::|||
Db       3650 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 3591

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::                        :::::::::|||::::::      |||
Db       3590 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 3531

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                |||||      :::||||||:::
Db       3530 ---CACCCCAAATCAACAGAATATACATTT 3504

RESULT 51
US-10-868-184C-10429/c
; Sequence 10429, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10429
;   LENGTH: 5691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10429

Alignment Scores:
Pred. No.:                6.28         Length:        5691
Score:                    87.00        Matches:       43
Percent Similarity:       43.1%        Conservative:  41
Best Local Similarity:    22.1%        Mismatches:    77
Query Match:              7.3%         Indels:        34
DB:                       6            Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10429 (1-5691)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ::::::||||||:::      |||         |||         :::   :::::::::
Db       4024 GTAAATGGGCTAAATGCTCCAAATTAAAGACACAGA---------CTGGCAAATTGGATA 3974

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                        |||  |||  ::::::||||||                      ::::::
Db       3973 AGGAGTCAAGACCCATCAGTGTGCTATATTCAGGAA-----------------ACCCAT 3932

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||         |||   ::::::          |||   |||       :::   |||
Db       3931 CTCATGTGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCA 3872

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                        ||||||:::   |||   :::         :::         |||
Db       3871 AATGGAAAACAAAAAAAGGCAGAGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAG 3812

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------ 98
                  ||||||:::   |||:::          ::::::   |||   |||:::
Db       3811 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAA 3752

Qy         99 ------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                    |||:::   ::::::      :::      :::|||           |||
Db       3751 CAAGAAGAGATAACTATCCTAAATATATATGCACTCAATACA---------GGAGCACCC 3701

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||          :::|||   ||||||:::   :::|||          :::|||   |||
Db       3700 AGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACA 3647

Qy        137 ValIlePheAlaGlyAspThrAsnLeu-------------------------------- 145
              :::|||      |||||   |||
Db       3646 ATAATAATG---GGAGACTTTAACACCCTACTGTCAACATTAGACAGATCAACAAGACAG 3590

Qy        146 ---------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
                      :::|||   :::                        ::::::::::|||
Db       3589 AAAGTTAACAAGGATATCCAGGAATTGAACTCGGCTCTGCACCAAGTGGACCTAATAGAC 3530

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::::::      |||      ||||||      :::||||||:::
Db       3529 ATCTACAGAATTCTC---CACCCCAAATCAACAGAATATACATTC 3488

RESULT 52
US-10-868-184D-10429/c
```

```
; Sequence 10429, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10429
;   LENGTH: 5691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10429

Alignment Scores:
Pred. No.:                6.28          Length:          5691
Score:                    87.00         Matches:         43
Percent Similarity:       43.1%         Conservative:    41
Best Local Similarity:    22.1%         Mismatches:      77
Query Match:              7.3%          Indels:          34
DB:                       6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10429 (1-5691)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ::::::||||||:::      |||         |||         :::   :::::::::
Db       4024 GTAAATGGGCTAAATGCTCCAAATTAAAGACACAGA---------CTGGCAAATTGGATA 3974

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                     |||   |||   ::::::||||||                    ::::::
Db       3973 AGGAGTCAAGACCCATCAGTGTGCTATATTCAGGAA-----------------ACCCAT 3932

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||     |||   ::::::         |||   |||        :::   |||
Db       3931 CTCATGTGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCA 3872

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                    ||||||:::   |||   :::         :::         |||
Db       3871 AATGGAAAACAAAAAAAGGCAGAGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAG 3812

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------ 98
                ||||||:::   |||:::         ::::::   |||   |||:::
Db       3811 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAA 3752

Qy         99 ------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                 |||:::   ::::::      :::      :::|||         |||
Db       3751 CAAGAAGAGATAACTATCCTAAATATATATGCACTCAATACA---------GGAGCACCC 3701

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||         :::|||   ||||||:::   :::|||         :::|||   |||
Db       3700 AGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACA 3647

Qy        137 ValIlePheAlaGlyAspThrAsnLeu------------------------------ 145
              :::|||      ||||||   |||
Db       3646 ATAATAATG---GGAGACTTTAACACCCTACTGTCAACATTAGACAGATCAACAAGACAG 3590

Qy        146 ---------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
                      :::|||   :::                         :::::::::|||
Db       3589 AAAGTTAACAAGGATATCCAGGAATTGAACTCGGCTCTGCACCAAGTGGACCTAATAGAC 3530

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::::::      |||     ||||||     :::||||||:::
Db       3529 ATCTACAGAATTCTC---CACCCCAAATCAACAGAATATACATTC 3488

RESULT 53
US-10-868-184C-10437/c
; Sequence 10437, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10437
;   LENGTH: 10367
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10437

Alignment Scores:
Pred. No.:                16.4          Length:          10367
Score:                    87.00         Matches:         49
Percent Similarity:       45.6%         Conservative:    40
Best Local Similarity:    25.1%         Mismatches:      72
Query Match:              7.3%          Indels:          34
DB:                       6             Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10437 (1-10367)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   |||   :::
Db       8222 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAGTTTGATAAAG 8163

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                    |||   |||      :::||||||                    :::::|||
Db       8162 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 8121

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                  |||   :::             |||   |||       :::   |||   |||
Db       8120 ATGTGCAGAGACACATATAGGGTCAAAATAAAAGGATGGAGAAAGATCTACCAAACAAAT 8061
```

```
Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
Db       8060 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 8001

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
Db       8000 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 7941

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
Db       7940 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 7890

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
Db       7889 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCAAGCAATA 7836

Qy        138 IlePheAlaGlyAsp----------------------ThrAsnLeuArgAspArgGlu 149
Db       7835 ATAATG---GGAGACTTTAACACCCACTGTCAACATTAGACAGATCAACAAGACAGAAAG 7779

Qy        150 ValThrArgCys-GlyGlyLeu---------------ProAsn-AsnIleValAspValT 164
Db       7778 TTAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATCT 7719

Qy        164 rpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
Db       7718 ACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 7681
```

RESULT 54
US-10-868-184D-10437/c
; Sequence 10437, Application US/10868184D
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184D
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10437
; LENGTH: 10367
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-868-184D-10437

Alignment Scores:

Pred. No.:	16.4	Length:	10367
Score:	87.00	Matches:	49
Percent Similarity:	45.6%	Conservative:	40
Best Local Similarity:	25.1%	Mismatches:	72
Query Match:	7.3%	Indels:	34
DB:	6	Gaps:	9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10437 (1-10367)

```
Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
Db       8222 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAGTTTGATAAAG 8163

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
Db       8162 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 8121

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
Db       8120 ATGTGCAGAGACACATATAGGGTCAAAATAAAAGGATGGAGAAAGATCTACCAAACAAAT 8061

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
Db       8060 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 8001

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
Db       8000 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 7941

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
Db       7940 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 7890

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
Db       7889 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCAAGCAATA 7836

Qy        138 IlePheAlaGlyAsp----------------------ThrAsnLeuArgAspArgGlu 149
Db       7835 ATAATG---GGAGACTTTAACACCCACTGTCAACATTAGACAGATCAACAAGACAGAAAG 7779

Qy        150 ValThrArgCys-GlyGlyLeu---------------ProAsn-AsnIleValAspValT 164
Db       7778 TTAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATCT 7719

Qy        164 rpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
Db       7718 ACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 7681
```

RESULT 55
US-10-868-184C-10373/c
; Sequence 10373, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.

```
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10373
;   LENGTH: 5647
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10373

Alignment Scores:
Pred. No.:                7.27          Length:          5647
Score:                    86.50         Matches:         51
Percent Similarity:       47.7%         Conservative:    43
Best Local Similarity:    25.9%         Mismatches:      67
Query Match:              7.3%          Indels:          36
DB:                       6             Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10373 (1-5647)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ::::::||||||         |||         :::         :::   :::::::::
Db       4016 GTAAATGGGCTA---------AATGCTCCAGTTAAAAGACACAGACTGGGAAATTGGATA 3966

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                       |||   |||      :::||||||                    ::::::
Db       3965 AACAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCAT 3924

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||  |||  ::::::         |||  |||      :::   |||
Db       3923 CTCAAATGCAGAGACACACACAGGCTCACAATAAAGGGATGGAGGAAGATCTACCAAGCA 3864

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                     ||||||:::  |||  :::         :::          |||
Db       3863 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAA 3804

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------ 98
              :::||||||:::   |||:::         ::::::   |||   |||:::
Db       3803 ACAACAAAGATAAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAA 3744

Qy         99 ------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                   ||||||  ::::::      :::      :::||||||           |||
Db       3743 CAAGAAGAGCTAACTATCCTAAATATATATGCACCCAATACAAGA---------GCACGC 3693

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||         :::|||   ||||||:::  :::|||        :::|||   |||
Db       3692 AGATTC------ATAAAGCAAGTCCTTAGAAACCTACAAAGAGACTTAGACTCCCACACA 3639

Qy        137 ValIlePheAlaGlyAsp------------------------ThrAsnLeuArgAspAr 148
              :::|||      ||||||                       |||:::   ||||||||
Db       3638 ATAATAATG---GGAGACTTTACCACCCCACTGTCAACATTAGACAGATCAACGAGACAG 3582

Qy        148 gGluValThr-ArgCysGlyGlyLeu--------------ProAsn-AsnIleValAsp 162
              |::::::||| |||      |||:::              |||::: :::::::|||
Db       3581 AAAGTTAACAAAGATATCCAGGAATTGAACTAAGTTCTGCACCAAGCGGACCTAATAGAC 3522

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::::::      |||      ||||||     :::||||||:::
Db       3521 ATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 3480

RESULT 56
US-10-868-184D-10373/c
; Sequence 10373, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10373
;   LENGTH: 5647
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10373

Alignment Scores:
Pred. No.:                7.27          Length:          5647
Score:                    86.50         Matches:         51
Percent Similarity:       47.7%         Conservative:    43
Best Local Similarity:    25.9%         Mismatches:      67
Query Match:              7.3%          Indels:          36
DB:                       6             Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10373 (1-5647)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ::::::||||||         |||         :::         :::   :::::::::
Db       4016 GTAAATGGGCTA---------AATGCTCCAGTTAAAAGACACAGACTGGGAAATTGGATA 3966

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                       |||   |||      :::||||||                    ::::::
Db       3965 AACAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCAT 3924

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||  |||  ::::::         |||  |||      :::   |||
Db       3923 CTCAAATGCAGAGACACACACAGGCTCACAATAAAGGGATGGAGGAAGATCTACCAAGCA 3864

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                     ||||||:::  |||  :::         :::          |||
Db       3863 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAA 3804

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------ 98
              :::||||||:::   |||:::         ::::::   |||   |||:::
Db       3803 ACAACAAAGATAAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAA 3744

Qy         99 ------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                   ||||||  ::::::      :::      :::||||||           |||
Db       3743 CAAGAAGAGCTAACTATCCTAAATATATATGCACCCAATACAAGA---------GCACGC 3693

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||         :::|||   ||||||:::  :::|||        :::|||   |||
Db       3692 AGATTC------ATAAAGCAAGTCCTTAGAAACCTACAAAGAGACTTAGACTCCCACACA 3639

Qy        137 ValIlePheAlaGlyAsp------------------------ThrAsnLeuArgAspAr 148
              :::|||      ||||||                       |||:::   ||||||||
Db       3638 ATAATAATG---GGAGACTTTACCACCCCACTGTCAACATTAGACAGATCAACGAGACAG 3582

Qy        148 gGluValThr-ArgCysGlyGlyLeu--------------ProAsn-AsnIleValAsp 162
              |::::::||| |||      |||:::              |||::: :::::::|||
Db       3581 AAAGTTAACAAAGATATCCAGGAATTGAACTAAGTTCTGCACCAAGCGGACCTAATAGAC 3522

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::::::      |||      ||||||     :::||||||:::
Db       3521 ATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 3480
```

```
RESULT 57
US-10-868-184C-10392/c
; Sequence 10392, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10392
;   LENGTH: 36783
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10392

Alignment Scores:
Pred. No.:              146             Length:         36783
Score:                  86.50           Matches:        45
Percent Similarity:     43.3%           Conservative:   39
Best Local Similarity:  23.2%           Mismatches:     77
Query Match:            7.3%            Indels:         33
DB:                     6               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10392 (1-36783)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::         :::         :::   :::::::::
Db      34870 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 34811

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                 |||   |||     :::||||||                 :::::::|||
Db      34810 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 34769

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
              |||   |||   ::::::         |||   |||       :::   |||
Db      34768 AAGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 34709

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                   ||||||:::   |||   :::         :::   |||   |||
Db      34708 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTGTGATGAAATAGACTTTAAACCA 34649

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::         ::::::   |||   |||:::
Db      34648 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 34589

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||           |||   |||
Db      34588 GAAGAGCTAACTATCCTAAATGTATATGCACCCAATACA---------GGAGCATCCAGA 34538

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                    :::|||   ||||||      :::|||         :::|||   |||:::
Db      34537 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 34484

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
                |||       ||||||   |||
Db      34483 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 34427

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                 :::::::::|||:::
Db      34426 GTTAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGTGGACCTAATAGACATC 34367

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
                :::       |||       ||||||      :::|||||:::
Db      34366 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 34328

RESULT 58
US-10-868-184D-10392/c
; Sequence 10392, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10392
;   LENGTH: 36783
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10392

Alignment Scores:
Pred. No.:              146             Length:         36783
Score:                  86.50           Matches:        45
Percent Similarity:     43.3%           Conservative:   39
Best Local Similarity:  23.2%           Mismatches:     77
Query Match:            7.3%            Indels:         33
DB:                     6               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10392 (1-36783)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db      34870 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 34811

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                 |||   |||     :::||||||                 :::::::|||
Db      34810 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 34769

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
              |||   |||   ::::::         |||   |||       :::   |||
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:22:00 ; Search time 102.946 Seconds
(without alignments)
1013.401 Million cell updates/sec

Title: US-10-757-745-6_COPY_2_45
Perfect score: 229
Sequence: 1 GDPGSEIIESVPPAGPEASE.....ENEDDIQFVSEGPSRPVLEY 44

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5007642 seqs, 1185514400 residues

Total number of hits satisfying chosen parameters: 10015284

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170335_4211/app_query.fasta_1
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_240_@runat_09032006_170335_4211 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq1:*
9: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	229	100.0	820	9	US-11-266-748A-73025	Sequence 73025, A
	2	229	100.0	820	9	US-11-266-748A-107600	Sequence 107600,
c	3	229	100.0	820	9	US-11-266-748A-125836	Sequence 125836,
	4	229	100.0	822	9	US-11-266-748A-2119	Sequence 2119, Ap
	5	229	100.0	822	9	US-11-266-748A-178504	Sequence 178504,
	6	229	100.0	834	9	US-11-266-748A-252150	Sequence 252150,
	7	229	100.0	834	9	US-11-266-748A-275725	Sequence 275725,
c	8	229	100.0	834	9	US-11-266-748A-312667	Sequence 312667,
	9	229	100.0	2878	9	US-11-266-748A-27338	Sequence 27338, A
	10	223	97.4	460	9	US-11-266-748A-73024	Sequence 73024, A
	11	223	97.4	460	9	US-11-266-748A-107599	Sequence 107599,
c	12	223	97.4	460	9	US-11-266-748A-125835	Sequence 125835,
	13	223	97.4	4999	6	US-10-567-867-2360	Sequence 2360, Ap
	14	223	97.4	5060	11	US-60-742-219-2547	Sequence 2547, Ap
	15	223	97.4	5097	11	US-60-742-219-2549	Sequence 2549, Ap
	16	223	97.4	5100	9	US-11-266-748A-28533	Sequence 28533, A
	17	169	73.8	485	9	US-11-266-748A-252149	Sequence 252149,
	18	169	73.8	485	9	US-11-266-748A-275724	Sequence 275724,
c	19	169	73.8	485	9	US-11-266-748A-312666	Sequence 312666,
	20	141	61.6	2300	11	US-60-762-056-36462	Sequence 36462, A
	21	85.5	37.3	1807	9	US-11-266-748A-357071	Sequence 357071,
c	22	85.5	37.3	1807	9	US-11-266-748A-440450	Sequence 440450,
	23	85.5	37.3	4860	9	US-11-266-748A-29410	Sequence 29410, A
	24	68	29.7	157	7	US-11-330-083-14432	Sequence 14432, A
c	25	68	29.7	491	9	US-11-266-748A-261855	Sequence 261855,
	26	68	29.7	491	9	US-11-266-748A-322372	Sequence 322372,
	27	68	29.7	559	9	US-11-266-748A-2021	Sequence 2021, Ap
	28	68	29.7	707	9	US-11-266-748A-48763	Sequence 48763, A
	29	68	29.7	991	9	US-11-266-748A-81722	Sequence 81722, A
c	30	68	29.7	991	9	US-11-266-748A-134533	Sequence 134533,
	31	68	29.7	1000	9	US-11-266-748A-281639	Sequence 281639,
c	32	68	29.7	1000	9	US-11-266-748A-308279	Sequence 308279,
	33	68	29.7	2311	9	US-11-266-748A-25318	Sequence 25318, A
	34	68	29.7	3456	9	US-11-266-748A-28122	Sequence 28122, A
c	35	68	29.7	3818	6	US-10-461-673-5739	Sequence 5739, Ap
c	36	68	29.7	3818	6	US-10-498-451-2220	Sequence 2220, Ap
c	37	66	28.8	1000	9	US-11-266-748A-196541	Sequence 196541,
c	38	64	27.9	558	9	US-11-266-748A-480480	Sequence 480480,
	39	64	27.9	1111	9	US-11-266-748A-368562	Sequence 368562,
c	40	64	27.9	1111	9	US-11-266-748A-451941	Sequence 451941,
	41	64	27.9	7078	6	US-10-276-817B-5095	Sequence 5095, Ap
	42	64	27.9	7666	6	US-10-461-673-164	Sequence 164, App
	43	64	27.9	7758	11	US-60-751-455-3391	Sequence 3391, Ap
	44	64	27.9	7762	9	US-11-266-748A-32565	Sequence 32565, A
	45	64	27.9	8862	6	US-10-276-817B-4479	Sequence 4479, Ap
	46	63	27.5	310	9	US-11-266-748A-167917	Sequence 167917,
	47	63	27.5	3120	6	US-10-461-673-5511	Sequence 5511, Ap
	48	62	27.1	963	9	US-11-266-748A-170046	Sequence 170046,
	49	62	27.1	1000	9	US-11-266-748A-222213	Sequence 222213,
	50	62	27.1	1000	9	US-11-266-748A-288201	Sequence 288201,
c	51	62	27.1	1000	9	US-11-266-748A-339630	Sequence 339630,
	52	62	27.1	1000	9	US-11-266-748A-399037	Sequence 399037,
c	53	62	27.1	1000	9	US-11-266-748A-470083	Sequence 470083,
	54	61.5	26.9	465	9	US-11-311-886-771	Sequence 771, App
	55	61	26.6	823	9	US-11-266-748A-169190	Sequence 169190,
	56	61	26.6	898	9	US-11-266-748A-55069	Sequence 55069, A
	57	61	26.6	1139	9	US-11-266-748A-85856	Sequence 85856, A
	58	61	26.6	1139	9	US-11-266-748A-112092	Sequence 112092,
c	59	61	26.6	1139	9	US-11-266-748A-138667	Sequence 138667,
	60	61	26.6	1185	9	US-11-266-748A-169191	Sequence 169191,
	61	61	26.6	1335	9	US-11-266-748A-251111	Sequence 251111,
c	62	61	26.6	1335	9	US-11-266-748A-311628	Sequence 311628,
c	63	61	26.6	1414	9	US-11-266-748A-358938	Sequence 358938,
	64	61	26.6	1414	9	US-11-266-748A-442317	Sequence 442317,
	65	61	26.6	1505	9	US-11-266-748A-26928	Sequence 26928, A
	66	61	26.6	1505	9	US-11-266-748A-31320	Sequence 31320, A
	67	61	26.6	2300	11	US-60-762-056-1344	Sequence 1344, Ap
	68	61	26.6	2300	11	US-60-762-056-1879	Sequence 1879, Ap
	69	61	26.6	2502	11	US-60-751-455-2721	Sequence 2721, Ap
	70	61	26.6	3775	6	US-10-531-147-1235	Sequence 1235, Ap
	71	60.5	26.4	439	9	US-11-266-748A-413970	Sequence 413970,
	72	60.5	26.4	1165	6	US-10-461-673-5295	Sequence 5295, Ap
c	73	60.5	26.4	1538	6	US-10-461-673-4943	Sequence 4943, Ap
c	74	60.5	26.4	2300	11	US-60-762-056-22047	Sequence 22047, A
	75	60.5	26.4	2300	11	US-60-762-056-22048	Sequence 22048, A

ALIGNMENTS

RESULT 1

```
US-11-266-748A-73025
; Sequence 73025, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 73025
;   LENGTH: 820
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-73025

Alignment Scores:
Pred. No.:                2.11e-22        Length:          820
Score:                    229.00          Matches:         44
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-73025 (1-820)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 119

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 179

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        180 GTTCTTGAATAC 191

RESULT 2
US-11-266-748A-107600
; Sequence 107600, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 107600
;   LENGTH: 820
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-107600

Alignment Scores:
Pred. No.:                2.11e-22        Length:          820
Score:                    229.00          Matches:         44
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-107600 (1-820)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 119

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 179

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        180 GTTCTTGAATAC 191

RESULT 3
US-11-266-748A-125836/c
; Sequence 125836, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 125836
```

```
;   LENGTH: 820
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-125836

Alignment Scores:
Pred. No.:              2.11e-22        Length:         820
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-125836 (1-820)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 702

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 642

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        641 GTTCTTGAATAC 630

RESULT 4
US-11-266-748A-2119
; Sequence 2119, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2119
;   LENGTH: 822
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (533)..(533)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (586)..(586)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (648)..(648)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (671)..(671)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (771)..(771)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (775)..(775)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (777)..(777)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (782)..(782)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (784)..(785)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-2119

Alignment Scores:
Pred. No.:              2.11e-22        Length:         822
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-2119 (1-822)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 119

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 179

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        180 GTTCTTGAATAC 191

RESULT 5
US-11-266-748A-178504
; Sequence 178504, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
```

```
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 178504
;   LENGTH: 822
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (533)..(533)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (586)..(586)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (648)..(648)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (671)..(671)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (771)..(771)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (775)..(775)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (777)..(777)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (782)..(782)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (784)..(785)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-178504

Alignment Scores:
Pred. No.:              2.11e-22        Length:         822
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-178504 (1-822)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 119

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 179

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        180 GTTCTTGAATAC 191

RESULT 6
US-11-266-748A-252150
; Sequence 252150, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 252150
;   LENGTH: 834
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (545)..(545)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (598)..(598)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (660)..(660)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (683)..(683)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (783)..(783)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (787)..(787)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (789)..(789)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (794)..(794)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (796)..(797)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-252150

Alignment Scores:
Pred. No.:              2.15e-22        Length:         834
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
```

```
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-252150 (1-834)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 134

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 194

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        195 GTTCTTGAATAC 206

RESULT 7
US-11-266-748A-275725
; Sequence 275725, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 275725
;   LENGTH: 834
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (545)..(545)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (598)..(598)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (660)..(660)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (683)..(683)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (783)..(783)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (787)..(787)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (789)..(789)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (794)..(794)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (796)..(797)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-275725

Alignment Scores:
Pred. No.:              2.15e-22        Length:         834
Score:                  229.00          Matches:        44
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-275725 (1-834)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 134

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 194

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        195 GTTCTTGAATAC 206

RESULT 8
US-11-266-748A-312667/c
; Sequence 312667, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 312667
;   LENGTH: 834
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (38)..(39)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (41)..(41)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (46)..(46)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (48)..(48)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (52)..(52)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152)..(152)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (175)..(175)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (237)..(237)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (290)..(290)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-312667

Alignment Scores:
Pred. No.:               2.15e-22        Length:          834
Score:                   229.00          Matches:         44
Percent Similarity:      100.0%          Conservative:    0
Best Local Similarity:   100.0%          Mismatches:      0
Query Match:             100.0%          Indels:          0
DB:                      9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-312667 (1-834)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 701

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 641

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        640 GTTCTTGAATAC 629

RESULT 9
US-11-266-748A-27338
; Sequence 27338, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27338
;   LENGTH: 2878
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-27338

Alignment Scores:
Pred. No.:               9.9e-22         Length:          2878
Score:                   229.00          Matches:         44
Percent Similarity:      100.0%          Conservative:    0
Best Local Similarity:   100.0%          Mismatches:      0
Query Match:             100.0%          Indels:          0
DB:                      9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-27338 (1-2878)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 233

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATCGAGACCT 293

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        294 GTTCTTGAATAC 305

RESULT 10
US-11-266-748A-73024
; Sequence 73024, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
```

```
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 73024
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-73024

Alignment Scores:
Pred. No.:                7.21e-22        Length:          460
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-73024 (1-460)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 240

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db         241 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 300

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         301 GTTCTTGAATAC 312

RESULT 11
US-11-266-748A-107599
; Sequence 107599, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 107599
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-107599

Alignment Scores:
Pred. No.:                7.21e-22        Length:          460
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-107599 (1-460)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 240

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db         241 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 300

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         301 GTTCTTGAATAC 312

RESULT 12
US-11-266-748A-125835/c
; Sequence 125835, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 125835
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-125835

Alignment Scores:
Pred. No.:                7.21e-22        Length:          460
Score:                    223.00          Matches:         43
Percent Similarity:       97.7%           Conservative:    0
Best Local Similarity:    97.7%           Mismatches:      1
Query Match:              97.4%           Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-125835 (1-460)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         280 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 221

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db         220 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 161

Qy          41 ValLeuGluTyr 44
               ||||||||||||
Db         160 GTTCTTGAATAC 149

RESULT 13
```

```
US-10-567-867-2360
; Sequence 2360, Application US/10567867
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PCT
;  CURRENT APPLICATION NUMBER: US/10/567,867
;  CURRENT FILING DATE:  2006-01-27
;  PRIOR APPLICATION NUMBER: US 60/490,890
;  PRIOR FILING DATE: 2003-07-29
;  NUMBER OF SEQ ID NOS: 2786
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2360
;   LENGTH: 4999
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-567-867-2360

Alignment Scores:
Pred. No.:              1.36e-20        Length:         4999
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-567-867-2360 (1-4999)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 288

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        289 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 348

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        349 GTTCTTGAATAC 360

RESULT 14
US-60-742-219-2547
; Sequence 2547, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2547
;   LENGTH: 5060
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-2547

Alignment Scores:
Pred. No.:              1.38e-20        Length:         5060
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-742-219-2547 (1-5060)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 GGAGACCCGGGGTCGGAAATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 106

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        107 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 166

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        167 GTTCTTGAATAC 178

RESULT 15
US-60-742-219-2549
; Sequence 2549, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2549
;   LENGTH: 5097
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-2549

Alignment Scores:
Pred. No.:              1.4e-20         Length:         5097
Score:                  223.00          Matches:        43
Percent Similarity:     97.7%           Conservative:   0
Best Local Similarity:  97.7%           Mismatches:     1
Query Match:            97.4%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-742-219-2549 (1-5097)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 287

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        288 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 347

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        348 GTTCTTGAATAC 359

RESULT 16
US-11-266-748A-28533
```

```
; Sequence 28533, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28533
;   LENGTH: 5100
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-28533

Alignment Scores:
Pred. No.:                1.4e-20       Length:          5100
Score:                    223.00        Matches:         43
Percent Similarity:       97.7%         Conservative:    0
Best Local Similarity:    97.7%         Mismatches:      1
Query Match:              97.4%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-28533 (1-5100)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 GGAGACCCGGGGTCGGAGATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAG 133

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        134 TCAACAACGGATGAAAATGAAGACGACATTCAGTTTGTCAGTGAAGGACCATTACGACCT 193

Qy         41 ValLeuGluTyr 44
              ||||||||||||
Db        194 GTTCTTGAATAC 205

RESULT 17
US-11-266-748A-252149
; Sequence 252149, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 252149
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-252149

Alignment Scores:
Pred. No.:                2.99e-14      Length:          485
Score:                    169.00        Matches:         34
Percent Similarity:       85.7%         Conservative:    2
Best Local Similarity:    81.0%         Mismatches:      6
Query Match:              73.8%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-252149 (1-485)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||||||||||||||   ||||||||||   ||||||:::|||||||||||||||
Db         91 CCGGGGTCGGAAATAATAGGGTCTGTCCCTCCAATTGGCCCTGATGCATCTGAATCAACA 150

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
              ||||||||||||||:::   |||||||||||||||||||||||||         ||||||
Db        151 ACGGATGAAAATGAAAACCACATTCAATTTGTCAGTGAAGGACCATTACCAACTGTTCTT 210

Qy         43 GluTyr 44
              ||||||
Db        211 GAATAC 216

RESULT 18
US-11-266-748A-275724
; Sequence 275724, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 275724
;   LENGTH: 485
```

```
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-275724

Alignment Scores:
Pred. No.:              2.99e-14        Length:         485
Score:                  169.00          Matches:        34
Percent Similarity:     85.7%           Conservative:   2
Best Local Similarity:  81.0%           Mismatches:     6
Query Match:            73.8%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-275724 (1-485)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||||||||||   ||||||||||||   ||||||:::|||||||||||||||
Db         91 CCGGGGTCGGAAATAATAGGGTCTGTCCCTCCAATTGGCCCTGATGCATCTGAATCAACA 150

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
              ||||||||||||||:::   |||||||||||||||||||||||||          ||||||
Db        151 ACGGATGAAAATGAAAACCACATTCAATTTGTCAGTGAAGGACCATTACCAACTGTTCTT 210

Qy         43 GluTyr 44
              ||||||
Db        211 GAATAC 216

RESULT 19
US-11-266-748A-312666/c
; Sequence 312666, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 312666
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-312666

Alignment Scores:
Pred. No.:              2.99e-14        Length:         485
Score:                  169.00          Matches:        34
Percent Similarity:     85.7%           Conservative:   2
Best Local Similarity:  81.0%           Mismatches:     6
Query Match:            73.8%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-312666 (1-485)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||||||||||   ||||||||||||   ||||||:::|||||||||||||||
Db        395 CCGGGGTCGGAAATAATAGGGTCTGTCCCTCCAATTGGCCCTGATGCATCTGAATCAACA 336

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
              ||||||||||||||:::   |||||||||||||||||||||||||          ||||||
Db        335 ACGGATGAAAATGAAAACCACATTCAATTTGTCAGTGAAGGACCATTACCAACTGTTCTT 276

Qy         43 GluTyr 44
              ||||||
Db        275 GAATAC 270

RESULT 20
US-60-762-056-36462
; Sequence 36462, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 36462
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-36462

Alignment Scores:
Pred. No.:              1.75e-09        Length:         2300
Score:                  141.00          Matches:        28
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            61.6%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-762-056-36462 (1-2300)

Qy          7 IleIleGluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsn 26
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1474 ATAATAGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAAAT 1533

Qy         27 GluAspAspIleGlnPheValSer 34
              |||||||||||||||||||||||
Db       1534 GAAGACGACATTCAGTTTGTCAGT 1557

RESULT 21
US-11-266-748A-357071
; Sequence 357071, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
```

```
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 357071
;   LENGTH: 1807
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-357071
```

Alignment Scores:

Pred. No.:	0.0822	Length:	1807
Score:	85.50	Matches:	19
Percent Similarity:	53.2%	Conservative:	6
Best Local Similarity:	40.4%	Mismatches:	15
Query Match:	37.3%	Indels:	7
DB:	9	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-357071 (1-1807)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAla----------------------Gly 15
              ||||||              ::::::   |||   |||                       |||
Db        993 CCTGGAGGACCAGCGGCTCAGGCTGCACCAGCAGCACCAGGGCTTTCCCTGGGTGAGGGC 1052

Qy         16 ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                ||||||:::   |||:::   ||||||   |||||||||:::   |||||||||
Db       1053 CGTGAAGCTGCCACAAGCAGTGACGACGAGGAGGAAGATGATGTTGTTTTTGTTTCCTCT 1112

Qy         36 GlyProSerArgProValLeu 42
                |||       |||:::|||
Db       1113 AAGCCTGGGAGCCCCCTACTC 1133
```

```
RESULT 22
US-11-266-748A-440450/c
; Sequence 440450, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 440450
;   LENGTH: 1807
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-440450
```

Alignment Scores:

Pred. No.:	0.0822	Length:	1807
Score:	85.50	Matches:	19
Percent Similarity:	53.2%	Conservative:	6
Best Local Similarity:	40.4%	Mismatches:	15
Query Match:	37.3%	Indels:	7
DB:	9	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-440450 (1-1807)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAla----------------------Gly 15
              ||||||              ::::::   |||   |||                       |||
Db        815 CCTGGAGGACCAGCGGCTCAGGCTGCACCAGCAGCACCAGGGCTTTCCCTGGGTGAGGGC 756

Qy         16 ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                ||||||:::   |||:::   ||||||   |||||||||:::   |||||||||
Db        755 CGTGAAGCTGCCACAAGCAGTGACGACGAGGAGGAAGATGATGTTGTTTTTGTTTCCTCT 696

Qy         36 GlyProSerArgProValLeu 42
                |||       |||:::|||
Db        695 AAGCCTGGGAGCCCCCTACTC 675
```

```
RESULT 23
US-11-266-748A-29410
; Sequence 29410, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 29410
;   LENGTH: 4860
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-29410
```

Alignment Scores:

Pred. No.:	0.278	Length:	4860
Score:	85.50	Matches:	19
Percent Similarity:	53.2%	Conservative:	6
Best Local Similarity:	40.4%	Mismatches:	15
Query Match:	37.3%	Indels:	7
DB:	9	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-29410 (1-4860)

```
Qy          3 ProGlySerGluIleIleGluSerValProProAla----------------------Gly 15
              ||||||              ::::::   |||   |||                       |||
Db        993 CCTGGAGGACCAGCGGCTCAGGCTGCACCAGCAGCACCAGGGCTTTCCCTGGGTGAGGGC 1052

Qy         16 ProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGlu 35
                ||||||:::   |||:::   ||||||   |||||||||:::   |||||||||
Db       1053 CGTGAAGCTGCCACAAGCAGTGACGACGAGGAGGAAGATGATGTTGTTTTTGTTTCCTCT 1112
```

```
Qy         36 GlyProSerArgProValLeu 42
                 |||       |||:::|||
Db       1113 AAGCCTGGGAGCCCCCTACTC 1133

RESULT 24
US-11-330-083-14432
; Sequence 14432, Application US/11330083
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.355 - 38-21(15401)C/US
;  CURRENT APPLICATION NUMBER: US/11/330,083
;  CURRENT FILING DATE:  2006-01-12
;  PRIOR APPLICATION NUMBER: US 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 14432
;   LENGTH: 157
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 701068758H1
US-11-330-083-14432

Alignment Scores:
Pred. No.:              1.17          Length:         157
Score:                  68.00         Matches:        17
Percent Similarity:     52.4%         Conservative:   5
Best Local Similarity:  40.5%         Mismatches:     17
Query Match:            29.7%         Indels:         3
DB:                     7             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-330-083-14432 (1-157)

Qy          2 AspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSer 21
              ||||||                |||      ||||||      |||   |||:::|||::: 
Db         28 GATCCGTTCGAAGCAGCGTTCGAGGAATCGCCGCCAGAATCCCCAACGGCTACGGAGACA 87

Qy         22 ThrThrAspGluAsnGluAsp------AspIleGlnPheVal-SerGluGlyProSerAr 39
                |||:::|||   ||||||      |||::::::       |||      |||   ||
Db         88 GAGACAGAGGAGGAGGAGGACAACATGGACGTGGAGCTCGCCAAGCTCCCTTCCACGGCG 147

Qy         39 gPro 40
              ||||
Db        148 ACCC 151

RESULT 25
US-11-266-748A-261855/c
; Sequence 261855, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 261855
;   LENGTH: 491
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-261855

Alignment Scores:
Pred. No.:              4.76          Length:         491
Score:                  68.00         Matches:        14
Percent Similarity:     48.8%         Conservative:   7
Best Local Similarity:  32.6%         Mismatches:     22
Query Match:            29.7%         Indels:         0
DB:                     9             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-261855 (1-491)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||::::::|||       |||   |||   |||         |||
Db        221 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 162

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                     |||::::::      |||    :::|||       |||    |||:::
Db        161 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 102

Qy         41 ValLeuGlu 43
                 :::|||
Db        101 GAGATGGAG 93

RESULT 26
US-11-266-748A-322372
; Sequence 322372, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 322372
;   LENGTH: 491
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-322372

Alignment Scores:
Pred. No.:              4.76          Length:         491
Score:                  68.00         Matches:        14
```

```
Percent Similarity:     48.8%          Conservative:  7
Best Local Similarity:  32.6%          Mismatches:    22
Query Match:            29.7%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-322372 (1-491)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||  |||      |||:::::::|||      |||  |||  |||          |||
Db        271 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 330

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||::::::      |||   :::|||          |||   |||:::
Db        331 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 390

Qy         41 ValLeuGlu 43
                 :::|||
Db        391 GAGATGGAG 399

RESULT 27
US-11-266-748A-2021
; Sequence 2021, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2021
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-2021

Alignment Scores:
Pred. No.:              5.58           Length:        559
Score:                  68.00          Matches:       14
Percent Similarity:     48.8%          Conservative:  7
Best Local Similarity:  32.6%          Mismatches:    22
Query Match:            29.7%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-2021 (1-559)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||  |||      |||:::::::|||      |||  |||  |||          |||
Db        244 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 303

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||::::::      |||   :::|||          |||   |||:::
Db        304 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 363

Qy         41 ValLeuGlu 43
                 :::|||
Db        364 GAGATGGAG 372

RESULT 28
US-11-266-748A-48763
; Sequence 48763, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 48763
;   LENGTH: 707
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (679)..(679)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (685)..(686)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (688)..(688)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-48763

Alignment Scores:
Pred. No.:              7.45           Length:        707
Score:                  68.00          Matches:       14
Percent Similarity:     48.8%          Conservative:  7
Best Local Similarity:  32.6%          Mismatches:    22
Query Match:            29.7%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-48763 (1-707)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||  |||      |||:::::::|||      |||  |||  |||          |||
Db        368 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 427

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                  |||::::::      |||   :::|||          |||   |||:::
Db        428 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 487

Qy         41 ValLeuGlu 43
                 :::|||
Db        488 GAGATGGAG 496
```

```
RESULT 29
US-11-266-748A-81722
; Sequence 81722, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81722
;   LENGTH: 991
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (286)..(358)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-81722

Alignment Scores:
Pred. No.:                11.3          Length:          991
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-81722 (1-991)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||         |||
Db        684 GGGCCCCCTCCCCAGGAGCTTGTTGAAAAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 743

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::      |||   :::|||         |||   |||:::
Db        744 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 803

Qy         41 ValLeuGlu 43
                 :::|||
Db        804 GAGATGGAG 812

RESULT 30
US-11-266-748A-134533/c
; Sequence 134533, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 134533
;   LENGTH: 991
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (634)..(706)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-134533

Alignment Scores:
Pred. No.:                11.3          Length:          991
Score:                    68.00         Matches:         14
Percent Similarity:       48.8%         Conservative:    7
Best Local Similarity:    32.6%         Mismatches:      22
Query Match:              29.7%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-134533 (1-991)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||         |||
Db        308 GGGCCCCCTCCCCAGGAGCTTGTTGAAAAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 249

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::      |||   :::|||         |||   |||:::
Db        248 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 189

Qy         41 ValLeuGlu 43
                 :::|||
Db        188 GAGATGGAG 180

RESULT 31
US-11-266-748A-281639
; Sequence 281639, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
```

```
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 281639
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (87)..(87)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-281639

Alignment Scores:
Pred. No.:              11.4          Length:         1000
Score:                  68.00         Matches:        14
Percent Similarity:     48.8%         Conservative:   7
Best Local Similarity:  32.6%         Mismatches:     22
Query Match:            29.7%         Indels:         0
DB:                     9             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-281639 (1-1000)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||::::::|||       |||   |||   |||           |||
Db        691 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 750

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::      |||   :::|||         |||   |||:::
Db        751 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 810

Qy         41 ValLeuGlu 43
                 :::|||
Db        811 GAGATGGAG 819

RESULT 32
US-11-266-748A-308279/c
; Sequence 308279, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 308279
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (914)..(914)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-308279

Alignment Scores:
Pred. No.:              11.4          Length:         1000
Score:                  68.00         Matches:        14
Percent Similarity:     48.8%         Conservative:   7
Best Local Similarity:  32.6%         Mismatches:     22
Query Match:            29.7%         Indels:         0
DB:                     9             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-308279 (1-1000)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||       |||::::::|||       |||   |||   |||           |||
Db        310 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 251

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::      |||   :::|||         |||   |||:::
Db        250 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 191

Qy         41 ValLeuGlu 43
                 :::|||
Db        190 GAGATGGAG 182

RESULT 33
US-11-266-748A-25318
; Sequence 25318, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25318
;   LENGTH: 2311
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-25318

Alignment Scores:
Pred. No.:              32.1          Length:         2311
Score:                  68.00         Matches:        14
Percent Similarity:     48.8%         Conservative:   7
Best Local Similarity:  32.6%         Mismatches:     22
Query Match:            29.7%         Indels:         0
DB:                     9             Gaps:           0
```

```
US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-25318 (1-2311)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::::|||      |||   |||   |||         |||
Db       1980 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 2039

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||:::::      |||   :::|||         |||   |||:::
Db       2040 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 2099

Qy         41 ValLeuGlu 43
                :::|||
Db       2100 GAGATGGAG 2108

RESULT 34
US-11-266-748A-28122
; Sequence 28122, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28122
;   LENGTH: 3456
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-28122

Alignment Scores:
Pred. No.:              52.7          Length:         3456
Score:                  68.00         Matches:        14
Percent Similarity:     48.8%         Conservative:   7
Best Local Similarity:  32.6%         Mismatches:     22
Query Match:            29.7%         Indels:         0
DB:                     9             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-28122 (1-3456)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::::|||      |||   |||   |||         |||
Db       3139 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 3198

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                ||:::::      |||   :::|||         |||   |||:::
Db       3199 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 3258

Qy         41 ValLeuGlu 43
                :::|||
Db       3259 GAGATGGAG 3267

RESULT 35
US-10-461-673-5739/c
; Sequence 5739, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 5739
;   LENGTH: 3818
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-461-673-5739

Alignment Scores:
Pred. No.:              59.6          Length:         3818
Score:                  68.00         Matches:        14
Percent Similarity:     48.8%         Conservative:   7
Best Local Similarity:  32.6%         Mismatches:     22
Query Match:            29.7%         Indels:         0
DB:                     6             Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-461-673-5739 (1-3818)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::::|||      |||   |||   |||         |||
Db        316 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 257
```

```
Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               |||::::::     |||   :::|||          |||  |||:::
Db         256 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 197

Qy          41 ValLeuGlu 43
                :::|||
Db         196 GAGATGGAG 188
```

RESULT 36
US-10-498-451-2220/c
; Sequence 2220, Application US/10498451
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-rui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Ghosh, Malabika
; APPLICANT: Xue, Aidong J.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Dunrui
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wehrman, Tom
; APPLICANT: Weng, Gezhi
; APPLICANT: Xu, Chongjun
; APPLICANT: Boyle, Bryan J
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 820/PCT
; CURRENT APPLICATION NUMBER: US/10/498,451
; CURRENT FILING DATE: 2004-06-09
; PRIOR APPLICATION NUMBER: US 60/339,739
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 10/128,558
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: US 60/339,453
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/365,384
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US 60/365,091
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US 60/372,615
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US 60/376,045
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 60/372,381
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 3134
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 2220
; LENGTH: 3818
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-498-451-2220

Alignment Scores:

Pred. No.:	59.6	Length:	3818
Score:	68.00	Matches:	14
Percent Similarity:	48.8%	Conservative:	7
Best Local Similarity:	32.6%	Mismatches:	22
Query Match:	29.7%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-498-451-2220 (1-3818)

```
Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
               |||   |||       |||::::::|||       |||   |||   |||          |||
Db         316 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 257

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
               |||::::::     |||   :::|||          |||  |||:::
Db         256 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 197

Qy          41 ValLeuGlu 43
                :::|||
Db         196 GAGATGGAG 188
```

RESULT 37
US-11-266-748A-196541/c
; Sequence 196541, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 196541
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-196541

Alignment Scores:

Pred. No.:	21.8	Length:	1000
Score:	66.00	Matches:	15
Percent Similarity:	55.8%	Conservative:	9
Best Local Similarity:	34.9%	Mismatches:	13
Query Match:	28.8%	Indels:	6
DB:	9	Gaps:	1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-196541 (1-1000)

```
Qy           4 GlySerGluIleIleGluSerValPro----------------ProAlaGlyProGlu 17
               ||||||:::   :::   |||  |||                |||      |||
Db         839 GGGAGCAAGACTGTGCCCTCCTCACCACGACATCCCCCTACTTTACCACCGTGGCCAGCA 780

Qy          18 AlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyPro 37
                :::   |||:::   :::   :::||||||:::       |||:::   :::||||||
Db         779 CGTAACAGGTCATCTGTAAACGCTGATGAAGATGAAGGTGGTTTCCTTCTAAAAGGGCCC 720

Qy          38 SerArgPro 40
               |||||||||
Db         719 AGTAGGCCC 711
```

RESULT 38

```
US-11-266-748A-480480/c
; Sequence 480480, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 480480
;   LENGTH: 558
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-480480

Alignment Scores:
Pred. No.:                20.3          Length:          558
Score:                    64.00         Matches:         13
Percent Similarity:       50.0%         Conservative:    6
Best Local Similarity:    34.2%         Mismatches:      19
Query Match:              27.9%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-480480 (1-558)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::::::|||      |||   |||   |||         |||
Db        305 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 246

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSer 38
                |||::::::      |||  :::|||         |||   |||:::
Db        245 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCA 192

RESULT 39
US-11-266-748A-368562
; Sequence 368562, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 368562
;   LENGTH: 1111
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-368562

Alignment Scores:
Pred. No.:                47.5          Length:          1111
Score:                    64.00         Matches:         17
Percent Similarity:       45.7%         Conservative:    4
Best Local Similarity:    37.0%         Mismatches:      17
Query Match:              27.9%         Indels:          8
DB:                       9             Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-368562 (1-1111)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        306 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 365

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||      ||||||||||||   |||   |||:::   |||            :::
Db        366 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 425

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db        426 GAGTCGCCCTCGCACCCG 443

RESULT 40
US-11-266-748A-451941/c
; Sequence 451941, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 451941
;   LENGTH: 1111
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-451941
```

```
Alignment Scores:
Pred. No.:                47.5          Length:         1111
Score:                    64.00         Matches:        17
Percent Similarity:       45.7%         Conservative:   4
Best Local Similarity:    37.0%         Mismatches:     17
Query Match:              27.9%         Indels:         8
DB:                       9             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-451941 (1-1111)

Qy           3 ProGlySerGluIleIleGluSerValProProAla------------------------ 14
               |||   |||      :::|||:::   ||||||
Db         806 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 747

Qy          15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
               |||       ||||||||||||   |||   |||:::   |||           :::
Db         746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 687

Qy          35 GluGlyProSerArgPro 40
               |||   ||||||   |||
Db         686 GAGTCGCCCTCGCACCCG 669

RESULT 41
US-10-276-817B-5095
; Sequence 5095, Application US/10276817B
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-104
;  CURRENT APPLICATION NUMBER: US/10/276,817B
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: PCT/US01/14827
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 16102
;  SOFTWARE: Custom
; SEQ ID NO 5095
;   LENGTH: 7078
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1104)..(1403)
;   OTHER INFORMATION: 62% homologous to Homo sapiens Human myelin oligodendrocyte
;   OTHER INFORMATION: glycoprotein gene,accession number AAZ29204,Smith-Waterman Score=
;   OTHER INFORMATION: 263.
US-10-276-817B-5095

Alignment Scores:
Pred. No.:                465           Length:         7078
Score:                    64.00         Matches:        17
Percent Similarity:       45.7%         Conservative:   4
Best Local Similarity:    37.0%         Mismatches:     17
Query Match:              27.9%         Indels:         8
DB:                       6             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-276-817B-5095 (1-7078)

Qy           3 ProGlySerGluIleIleGluSerValProProAla------------------------ 14
               |||   |||      :::|||:::   ||||||
Db        1716 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 1775

Qy          15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
               |||       ||||||||||||   |||   |||:::   |||           :::
Db        1776 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 1835

Qy          35 GluGlyProSerArgPro 40
               |||   ||||||   |||
Db        1836 GAGTCGCCCTCGCACCCG 1853

RESULT 42
US-10-461-673-164
; Sequence 164, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 164
;   LENGTH: 7666
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (632)..(1552)
US-10-461-673-164

Alignment Scores:
Pred. No.:                513           Length:         7666
Score:                    64.00         Matches:        17
Percent Similarity:       45.7%         Conservative:   4
Best Local Similarity:    37.0%         Mismatches:     17
Query Match:              27.9%         Indels:         8
DB:                       6             Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-461-673-164 (1-7666)

Qy           3 ProGlySerGluIleIleGluSerValProProAla------------------------ 14
               |||   |||      :::|||:::   ||||||
```

```
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||       ||||||||||||   |||   |||:::   |||            :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||  |||
Db        806 GAGTCGCCCTCGCACCCG 823

RESULT 43
US-60-751-455-3391
; Sequence 3391, Application US/60751455
; GENERAL INFORMATION:
;  APPLICANT: Radich, Jerald P.
;  APPLICANT:  Dai, Hongyue
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Schelter, Janell
;  APPLICANT:  Linsley, Peter S.
;  TITLE OF INVENTION: GENES ASSOCIATED WITH PROGRESSION AND RESPONSE IN
;  TITLE OF INVENTION:  CHRONIC MYELOID LEUKEMIA AND USES THEREOF
;  FILE REFERENCE: 9301-253-888
;  CURRENT APPLICATION NUMBER: US/60/751,455
;  CURRENT FILING DATE:  2005-12-15
;  NUMBER OF SEQ ID NOS: 7936
;  SOFTWARE: FastSEQ 4.01
; SEQ ID NO 3391
;   LENGTH: 7758
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-751-455-3391

Alignment Scores:
Pred. No.:              521          Length:         7758
Score:                  64.00        Matches:        17
Percent Similarity:     45.7%        Conservative:   4
Best Local Similarity:  37.0%        Mismatches:     17
Query Match:            27.9%        Indels:         8
DB:                     11           Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-60-751-455-3391 (1-7758)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||       ||||||||||||   |||   |||:::   |||            :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||  |||
Db        806 GAGTCGCCCTCGCACCCG 823

RESULT 44
US-11-266-748A-32565
; Sequence 32565, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32565
;   LENGTH: 7762
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-32565

Alignment Scores:
Pred. No.:              521          Length:         7762
Score:                  64.00        Matches:        17
Percent Similarity:     45.7%        Conservative:   4
Best Local Similarity:  37.0%        Mismatches:     17
Query Match:            27.9%        Indels:         8
DB:                     9            Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-32565 (1-7762)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db        686 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 745

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||       ||||||||||||   |||   |||:::   |||            :::
Db        746 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 805

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||  |||
Db        806 GAGTCGCCCTCGCACCCG 823

RESULT 45
US-10-276-817B-4479
; Sequence 4479, Application US/10276817B
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-104
;  CURRENT APPLICATION NUMBER: US/10/276,817B
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: PCT/US01/14827
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 16102
;  SOFTWARE: Custom
; SEQ ID NO 4479
;   LENGTH: 8862
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (8052)..(8336)
;   OTHER INFORMATION: 63% homologous to Homo sapiens Human myelin oligodendrocyte
;   OTHER INFORMATION: glycoprotein gene,accession number AAZ29204,Smith-Waterman Score=
;   OTHER INFORMATION: 257.
US-10-276-817B-4479

Alignment Scores:
Pred. No.:              614          Length:         8862
Score:                  64.00        Matches:        17
Percent Similarity:     45.7%        Conservative:   4
Best Local Similarity:  37.0%        Mismatches:     17
Query Match:            27.9%        Indels:         8
DB:                     6            Gaps:           1
```

```
US-10-757-745-6_COPY_2_45 (1-44) x US-10-276-817B-4479 (1-8862)

Qy          3 ProGlySerGluIleIleGluSerValProProAla---------------------- 14
              |||   |||      :::|||:::   ||||||
Db       3260 CCCACCTCTTCTCCCCTGGAGGCACCACCCCCTGTGCTGCAGGACGGCGATGGCTCCCTG 3319

Qy         15 GlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGlnPheValSer 34
              |||       ||||||||||||   |||   |||:::   |||           :::
Db       3320 GGGGACGGTGCATCAGAGAGTGAGACCACTGAGTCTGCGGACAGTGAGAATGACATGGGC 3379

Qy         35 GluGlyProSerArgPro 40
              |||   ||||||   |||
Db       3380 GAGTCGCCCTCGCACCCG 3397

RESULT 46
US-11-266-748A-167917
; Sequence 167917, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 167917
;   LENGTH: 310
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-167917

Alignment Scores:
Pred. No.:              13.6          Length:          310
Score:                  63.00         Matches:         15
Percent Similarity:     48.8%         Conservative:    6
Best Local Similarity:  34.9%         Mismatches:      18
Query Match:            27.5%         Indels:          4
DB:                     9             Gaps:            1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-167917 (1-310)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||:::||||||      |||   |||  |||          |||
Db          4 GGGCCCCCTCCCCAGGAGCTTATTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 63

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 ||||::::::      |||   :::|||               |||:::|||
Db         64 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAG-----------GCCCCCACCAGAGTA 111

Qy         41 ValLeuGlu 43
                 |||:::
Db        112 CACCTGCAG 120
```

```
RESULT 47
US-10-461-673-5511
; Sequence 5511, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 5511
;   LENGTH: 3120
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (115)..(3066)
US-10-461-673-5511

Alignment Scores:
Pred. No.:              234           Length:          3120
Score:                  63.00         Matches:         13
Percent Similarity:     46.5%         Conservative:    7
Best Local Similarity:  30.2%         Mismatches:      23
Query Match:            27.5%         Indels:          0
DB:                     6             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-461-673-5511 (1-3120)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
```

```
                ||| ||| |||::::::|||  |||  |||  |||
Db       2812 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTCCTCCCCCAACCCTGTTGGAA 2871

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 |||::::::     |||  :::|||     |||  |||:::
Db       2872 GAGGAGACTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 2931

Qy         41 ValLeuGlu 43
                :::|||
Db       2932 GAGATGGAG 2940

RESULT 48
US-11-266-748A-170046
; Sequence 170046, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 170046
;   LENGTH: 963
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-170046

Alignment Scores:
Pred. No.:                76              Length:         963
Score:                    62.00           Matches:        14
Percent Similarity:       45.0%           Conservative:   4
Best Local Similarity:    35.0%           Mismatches:     22
Query Match:              27.1%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-170046 (1-963)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||:::|||   |||  ||||||  |||||||||   |||
Db        402 CCAGGATCCCAGATTAAAAACTCCACCCCGCCCCCAGGGCCAGAGACCAGCTCCGGAGGA 461

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
                :::   :::         :::                          |||    ||||||
Db        462 GCAGAGGGGGATTCCAGCCATTTGGTGATACCGGGGGCTTCCACTTCTCATTTGGTGTTG 521

RESULT 49
US-11-266-748A-222213
; Sequence 222213, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 222213
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-222213

Alignment Scores:
Pred. No.:                79.7            Length:         1000
Score:                    62.00           Matches:        13
Percent Similarity:       46.5%           Conservative:   7
Best Local Similarity:    30.2%           Mismatches:     23
Query Match:              27.1%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-222213 (1-1000)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||  |||      |||::::::|||      |||  |||  |||         |||
Db        746 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 805

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 ::::::     |||  :::|||     |||  |||:::
Db        806 GAGGAGCCTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 865

Qy         41 ValLeuGlu 43
                :::|||
Db        866 GAGATGGAG 874

RESULT 50
US-11-266-748A-288201
; Sequence 288201, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
```

```
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 288201
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-288201

Alignment Scores:
Pred. No.:              79.7          Length:        1000
Score:                  62.00         Matches:       13
Percent Similarity:     46.5%         Conservative:  7
Best Local Similarity:  30.2%         Mismatches:    23
Query Match:            27.1%         Indels:        0
DB:                     9             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-288201 (1-1000)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||          |||
Db        746 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 805

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                     ::::::      |||   :::|||          |||   |||:::
Db        806 GAGGAGCCTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 865

Qy         41 ValLeuGlu 43
                 :::|||
Db        866 GAGATGGAG 874

RESULT 51
US-11-266-748A-339630/c
; Sequence 339630, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 339630
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-339630

Alignment Scores:
Pred. No.:              79.7          Length:        1000
Score:                  62.00         Matches:       13
Percent Similarity:     46.5%         Conservative:  7
Best Local Similarity:  30.2%         Mismatches:    23
Query Match:            27.1%         Indels:        0
DB:                     9             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-339630 (1-1000)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||          |||
Db        255 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 196

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                     ::::::      |||   :::|||          |||   |||:::
Db        195 GAGGAGCCTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 136

Qy         41 ValLeuGlu 43
                 :::|||
Db        135 GAGATGGAG 127

RESULT 52
US-11-266-748A-399037
; Sequence 399037, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 399037
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-399037

Alignment Scores:
Pred. No.:              79.7          Length:        1000
Score:                  62.00         Matches:       13
Percent Similarity:     46.5%         Conservative:  7
Best Local Similarity:  30.2%         Mismatches:    23
Query Match:            27.1%         Indels:        0
DB:                     9             Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-399037 (1-1000)

Qy          1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
              |||   |||      |||::::::|||      |||   |||   |||          |||
Db        746 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 805

Qy         21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                     ::::::      |||   :::|||          |||   |||:::
Db        806 GAGGAGCCTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 865
```

```
Qy          41 ValLeuGlu 43
                  :::|||
Db         866 GAGATGGAG 874

RESULT 53
US-11-266-748A-470083/c
; Sequence 470083, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 470083
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-470083

Alignment Scores:
Pred. No.:                79.7          Length:          1000
Score:                    62.00         Matches:         13
Percent Similarity:       46.5%         Conservative:    7
Best Local Similarity:    30.2%         Mismatches:      23
Query Match:              27.1%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-470083 (1-1000)

Qy           1 GlyAspProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGlu 20
                 |||   |||      |||:::::::|||      |||   |||   |||         |||
Db         255 GGGCCCCCTCCCCAGGAGCTTGTTGAAGAAGAGCCCTCTGCTCCCCCAACCCTGTTGGAA 196

Qy          21 SerThrThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                   ::::::      |||   :::|||          |||   |||:::
Db         195 GAGGAGCCTGAGGATGGGAGTGACAAGGTGCAGCCCCCACCAGAGACACCTGCAGAAGAA 136

Qy          41 ValLeuGlu 43
                  :::|||
Db         135 GAGATGGAG 127

RESULT 54
US-11-311-886-771
; Sequence 771, Application US/11311886
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
;  TITLE OF INVENTION:  PLANTS
;  FILE REFERENCE: 16517.358 - 38-21(15598)C/US
;  CURRENT APPLICATION NUMBER: US/11/311,886
;  CURRENT FILING DATE:  2005-12-20
;  PRIOR APPLICATION NUMBER: 09/421,106
;  PRIOR FILING DATE: 1999-10-15
;  NUMBER OF SEQ ID NOS: 36938
; SEQ ID NO 771
;   LENGTH: 465
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
US-11-311-886-771

Alignment Scores:
Pred. No.:                36.5          Length:          465
Score:                    61.50         Matches:         18
Percent Similarity:       43.3%         Conservative:    8
Best Local Similarity:    30.0%         Mismatches:      15
Query Match:              26.9%         Indels:          19
DB:                       9             Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x US-11-311-886-771 (1-465)

Qy           2 AspProGlySerGluIleIleGluSerValProProAlaGlyPro--------------- 16
                 ||||||   |||        :::|||:::   ||||||:::   |||
Db         115 GATCCATCTTCTCCCGAACTTGAAGCAGCTCCCCCATCTTCACCTATTATTATCATCTCT 174

Qy          17 ---GluAlaSerGluSerThrThrAspGlu------------------------------ 25
                   :::   :::|||||||||       |||
Db         175 GAAGACCCTACAGAGTCAACATTTGGAGAAGCTGTTGCTCTCTCTGATTCCCCTGTTNTT 234

Qy          26 ---------AsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
                        ||||||::::::   |||      |||:::           |||||||||
Db         235 CATCTGATGAATGAGGAGGAGACACAGGATCAGTCACAGGATTCTTANATTCCTGTCCTT 294

RESULT 55
US-11-266-748A-169190
; Sequence 169190, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 169190
;   LENGTH: 823
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-169190

Alignment Scores:
Pred. No.:                86.6          Length:          823
Score:                    61.00         Matches:         11
Percent Similarity:       62.1%         Conservative:    7
```

```
Best Local Similarity:  37.9%          Mismatches:    11
Query Match:            26.6%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-169190 (1-823)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
              ||||||         :::|||:::||||||:::::|||   ::::::   |||
Db        377 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 436

Qy         32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db        437 GAGGATGATGAAGCTCCTTCTAAGCCA 463

RESULT 56
US-11-266-748A-55069
; Sequence 55069, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 55069
;   LENGTH: 898
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (870)..(875)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-55069

Alignment Scores:
Pred. No.:              96.4           Length:        898
Score:                  61.00          Matches:       11
Percent Similarity:     62.1%          Conservative:  7
Best Local Similarity:  37.9%          Mismatches:    11
Query Match:            26.6%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-55069 (1-898)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
              ||||||         :::|||:::||||||:::::|||   ::::::   |||
Db         52 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 111

Qy         32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db        112 GAGGATGATGAAGCTCCTTCTAAGCCA 138

RESULT 57
US-11-266-748A-85856
; Sequence 85856, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 85856
;   LENGTH: 1139
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-85856

Alignment Scores:
Pred. No.:              129            Length:        1139
Score:                  61.00          Matches:       15
Percent Similarity:     47.5%          Conservative:  4
Best Local Similarity:  37.5%          Mismatches:    15
Query Match:            26.6%          Indels:        6
DB:                     9              Gaps:          1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-85856 (1-1139)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||                  ||||||||||||:::      ||||||:::|||
Db        269 CCAGCTCCTAACGCCTTCTCTTCTGTGCCCCCATCTCCAGAGGAGGCTGCGGAGGAGGAG 328

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
                 :::|||   |||                              ||||||||||||||:::
Db        329 GAGGAAGAAGAGGAA-----------------CGCCCCAAACCAAGCCGCCCCGTGGTG 370

RESULT 58
US-11-266-748A-112092
; Sequence 112092, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
```

```
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 112092
;    LENGTH: 1139
;    TYPE: DNA
;    ORGANISM: Homo Sapiens
US-11-266-748A-112092

Alignment Scores:
Pred. No.:                129             Length:         1139
Score:                    61.00           Matches:        15
Percent Similarity:       47.5%           Conservative:   4
Best Local Similarity:    37.5%           Mismatches:     15
Query Match:              26.6%           Indels:         6
DB:                       9               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-112092 (1-1139)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||                  |||||||||||:::      ||||||:::|||
Db        269 CCAGCTCCTAACGCCTTCTCTTCTGTGCCCCCATCTCCAGAGGAGGCTGCGGAGGAGGAG 328

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
                ::|||   |||                 ||||||||||||||:::
Db        329 GAGGAAGAAGAGGAA-----------------CGCCCCAAACCAAGCCGCCCCGTGGTG 370

RESULT 59
US-11-266-748A-138667/c
; Sequence 138667, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 138667
;    LENGTH: 1139
;    TYPE: DNA
;    ORGANISM: Homo Sapiens
US-11-266-748A-138667

Alignment Scores:
Pred. No.:                129             Length:         1139
Score:                    61.00           Matches:        15
Percent Similarity:       47.5%           Conservative:   4
Best Local Similarity:    37.5%           Mismatches:     15
Query Match:              26.6%           Indels:         6
DB:                       9               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-138667 (1-1139)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||                  |||||||||||:::      ||||||:::|||
Db        871 CCAGCTCCTAACGCCTTCTCTTCTGTGCCCCCATCTCCAGAGGAGGCTGCGGAGGAGGAG 812

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgProValLeu 42
                ::|||   |||                 ||||||||||||||:::
Db        811 GAGGAAGAAGAGGAA-----------------CGCCCCAAACCAAGCCGCCCCGTGGTG 770

RESULT 60
US-11-266-748A-169191
; Sequence 169191, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 169191
;    LENGTH: 1185
;    TYPE: DNA
;    ORGANISM: Homo Sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (42)..(42)
;    OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-169191

Alignment Scores:
Pred. No.:                136             Length:         1185
Score:                    61.00           Matches:        11
Percent Similarity:       62.1%           Conservative:   7
Best Local Similarity:    37.9%           Mismatches:     11
Query Match:              26.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-169191 (1-1185)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
              ||||||          :::|||:::||||||:::::::|||     ::::::     |||
Db        438 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 497

Qy         32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db        498 GAGGATGATGAAGCTCCTTCTAAGCCA 524
```

```
RESULT 61
US-11-266-748A-251111
; Sequence 251111, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 251111
;   LENGTH: 1335
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (664)..(705)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-251111

Alignment Scores:
Pred. No.:              157          Length:        1335
Score:                  61.00        Matches:       11
Percent Similarity:     62.1%        Conservative:  7
Best Local Similarity:  37.9%        Mismatches:    11
Query Match:            26.6%        Indels:        0
DB:                     9            Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-251111 (1-1335)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
              ||||||         :::|||:::||||||:::::|||   ::::::   |||
Db       1022 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 1081

Qy         32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db       1082 GAGGATGATGAAGCTCCTTCTAAGCCA 1108

RESULT 62
US-11-266-748A-311628/c
; Sequence 311628, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 311628
;   LENGTH: 1335
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (631)..(672)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-311628

Alignment Scores:
Pred. No.:              157          Length:        1335
Score:                  61.00        Matches:       11
Percent Similarity:     62.1%        Conservative:  7
Best Local Similarity:  37.9%        Mismatches:    11
Query Match:            26.6%        Indels:        0
DB:                     9            Gaps:          0

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-311628 (1-1335)

Qy         12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
              ||||||         :::|||:::||||||:::::|||   ::::::   |||
Db        314 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 255

Qy         32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db        254 GAGGATGATGAAGCTCCTTCTAAGCCA 228

RESULT 63
US-11-266-748A-358938/c
; Sequence 358938, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
```

```
; SEQ ID NO 358938
;   LENGTH: 1414
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-358938

Alignment Scores:
Pred. No.:                169             Length:         1414
Score:                    61.00           Matches:        17
Percent Similarity:       51.0%           Conservative:   8
Best Local Similarity:    34.7%           Mismatches:     16
Query Match:              26.6%           Indels:         8
DB:                       9               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-358938 (1-1414)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||    |||:::|||   :::       |||||||     ||||||||||:::
Db       1407 CCCGGGTGTGGTATTTTGGAGTTAATGGAGTCAGAAGGACCTCCAGAGTCAGAGAGTTCA 1348

Qy         23 ------------------ThrAspGluAsnGluAspAspIleGlnPheValSerGlu--- 35
                               :::||||||||:::::::                  |||
Db       1347 GAATTTTTCTCACAGCAAGAAGAGGAAAATGAAGAAGAAGAAGCCCAAGAACCAGAGGAA 1288

Qy         36 ---GlyProSerArgProValLeuGlu 43
                 ||||||     |||:::|||:::
Db       1287 ACAGGCCCTAAGAACCCCCTTTTACAA 1261

RESULT 64
US-11-266-748A-442317
; Sequence 442317, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 442317
;   LENGTH: 1414
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-442317

Alignment Scores:
Pred. No.:                169             Length:         1414
Score:                    61.00           Matches:        17
Percent Similarity:       51.0%           Conservative:   8
Best Local Similarity:    34.7%           Mismatches:     16
Query Match:              26.6%           Indels:         8
DB:                       9               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-442317 (1-1414)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              ||||||    |||:::|||   :::       |||||||     ||||||||||:::
Db          8 CCCGGGTGTGGTATTTTGGAGTTAATGGAGTCAGAAGGACCTCCAGAGTCAGAGAGTTCA 67

Qy         23 ------------------ThrAspGluAsnGluAspAspIleGlnPheValSerGlu--- 35
                               :::||||||||:::::::                  |||
Db         68 GAATTTTTCTCACAGCAAGAAGAGGAAAATGAAGAAGAAGAAGCCCAAGAACCAGAGGAA 127

Qy         36 ---GlyProSerArgProValLeuGlu 43
                 ||||||     |||:::|||:::
Db        128 ACAGGCCCTAAGAACCCCCTTTTACAA 154

RESULT 65
US-11-266-748A-26928
; Sequence 26928, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26928
;   LENGTH: 1505
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-26928

Alignment Scores:
Pred. No.:                182             Length:         1505
Score:                    61.00           Matches:        16
Percent Similarity:       60.5%           Conservative:   7
Best Local Similarity:    42.1%           Mismatches:     13
Query Match:              26.6%           Indels:         2
DB:                       9               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-26928 (1-1505)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||       :::   |||   |||||||||      |||::::::::::|||
Db          9 CCTGGTTCCCACTGCCTAGTTTCTGGGCCCCCGGCATCC---GAGTCGGCCAAAAGCCGG 65

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 :::|||:::      |||   ||||||         |||   |||:::|||
Db         66 GCAGAAGAGAGCGCCAAGGACTGGCAGTTC---CGGAACGGGCAATCCCAGCCG 116

RESULT 66
US-11-266-748A-31320
; Sequence 31320, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
```

```
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31320
;   LENGTH: 1505
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-31320

Alignment Scores:
Pred. No.:              182          Length:          1505
Score:                  61.00        Matches:         16
Percent Similarity:     60.5%        Conservative:    7
Best Local Similarity:  42.1%        Mismatches:      13
Query Match:            26.6%        Indels:          2
DB:                     9            Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-31320 (1-1505)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||       :::   |||   |||||||||      |||::::::::::|||
Db          9 CCTGGTTCCCACTGCCTAGTTTCTGGGCCCCCGGCATCC---GAGTCGGCCAAAAGCCGG 65

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 :::|||:::      |||   ||||||         |||   |||:::|||
Db         66 GCAGAAGAGAGCGCCAAGGACTGGCAGTTC---CGGAACGGGCAATCCCAGCCG 116

RESULT 67
US-60-762-056-1344
; Sequence 1344, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 1344
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-1344

Alignment Scores:
Pred. No.:              307          Length:          2300
Score:                  61.00        Matches:         16
Percent Similarity:     60.5%        Conservative:    7
Best Local Similarity:  42.1%        Mismatches:      13
Query Match:            26.6%        Indels:          2
DB:                     11           Gaps:            2

US-10-757-745-6_COPY_2_45 (1-44) x US-60-762-056-1344 (1-2300)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||       :::   |||   |||||||||      |||::::::::::|||
Db       2109 CCTGGTTCCCACTGCCTGGTTTCTGGGCCCCCGGCATCC---GAGTCGGCCAAAAGCCGG 2165

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArgPro 40
                 :::|||:::      |||   ||||||         |||   |||:::|||
Db       2166 GCAGAAGAGAGCGCCAAGGACTGGCAGTTC---CGGAACGGGCAATCCCAGCCG 2216

RESULT 68
US-60-762-056-1879
; Sequence 1879, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 1879
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-1879

Alignment Scores:
Pred. No.:              307          Length:          2300
Score:                  61.00        Matches:         12
Percent Similarity:     55.6%        Conservative:    3
Best Local Similarity:  44.4%        Mismatches:      12
Query Match:            26.6%        Indels:          0
DB:                     11           Gaps:            0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-762-056-1879 (1-2300)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||   |||       :::         ||||||       ||||||   ||||||   :::
Db       1590 CCCCAGAGCCATGGGCTCGCTCACGCCCCTCCGCACTCCCCGGAGGGCTCCGAGGACAGC 1649

Qy         23 ThrAspGluAsnGluAspAsp 29
                 :::|||   |||||||||
Db       1650 GGAGAGGAGGAGGAAGACGAC 1670

RESULT 69
US-60-751-455-2721
; Sequence 2721, Application US/60751455
; GENERAL INFORMATION:
;  APPLICANT: Radich, Jerald P.
;  APPLICANT:  Dai, Hongyue
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Schelter, Janell
;  APPLICANT:  Linsley, Peter S.
;  TITLE OF INVENTION: GENES ASSOCIATED WITH PROGRESSION AND RESPONSE IN
;  TITLE OF INVENTION:  CHRONIC MYELOID LEUKEMIA AND USES THEREOF
;  FILE REFERENCE: 9301-253-888
;  CURRENT APPLICATION NUMBER: US/60/751,455
;  CURRENT FILING DATE:  2005-12-15
;  NUMBER OF SEQ ID NOS: 7936
;  SOFTWARE: FastSEQ 4.01
; SEQ ID NO 2721
;   LENGTH: 2502
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-751-455-2721
```

```
Alignment Scores:
Pred. No.:                341             Length:         2502
Score:                    61.00           Matches:        11
Percent Similarity:       62.1%           Conservative:   7
Best Local Similarity:    37.9%           Mismatches:     11
Query Match:              26.6%           Indels:         0
DB:                       11              Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-60-751-455-2721 (1-2502)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||          :::|||:::||||||:::::|||   ::::::    |||
Db        1108 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 1167

Qy          32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db        1168 GAGGATGATGAAGCTCCTTCTAAGCCA 1194

RESULT 70
US-10-531-147-1235
; Sequence 1235, Application US/10531147
; GENERAL INFORMATION:
;  APPLICANT: THOMAS WU
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF TUMOR
;  FILE REFERENCE: P5102R1-US
;  CURRENT APPLICATION NUMBER: US/10/531,147
;  CURRENT FILING DATE:  2005-04-12
;  PRIOR APPLICATION NUMBER: PCT/US03/029126
;  PRIOR FILING DATE: 2003-10-15
;  PRIOR APPLICATION NUMBER: US 60/418,988
;  PRIOR FILING DATE: 2002-10-18
;  NUMBER OF SEQ ID NOS: 4622
; SEQ ID NO 1235
;   LENGTH: 3775
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-531-147-1235

Alignment Scores:
Pred. No.:                566             Length:         3775
Score:                    61.00           Matches:        11
Percent Similarity:       62.1%           Conservative:   7
Best Local Similarity:    37.9%           Mismatches:     11
Query Match:              26.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-6_COPY_2_45 (1-44) x US-10-531-147-1235 (1-3775)

Qy          12 ProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAspAspIleGln 31
               ||||||          :::|||:::||||||:::::|||   ::::::    |||
Db        1128 CCACCTCCAGCAAAGAAAGCAGCAGAGAGCTCTTCAGACAGCTCAGACTCTGACAGCTCT 1187

Qy          32 PheValSerGluGlyProSerArgPro 40
                  |||   ||||||:::|||
Db        1188 GAGGATGATGAAGCTCCTTCTAAGCCA 1214

RESULT 71
US-11-266-748A-413970
; Sequence 413970, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 413970
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (413)..(413)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-413970

Alignment Scores:
Pred. No.:                46.9            Length:         439
Score:                    60.50           Matches:        18
Percent Similarity:       46.5%           Conservative:   2
Best Local Similarity:    41.9%           Mismatches:     16
Query Match:              26.4%           Indels:         7
DB:                       9               Gaps:           2

US-10-757-745-6_COPY_2_45 (1-44) x US-11-266-748A-413970 (1-439)

Qy           3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
               |||||||||          |||||||||      ||||||          |||
Db          75 CCCGGGTCTACACCGAAGGGTAGCGTGCCCGCGCGGGGGCCGGCGGGTGGGAGATCGCGC 134

Qy          23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArg------Pro 40
                  :::|||   |||   |||                  ||||||:::|||      |||
Db         135 CCCGAGGAGGGCGAGGGAGAC--------------CCAGGCCCCGCGCGACGGTCCCCC 179

Qy          41 ValLeuGlu 43
                  ||||||
Db         180 TCGCTGGAG 188

RESULT 72
US-10-461-673-5295
; Sequence 5295, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
```

```
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 5295
;   LENGTH: 1165
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1165)
;   OTHER INFORMATION: n = a,t,c or g
US-10-461-673-5295

Alignment Scores:
Pred. No.:                156             Length:         1165
Score:                    60.50           Matches:        13
Percent Similarity:       62.5%           Conservative:   7
Best Local Similarity:    40.6%           Mismatches:     11
Query Match:              26.4%           Indels:         1
DB:                       6               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-461-673-5295 (1-1165)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp 28
              |||:::   ||||||      |||:::|||   ::::::||||||         :::
Db        982 GAGACACCCCCTCCCCCGGCCCCGGACGCT---GACACCACCACCGCCCTCAGGCAGTCA 1038

Qy         29 AspIleGlnPheValSerGluGlyProSerArgPro 40
                 :::         :::||||||||||||||||||
Db       1039 CATCTCCCGAGCTCCGCCGAGGGACCCAGCCGCCCC 1074

RESULT 73
US-10-461-673-4943/c
; Sequence 4943, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 4943
;   LENGTH: 1538
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1259)..(1447)
US-10-461-673-4943

Alignment Scores:
Pred. No.:                220             Length:         1538
Score:                    60.50           Matches:        13
Percent Similarity:       62.5%           Conservative:   7
Best Local Similarity:    40.6%           Mismatches:     11
Query Match:              26.4%           Indels:         1
DB:                       6               Gaps:           1

US-10-757-745-6_COPY_2_45 (1-44) x US-10-461-673-4943 (1-1538)

Qy          9 GluSerValProProAlaGlyProGluAlaSerGluSerThrThrAspGluAsnGluAsp 28
              |||:::   ||||||      |||:::|||   ::::::||||||         :::
Db        192 GAGACACCCCCTCCCCCGGCCCCGGACGCT---GACACCACCACCGCCCTCAGGCAGTCA 136

Qy         29 AspIleGlnPheValSerGluGlyProSerArgPro 40
                 :::         :::||||||||||||||||||
Db        135 CATCTCCCGAGCTCCGCCGAGGGACCCAGCCGCCCC 100

RESULT 74
US-60-762-056-22047/c
; Sequence 22047, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
```

```
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 22047
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-22047

Alignment Scores:
Pred. No.:              361          Length:        2300
Score:                  60.50        Matches:       18
Percent Similarity:     46.5%        Conservative:  2
Best Local Similarity:  41.9%        Mismatches:    16
Query Match:            26.4%        Indels:        7
DB:                     11           Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x US-60-762-056-22047 (1-2300)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||          |||||||||      ||||||           |||
Db       1956 CCCGGGTCTACACCGAAGGGTAGCGTGCCCGCGCGGGGGCCGGCGGGTGGGAGATCGCGC 1897

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArg------Pro 40
                 :::|||   |||   |||                 ||||||:::|||      |||
Db       1896 CCCGAGGAGGGCGAGGGAGAC--------------CCAGGCCCCGCGCGACGGTCCCCC 1852

Qy         41 ValLeuGlu 43
                |||||
Db       1851 TCGCTGGAG 1843

RESULT 75
US-60-762-056-22048
; Sequence 22048, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 22048
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-22048

Alignment Scores:
Pred. No.:              361          Length:        2300
Score:                  60.50        Matches:       18
Percent Similarity:     46.5%        Conservative:  2
Best Local Similarity:  41.9%        Mismatches:    16
Query Match:            26.4%        Indels:        7
DB:                     11           Gaps:          2

US-10-757-745-6_COPY_2_45 (1-44) x US-60-762-056-22048 (1-2300)

Qy          3 ProGlySerGluIleIleGluSerValProProAlaGlyProGluAlaSerGluSerThr 22
              |||||||||          |||||||||      ||||||           |||
Db         73 CCCGGGTCTACACCGAAGGGTAGCGTGCCCGCGCGGGGGCCGGCGGGTGGGAGATCGCGC 132

Qy         23 ThrAspGluAsnGluAspAspIleGlnPheValSerGluGlyProSerArg------Pro 40
                 :::|||   |||   |||                 ||||||:::|||      |||
Db        133 CCCGAGGAGGGCGAGGGAGAC--------------CCAGGCCCCGCGCGACGGTCCCCC 177

Qy         41 ValLeuGlu 43
                |||||
Db        178 TCGCTGGAG 186
```

```
Search completed: March 11, 2006, 00:16:11
Job time : 106.946 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:22:00 ; Search time 846.962 Seconds
(without alignments)
1013.401 Million cell updates/sec

Title: US-10-757-745-2
Perfect score: 1902
Sequence: 1 MELGSCLEGGREAAEEEGEP.....DCGRFPSDHWGLLCNLDIIL 362

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5007642 seqs, 1185514400 residues

Total number of hits satisfying chosen parameters: 10015284

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170335_4211/app_query.fasta_1
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_240_@runat_09032006_170335_4211 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq1:*
9: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1902	100.0	1940	6	US-10-567-867-921	Sequence 921, App
2	1902	100.0	1940	11	US-60-742-219-2071	Sequence 2071, Ap
3	1893	99.5	1898	6	US-10-917-503B-13209	Sequence 13209, A
4	1616	85.0	1288	9	US-11-266-748A-226184	Sequence 226184,
5	1182	62.1	752	6	US-10-917-503B-4908	Sequence 4908, Ap
6	729.5	38.4	849	9	US-11-266-748A-166646	Sequence 166646,
7	696.5	36.6	1417	6	US-10-276-817B-5136	Sequence 5136, Ap
8	607	31.9	1179	9	US-11-266-748A-19660	Sequence 19660, A
9	482	25.3	1205	9	US-11-266-748A-184613	Sequence 184613,
10	482	25.3	1205	9	US-11-266-748A-192113	Sequence 192113,
c 11	387	20.3	2300	11	US-60-762-056-35503	Sequence 35503, A
12	305	16.0	1000	9	US-11-266-748A-117422	Sequence 117422,
c 13	305	16.0	1000	9	US-11-266-748A-159586	Sequence 159586,
14	305	16.0	1000	9	US-11-266-748A-289180	Sequence 289180,
c 15	305	16.0	1000	9	US-11-266-748A-340609	Sequence 340609,
16	288	15.1	2300	11	US-60-762-056-35504	Sequence 35504, A
17	231	12.1	1643	6	US-10-953-349-4575	Sequence 4575, Ap
18	158.5	8.3	525	9	US-11-292-078-14543	Sequence 14543, A
19	140	7.4	2928	6	US-10-540-898-785	Sequence 785, App
20	128	6.7	1767	6	US-10-998-197A-33	Sequence 33, Appl
21	128	6.7	1769	6	US-10-540-898-788	Sequence 788, App
22	126	6.6	47040	6	US-10-540-898-784	Sequence 784, App
23	123	6.5	49355	6	US-10-540-898-787	Sequence 787, App
24	122	6.4	1000	9	US-11-266-748A-283521	Sequence 283521,
c 25	122	6.4	1000	9	US-11-266-748A-310161	Sequence 310161,
c 26	118.5	6.2	6198	6	US-10-868-184C-10580	Sequence 10580, A
c 27	118.5	6.2	6198	6	US-10-868-184D-10580	Sequence 10580, A
c 28	113.5	6.0	6198	6	US-10-868-184C-10303	Sequence 10303, A
c 29	113.5	6.0	6198	6	US-10-868-184D-10303	Sequence 10303, A
c 30	113	5.9	6075	6	US-10-868-184C-10659	Sequence 10659, A
c 31	113	5.9	6075	6	US-10-868-184D-10659	Sequence 10659, A
32	112.5	5.9	1421	6	US-10-953-349-28325	Sequence 28325, A
33	110.5	5.8	5130	6	US-10-276-817B-96	Sequence 96, Appl
34	110.5	5.8	5130	6	US-10-276-817B-3643	Sequence 3643, Ap
35	110.5	5.8	8065	6	US-10-461-673-4287	Sequence 4287, Ap
36	110.5	5.8	10136	6	US-10-461-673-7266	Sequence 7266, Ap
c 37	110	5.8	6063	6	US-10-868-184C-10371	Sequence 10371, A
c 38	110	5.8	6063	6	US-10-868-184D-10371	Sequence 10371, A
c 39	110	5.8	6093	6	US-10-868-184C-10476	Sequence 10476, A
c 40	110	5.8	6093	6	US-10-868-184D-10476	Sequence 10476, A
c 41	109	5.7	5381	6	US-10-868-184C-10470	Sequence 10470, A
c 42	109	5.7	5381	6	US-10-868-184C-10513	Sequence 10513, A
c 43	109	5.7	5381	6	US-10-868-184D-10470	Sequence 10470, A
c 44	109	5.7	5381	6	US-10-868-184D-10513	Sequence 10513, A
45	108	5.7	1434	1	PCT-US05-40306-111	Sequence 111, App
46	108	5.7	1434	1	PCT-US05-40255-124	Sequence 124, App
47	108	5.7	1434	1	PCT-US05-40256-152	Sequence 152, App
48	108	5.7	1434	9	US-11-264-737-152	Sequence 152, App
49	108	5.7	1434	9	US-11-265-761-124	Sequence 124, App
50	108	5.7	1434	9	US-11-264-784-111	Sequence 111, App
c 51	108	5.7	6044	6	US-10-868-184C-10381	Sequence 10381, A
c 52	108	5.7	6044	6	US-10-868-184D-10381	Sequence 10381, A
53	108	5.7	8179	1	PCT-US05-40306-155	Sequence 155, App
54	108	5.7	8179	1	PCT-US05-40255-170	Sequence 170, App
55	108	5.7	8179	1	PCT-US05-40256-207	Sequence 207, App
56	108	5.7	8179	9	US-11-264-737-207	Sequence 207, App
57	108	5.7	8179	9	US-11-265-761-170	Sequence 170, App
58	108	5.7	8179	9	US-11-264-784-155	Sequence 155, App
c 59	108	5.7	14411	6	US-10-868-184C-10688	Sequence 10688, A
c 60	108	5.7	14411	6	US-10-868-184D-10688	Sequence 10688, A
c 61	107.5	5.7	2944528	9	US-11-045-004-1	Sequence 1, Appli
c 62	107	5.6	5691	6	US-10-868-184C-10429	Sequence 10429, A
c 63	107	5.6	5691	6	US-10-868-184D-10429	Sequence 10429, A
c 64	107	5.6	10789	6	US-10-868-184C-10456	Sequence 10456, A
c 65	107	5.6	10789	6	US-10-868-184D-10456	Sequence 10456, A
66	106.5	5.6	1207	9	US-11-266-748A-80889	Sequence 80889, A
c 67	106.5	5.6	1207	9	US-11-266-748A-133700	Sequence 133700,
68	106.5	5.6	3958	11	US-60-742-219-1685	Sequence 1685, Ap
69	106.5	5.6	6028	6	US-10-498-451-2320	Sequence 2320, Ap
c 70	106	5.6	2300	11	US-60-762-056-10029	Sequence 10029, A
c 71	106	5.6	5446	6	US-10-868-184C-10658	Sequence 10658, A
c 72	106	5.6	5446	6	US-10-868-184D-10658	Sequence 10658, A
c 73	106	5.6	5453	6	US-10-868-184C-10684	Sequence 10684, A
c 74	106	5.6	5453	6	US-10-868-184D-10684	Sequence 10684, A
c 75	106	5.6	5916	6	US-10-868-184C-10535	Sequence 10535, A

ALIGNMENTS

RESULT 1

```
US-10-567-867-921
; Sequence 921, Application US/10567867
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PCT
;  CURRENT APPLICATION NUMBER: US/10/567,867
;  CURRENT FILING DATE:  2006-01-27
;  PRIOR APPLICATION NUMBER: US 60/490,890
;  PRIOR FILING DATE: 2003-07-29
;  NUMBER OF SEQ ID NOS: 2786
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-567-867-921

Alignment Scores:
Pred. No.:              4.72e-217       Length:         1940
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2 (1-362) x US-10-567-867-921 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
              ||||||
Db       1106 ATATTG 1111

RESULT 2
US-60-742-219-2071
; Sequence 2071, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2071
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-2071

Alignment Scores:
Pred. No.:              4.72e-217       Length:         1940
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2 (1-362) x US-60-742-219-2071 (1-1940)
```

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
```

```
              ||||||
Db       1106 ATATTG 1111
```

RESULT 3
US-10-917-503B-13209
; Sequence 13209, Application US/10917503B
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/10/917,503B
; CURRENT FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US/09/629,469
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1086)
US-10-917-503B-13209

Alignment Scores:

Pred. No.:	5.48e-216	Length:	1898
Score:	1893.00	Matches:	361
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	99.5%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2 (1-362) x US-10-917-503B-13209 (1-1898)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220

Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240

Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260

Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280

Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340

Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360

Db       1021 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1080

Qy        361 IleLeu 362

Db       1081 ATATTG 1086

RESULT 4
US-11-266-748A-226184
; Sequence 226184, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 226184
;   LENGTH: 1288
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-226184

Alignment Scores:
Pred. No.:                4.96e-183       Length:          1288
Score:                    1616.00         Matches:         307
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    99.7%           Mismatches:      0
Query Match:              85.0%           Indels:          0
DB:                       9               Gaps:            0

US-10-757-745-2 (1-362) x US-11-266-748A-226184 (1-1288)

Qy         55 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 74

Db         15 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 74

Qy         75 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 94

Db         75 CCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 134

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114

Db        135 TCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTC 194

Qy        115 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 134

Db        195 TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 254

Qy        135 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 154

Db        255 GGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATT 314

Qy        155 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 174

Db        315 CCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCAT 374

Qy        175 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194

Db        375 GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAA 434

Qy        195 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 214

Db        435 GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAT 494

Qy        215 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 234

Db        495 GTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCT 554

Qy        235 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
```

```
Db        555 GCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCA 614

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 GCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGT 674

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAG 734

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 TATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGT 794

Qy        315 PheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeu 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 TTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 854

Qy        335 AspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHisTrpGlyLeu 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTT 914

Qy        355 LeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||||||
Db        915 CTGTGCAACTTAGATATAATATTG 938

RESULT 5
US-10-917-503B-4908
; Sequence 4908, Application US/10917503B
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4908
;   LENGTH: 752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (685)..(685)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (694)..(694)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (740)..(740)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-10-917-503B-4908

Alignment Scores:
Pred. No.:                2.4e-131      Length:          752
Score:                    1182.00       Matches:         238
Percent Similarity:       94.8%         Conservative:    1
Best Local Similarity:    94.4%         Mismatches:      10
Query Match:              62.1%         Indels:          3
DB:                       6             Gaps:            0

US-10-757-745-2 (1-362) x US-10-917-503B-4908 (1-752)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArg-AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 660

Qy        220 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 240
              |||||||   ||||||  |||    :::|||   ||||||        |||||||||||||
Db        661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy        240 lnLeuLysMetValLeu-LysLysMetGln 249
              ||   ||||||  |||   |||||||||
Db        721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750

RESULT 6
US-11-266-748A-166646
; Sequence 166646, Application US/11266748A
```

```
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 166646
;   LENGTH: 849
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-166646

Alignment Scores:
Pred. No.:              5.43e-77      Length:          849
Score:                  729.50        Matches:         148
Percent Similarity:     91.0%         Conservative:    3
Best Local Similarity:  89.2%         Mismatches:      10
Query Match:            38.4%         Indels:          5
DB:                     9             Gaps:            1

US-10-757-745-2 (1-362) x US-11-266-748A-166646 (1-849)

Qy        202 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal---SerGlyAsnGluLeu 220
              ||||||   :::    |||  |||||||||     :::      |||||||||
Db          8 ACCAAAAAGATAGAGGAACTTGGATGTGTGCATGAGCAACTGTGTCAGGACAATGAGCTA 67

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 127

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 187

Qy        261 GlyAspThrAsnLeuArgAspArg-GluValThr-ArgCysGlyGlyLeuProAsnAsnI 280
              ||||||||||||||||||||||| :::||||||   ||||||||||||||||||||||
Db        188 GGAGATACAAATCTAAGGGATCGAGCAGGTTACCAGCATGTGGTGGTTTACCCAACAACA 247

Qy        280 leValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp-AspThr 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        248 TTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGCATACA 307

Qy        300 GlnMet-AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePh 319
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CAAATGCAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATT 367

Qy        319 ePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLe 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 TTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATT 427

Qy        339 uGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAs 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        428 AGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGA 487

Qy        359 pIleIleLeu 362
              |||||||||
Db        488 TATAATATTG 497

RESULT 7
US-10-276-817B-5136
; Sequence 5136, Application US/10276817B
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-104
;  CURRENT APPLICATION NUMBER: US/10/276,817B
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: PCT/US01/14827
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 16102
;  SOFTWARE: Custom
; SEQ ID NO 5136
;   LENGTH: 1417
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (48)..(422)
;   OTHER INFORMATION: 79% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=494.
US-10-276-817B-5136

Alignment Scores:
Pred. No.:              1.11e-72      Length:          1417
Score:                  696.50        Matches:         145
Percent Similarity:     92.4%         Conservative:    1
Best Local Similarity:  91.8%         Mismatches:      6
Query Match:            36.6%         Indels:          6
DB:                     6             Gaps:            1

US-10-757-745-2 (1-362) x US-10-276-817B-5136 (1-1417)

Qy        211 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 230
              :::  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 CTGCAGGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACC 107

Qy        231 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 AGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAG 167

Qy        251 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluVal 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTT 227

Qy        271 ThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 ACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCT 287

Qy        291 LysHisCysGln-TyrThr-TrpAspThrGlnMetAsn---SerAsnLeuGlyIleThrA 309
              |||||||||  |||||| ||||||||||||           ||||||||||||||
Db        288 AAACATTGCCACGTATACATTGGGATACACAAACTGAACTCCTAATCTTGGGAATAACTG 347

Qy        309 la-AlaCysLysLeuArgPheAspArgIlePhePheArg-AlaAlaAlaGluGluGlyHi 328
              || ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        348 CCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAAGCAGCAGCAGAAGAGGGACA 407

Qy        328 sIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePr 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 CATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCC 467

Qy        348 oSerAspHisTrpGlyLeuLeuCysAsnLeu-AspIleIleLeu 362
              |||||||||||||||||||||||||||||| ||||||||||||
```

```
Db        468 TAGTGATCACTGGGGTCTTCTGTGCAACTTAGGATATAATATTG 511

RESULT 8
US-11-266-748A-19660
; Sequence 19660, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19660
;   LENGTH: 1179
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-19660

Alignment Scores:
Pred. No.:              4.49e-62        Length:         1179
Score:                  607.00          Matches:        111
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            31.9%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2 (1-362) x US-11-266-748A-19660 (1-1179)

Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACC 60

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAA 120

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGT 180

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 240

Qy        332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHis 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCAC 300

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||||||||||||||||||||||||
Db        301 TGGGGTCTTCTGTGCAACTTAGATATAATATTG 333

RESULT 9
US-11-266-748A-184613
; Sequence 184613, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 184613
;   LENGTH: 1205
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (64)..(81)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-184613

Alignment Scores:
Pred. No.:              4.56e-47        Length:         1205
Score:                  482.00          Matches:        90
Percent Similarity:     98.9%           Conservative:   3
Best Local Similarity:  95.7%           Mismatches:     1
Query Match:            25.3%           Indels:         1
DB:                     9               Gaps:           0

US-10-757-745-2 (1-362) x US-11-266-748A-184613 (1-1205)

Qy        269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 193

Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 253

Qy        309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328
              ||||||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db        254 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAAAGCAGCAGCAGAAGAGGGACAC 313

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348
              ||||||||||||||||   ||||||||||:::||||||||||||||||||||||||||||
Db        314 ATTATTCCCCGAAGTTTGGC-CTTCTTGGATTAAAAAAACTGGACTGTGGTAGATTTCCT 372

Qy        349 SerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||||||||||||||||||||||||||||||||
Db        373 AGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 414

RESULT 10
US-11-266-748A-192113
; Sequence 192113, Application US/11266748A
; GENERAL INFORMATION:
```

```
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 192113
;   LENGTH: 1205
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (64)..(81)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-192113

Alignment Scores:
Pred. No.:                4.56e-47        Length:          1205
Score:                    482.00          Matches:         90
Percent Similarity:       98.9%           Conservative:    3
Best Local Similarity:    95.7%           Mismatches:      1
Query Match:              25.3%           Indels:          1
DB:                       9               Gaps:            0

US-10-757-745-2 (1-362) x US-11-266-748A-192113 (1-1205)

Qy        269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 193

Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 253

Qy        309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328
              ||||||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db        254 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAAAGCAGCAGCAGAAGAGGGACAC 313

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348
              ||||||||||||||||   ||||||||||||:::||||||||||||||||||||||||||
Db        314 ATTATTCCCCGAAGTTTGGC-CTTCTTGGATTAAAAAAACTGGACTGTGGTAGATTTCCT 372

Qy        349 SerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||||||||||||||||||||||||
Db        373 AGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 414

RESULT 11
US-60-762-056-35503/c
; Sequence 35503, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 35503
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-35503

Alignment Scores:
Pred. No.:                3.14e-35        Length:          2300
Score:                    387.00          Matches:         82
Percent Similarity:       73.2%           Conservative:    0
Best Local Similarity:    73.2%           Mismatches:      1
Query Match:              20.3%           Indels:          30
DB:                       11              Gaps:            1

US-10-757-745-2 (1-362) x US-60-762-056-35503 (1-2300)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1916 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 1857

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1856 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 1797

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu---------------- 55
              ||||||||||||||||||||||||||||||||||||||||||||
Db       1796 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGA-AGTACGCTTTCGCTG 1738

Qy         55 ------------------------------------------------------------ 55

Db       1737 TCCTCGCTTCCGGCTCCATTACCTAGGTCGGTGGAGCGCACAAACCTCATCCCTTGTGCC 1678

Qy         56 ------------ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
                          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1677 TTTGTTTTTCAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTG 1618

Qy         72 GluArgArgProGluThrIleSerGluProLysThr 83
              ||||||||||||||||||||||||||||||||||||
Db       1617 GAACGCCGACCTGAAACCATCTCTGAGCCCAAGACC 1582

RESULT 12
US-11-266-748A-117422
; Sequence 117422, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
```

```
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 117422
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-117422

Alignment Scores:
Pred. No.:              5.76e-26        Length:         1000
Score:                  305.00          Matches:        57
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            16.0%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2 (1-362) x US-11-266-748A-117422 (1-1000)

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 325
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 60

Qy        326 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGly 345
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 120

Qy        346 ArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 171

RESULT 13
US-11-266-748A-159586/c
; Sequence 159586, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 159586
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-159586

Alignment Scores:
Pred. No.:              5.76e-26        Length:         1000
Score:                  305.00          Matches:        57
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            16.0%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2 (1-362) x US-11-266-748A-159586 (1-1000)

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 325
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 941

Qy        326 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGly 345
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 881

Qy        346 ArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 AGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 830

RESULT 14
US-11-266-748A-289180
; Sequence 289180, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 289180
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-289180

Alignment Scores:
Pred. No.:              5.76e-26        Length:         1000
Score:                  305.00          Matches:        57
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            16.0%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2 (1-362) x US-11-266-748A-289180 (1-1000)

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 325
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 60

Qy        326 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGly 345
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 120

Qy        346 ArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 AGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 171

RESULT 15
US-11-266-748A-340609/c
; Sequence 340609, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 340609
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-340609

Alignment Scores:
Pred. No.:              5.76e-26        Length:         1000
Score:                  305.00          Matches:        57
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            16.0%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2 (1-362) x US-11-266-748A-340609 (1-1000)

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 941

Qy        326 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGly 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 881

Qy        346 ArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 AGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTG 830

RESULT 16
US-60-762-056-35504
; Sequence 35504, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 35504
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-35504

Alignment Scores:
Pred. No.:              2.35e-23        Length:         2300
Score:                  288.00          Matches:        55
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            15.1%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2 (1-362) x US-60-762-056-35504 (1-2300)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2126 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 2185

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2186 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 2245

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||
Db       2246 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAA 2290

RESULT 17
US-10-953-349-4575
; Sequence 4575, Application US/10953349
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCONDED THERBY
;  FILE REFERENCE: 2750-1579PUS2
;  CURRENT APPLICATION NUMBER: US/10/953,349
;  CURRENT FILING DATE:  2004-09-30
;  NUMBER OF SEQ ID NOS: 40252
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4575
;   LENGTH: 1643
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-10-953-349-4575

Alignment Scores:
Pred. No.:              9.47e-17        Length:         1643
Score:                  231.00          Matches:        84
Percent Similarity:     44.2%           Conservative:   39
Best Local Similarity:  30.2%           Mismatches:     109
Query Match:            12.1%           Indels:         46
DB:                     6               Gaps:           11

US-10-757-745-2 (1-362) x US-10-953-349-4575 (1-1643)

Qy        113 MetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 132
              :::   |||          |||      :::   |||:::||||||           |||
Db        550 ATCTTGAGCTACAATGTTTGGTTTCGTGAAGACCTTGAGTTGAAT---------CTAAGG 600

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              |||  :::         :::   |||:::||||||   :::|||      ||||||
Db        601 ATGAGAGCGATTGGTCATCTTATTCAGCTTCATTCTCCTCATCTCATATGTTTTCAGGAA 660

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArg-----------------SerSer 166
              |||  |||      |||         :::|||                     |||
Db        661 GTCACTCCTGAGATATATGATATTTTCCGCAAGTCGAATTGGTGGAAAGCGTATTCTTGT 720

Qy        167 AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 186
              :::      :::                 ||||||:::      ::::::
Db        721 TCTGTCTCAGTTGATGTGGCAGTTTCGAGAGGCTACTACTGCATGCTGTTG--------- 771
```

```
Qy        187 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 206
              |||::::::   :::||||||                  |||   ::::::   |||   |||
Db        772 AGCAAGTTGGGAGTGAAATCTTTCAGCAGCAAAAGCTTTGGGAACTCGATAATGGGAAGA 831

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGlu---LeuCysLeuMetThrSer 225
                 |||             |||   |||   |||   :::   |||             ||||||
Db        832 GAACTATCTATCGCAGAAGTTGAAGTTCCAGGGAGAAAGCCACTTGTTTTTGCAACTAGC 891

Qy        226 HisLeuGluSer----------------------ThrArgGlyHisAlaAlaGluArg 237
              ||||||||||||                    :::       :::   ||||||
Db        892 CATCTTGAGAGTCCCTGTCCAGGTCCTCCAAAATGGGACCAAATGTTCAGTCGAGAACGC 951

Qy        238 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 257
              :::   |||   |||      ::::::  ::::::           :::|||   |||
Db        952 GTTGAACAAGCCAAGGAGGCAATTGAGATTCTCAGA---------CCTAATGCCAATGTG 1002

Qy        258 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly------ 275
              ||||||  ||||||  |||     |||           :::     |||
Db       1003 ATATTCGGGGGAGACATGAACTGGTGCGAC-----------AAGCTAGATGGGAAGTTT 1050

Qy        276 ---LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 294
                 ||||||:::     |||||||||||||||   |||   ||||||
Db       1051 CCTTTGCCAGACAAATGGGTGGATGTTTGGGAGGTTTTG---AAACCAGGTGATTTGGGT 1107

Qy        295 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 314
              :::|||:::||||||:::   |||       |||             |||   :::   |||
Db       1108 TTTACATATGACACAAAAGCAAACCCAATGTTGTCCGGTAACCGGGCATTGCAGAAACGG 1167

Qy        315 PheAspArgIlePhePheArg---------------------------------------- 321
                 ||||||||   |||
Db       1168 TTAGACCGGATTTTGTGTCGTCTGGATGATTACAAGCTTGGTGGAATCGAGATGGTTGGG 1227

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGly---LeuGlu 340
                 |||     |||           :::       :::   |||   ::::::
Db       1228 AAAGAGGCCATTCCTGGTTTATCATATGTAAAAGAGAAGAAAGTGAGGGGTGACATTAAG 1287

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeu 358
              ||||||:::           ||||||||||||:::||||||||      |||
Db       1288 AAGTTAGAGCTTCCGGTGTTGCCTAGTGATCATTTCGGTCTGCTTGTAACCCTT 1341
```

RESULT 18
US-11-292-078-14543
; Sequence 14543, Application US/11292078
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.346 - 38-21(52274)C
; CURRENT APPLICATION NUMBER: US/11/292,078
; CURRENT FILING DATE: 2005-12-02
; PRIOR APPLICATION NUMBER: US 10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255,619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 14543
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-034-Q1-K6-G12
US-11-292-078-14543

Alignment Scores:

Pred. No.:	7.76e-09	Length:	525
Score:	158.50	Matches:	53
Percent Similarity:	46.1%	Conservative:	29
Best Local Similarity:	29.8%	Mismatches:	63
Query Match:	8.3%	Indels:	33
DB:	9	Gaps:	7

US-10-757-745-2 (1-362) x US-11-292-078-14543 (1-525)

```
Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   :::   :::       :::   |||:::|||||||||:::|||
Db         19 AAGAGGATGAAAGCTATTGGTGACCTTATACAACTTCATTCCCCTGATATCATTTGTTTT 78

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              |||||||||   |||       |||                ||   ||||||:::
Db         79 CAAGAGGTTACTCCGATGATTTATGACATTTTT------CGTGGGTCTAATTGG------ 126

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIle---------------------- 182
                          :::|||||||      :::|||
Db        127 ------------TGGAAAGGGTACCGGTGCTCTATCTCGGACAACACGGCGAGTTTGAGA 174

Qy        183 ---------MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 199
                      ::::::  |||::::::  :::||||||                ||||||
Db        175 GCTTACTTTTGTATACAGTTAAGCAAATTGCCTGTGAAATCCTTTCGTCGTGAACCTTTT 234

Qy        200 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 219
              ::::::   |||   |||   |||             |||   |||||||||   |||
Db        235 GATAATTCGATAATGGGGCGAGAACTTTGCATGACTGAGGTTGAAGTTTCTGGGGGTGAA 294

Qy        220 ------LeuCysLeuMetThrSerHisLeuGluSer----------------------- 229
                    |||   :::   |||:::||||||||||||
Db        295 ACAACGCTTGTTGTTGCTACTACCCATCTCGAGAGTCCGTGTCCTGCTCCGCCGAAATGG 354

Qy        230 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 249
              :::   ::::::   ||||||:::   |||   |||:::   :::      ::::::
Db        355 GATCAAATGTATAGTAAGGAACGTGTAGAACAAGCGAAAGTAGCGATCAATATTCTCGAG 414

Qy        250 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
                 |||           |||||||||   ||||||   |||      |||:::
Db        415 AATAACCCG---------AATGTTATTTTTGGGGGTGATATGAATTGGGATGACAAATTA 465

Qy        270 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
                 |||         |||  :::      :::|||   |||         |||
Db        466 GATGGTCGATTT---CCTTTAACTGATGGATGGATTGATGCTTGGAACGAGTTA 516
```

RESULT 19
US-10-540-898-785
; Sequence 785, Application US/10540898
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)
; CURRENT APPLICATION NUMBER: US/10/540,898
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: US 10/330,773
; PRIOR FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785
; LENGTH: 2928
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2928)
; OTHER INFORMATION: n = A,T,C or G
US-10-540-898-785

Alignment Scores:

Pred. No.:	1.93e-05	Length:	2928
Score:	140.00	Matches:	89
Percent Similarity:	35.0%	Conservative:	55
Best Local Similarity:	21.7%	Mismatches:	163
Query Match:	7.4%	Indels:	104
DB:	6	Gaps:	19

US-10-757-745-2 (1-362) x US-10-540-898-785 (1-2928)

```
Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArg------------------ 25
              :::||||||    |||         :::   |||
Db        666 CAGGCGGCGGCGGCCTCGGGCGCCGCGAGGTCGCGTCCCCGAACAGTGAGTTCCATGGGA 725

Qy         26 ------------ArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
                        ||||||                ::|||   |||      ||||||
Db        726 AACGGCACCAGTCGACTCTACAGTGCTCTCGCCAAGACCGTCAACAGCAGTGCAGCTGCT 785

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
                     :::|||          |||   |||                    :::
Db        786 CAGCACCCGGAGTACTTGGTGTCAACTGACCCCGAA--------------------CAT 824

Qy         62 PheGluPro------ProValGlu---------------GluSerAlaLeuGluArgArg 74
              ||||||      |||   |||               :::   |||      |||
Db        825 CTGGAGCCCATCGATCCTAAGGAACTTCTGGAGGAATGCAGGGCTGTTCTGCACACTCGG 884

Qy         75 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 94
              |||                  :::   :::|||||||||                |||||
Db        885 CCACCCCGCTACCAG------CGGGATTTTGTGGACCTG------------AGGACAGAT 926

Qy         95 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 114
                 ::::::|||                                          |||
Db        927 TGCTCCAGCAGCCAC----------------------------------TCTCCCATT 950

Qy        115 SerLeuIleThrTrpAsnIle-----------------AspGlyLeuAsp--------- 125
                 ::::::   |||||||||               :::|||   |||
Db        951 CGCGTCATGCAGTGGAACATCCTCGCCCAAGCTCTCGGAGAAGGCAAAGACAACTTTGTG 1010

Qy        126 ---------------LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
                                      ||||||        :::         :::
Db       1011 CAGTGCCCCGTGGAAGCGCTCAAATGGGAAGAGAGGAAGTGCCTGATCCTGGAGGAGATC 1070

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
               |||   ||||||::::::   ||||||||||||         |||:::
Db       1071 CTGGCTTACCAGCCAGACATACTGTGCCTCCAGGAAGTG---GACCACTACTTTGACACC 1127

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe--- 179
                 :::    |||              |||::::::    :::|||
Db       1128 TTCCAGCCACTCCTCAGTAGA------------CTGGGCTACCAAGGCACGTTTTTCCCC 1175

Qy        180 ------------------------------------------------ThrAlaIle 182
                                                             |||:::
Db       1176 AAGCCCTGGTCACCATGTCTAGATGTGGAACACAACAACGGTCCAGATGGCTGTGCCTTA 1235

Qy        183 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 202
                  ::::::|||   ||||||   |||         |||         :::
Db       1236 TTTTTTCTCCAAAACCGATTCAAGCTTATCAGCAGCACCAATATTAGGCTGACAGCCATG 1295

Qy        203 LysMetMetArgAsnLeuLeuCysVal------------HisValAsnValSerGlyAsn 218
                 :::   |||  :::   :::                    |||||
Db       1296 ACCCTAAAAACCAATCAGGTGGCCATCGCACAGACCCTGGAGTGCAAGGAGTCTGGGCGA 1355

Qy        219 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 238
              :::   |||:::      :::||||||::::::      |||              |||
Db       1356 CAGTTCTGCATTGCTGTCACCCACTTAAAAGCCCGCACTGGCTGGGAGCGGTTCCGGTCA 1415

Qy        239 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal--- 257
                 |||         :::|||:::   :::|||         :::   |||   :::
Db       1416 GCTCAGGGCTGTGACCTCCTCCAGAACCTGCAGAACATCACCCAGGGAGCAAAGATCCCC 1475

Qy        258 ---IlePheAlaGlyAspThrAsnLeuArgAspArgGlu------ValThrArgCysGly 274
                 |||     ||||||  |||          |||
Db       1476 CTGATCGTCTGCGGGGACTTCAACGCAGAGCCAACCGAAGAGGTCTACAAACACTTTGCG 1535

Qy        275 GlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------GlyLysProLys 291
                    |||:::         ::::::   |||          |||:::   :::
Db       1536 TCCTCCAGCCTCAACCTCAACAGCGCCTACAAGCTGCTGAGTCCCGATGGACAGTCGGAG 1595

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
                    ||||||                  :::      :::      |||
Db       1596 CCTCCCTACACAACCTGG-----------------AAGATCCGGACCTCAGGCGAGTGT 1637

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
```

```
                 :::         |||   |||::::::         |||                :::
Db       1638 CGCCACACGCTGGACTATATCTGGTACTCCAGACATGCTCTGAGT---------GTCACG 1688

Qy        332 ArgSerLeuAspLeuLeu---------------GlyLeuGluLysLeuAspCysGlyArg 346
                  :::||||||||||||               |||      :::|||
Db       1689 TCCGCCCTGGATCTACTCACTGAAGAACAGATTGGGCCCAACCGGCTCCCATCCTTCCAT 1748

Qy        347 PheProSerAspHisTrpGlyLeuLeuCysAsn 357
                  :::||||||||||||      |||:::|||:::
Db       1749 TACCCCTCGGACCACCTGTCCCTGGTGTGTGAC 1781

RESULT 20
US-10-998-197A-33
; Sequence 33, Application US/10998197A
; GENERAL INFORMATION:
;  APPLICANT: Dohrmann, Cord
;  APPLICANT:  Austen, Matthias
;  TITLE OF INVENTION: Pancreas-Specific Proteins
;  FILE REFERENCE: 2923-669
;  CURRENT APPLICATION NUMBER: US/10/998,197A
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: PCT/EP03/05700
;  PRIOR FILING DATE: 2003-05-30
;  PRIOR APPLICATION NUMBER: EP 02020829.4
;  PRIOR FILING DATE: 2002-09-17
;  PRIOR APPLICATION NUMBER: EP 02011963.2
;  PRIOR FILING DATE: 2002-05-29
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
;   LENGTH: 1767
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-998-197A-33

Alignment Scores:
Pred. No.:              0.00024       Length:          1767
Score:                  128.00        Matches:         83
Percent Similarity:     33.0%         Conservative:    55
Best Local Similarity:  19.9%         Mismatches:      164
Query Match:            6.7%          Indels:          116
DB:                     6             Gaps:            18

US-10-757-745-2 (1-362) x US-10-998-197A-33 (1-1767)

Qy         20 ProGluValLysLysArgArgLeuLeuCysValGluPheAla---------SerValAla 36
              |||                 |||||||||             |||         :::
Db        308 CCCAGGCCCGCATCCCCCCGGCTGCTGGCGGCGGCCTCGGCGGCCTCGGGCGCCGCGAGG 367

Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
              ||||||         |||       |||   :::                 :::
Db        368 TCGTGTTCCCGAACAGTG------TGTTCCATGGGAACCGGTACAAGCAGACTCTATAGT 421

Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
              ||||||                  :::     |||||     :::     ||||||
Db        422 GCTCTC------------GCCAAGACACTGAACAGCAGCGCTGCCTCCCAGCACCCAGAG 469

Qy         77 ThrIleSerGluPro---------------------------------------------- 81
                 :::       |||
Db        470 TATTTGGTGTCACCTGACCCAGAGCATCTGGAGCCCATTGATCCTAAAGAGCTTCTTGAG 529

Qy         82 ------------------------------------------LysThrTyrValAspLeu 87
                                                       :::   :::||||||||||
Db        530 GAATGCAGGGCCGTCCTGCACACCCGACCTCCCCGGTTCCAGAGGGATTTTGTGGATCTG 589

Qy         88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 107
                    :::            ||||||          |||
Db        590 AGGACAGAT------TGCCCTAGTACCCACCCACCTATC---------------------- 622

Qy        108 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle------------------ 121
                                     ::::::   |||||||||
Db        623 --------------------AGGGTTATGCAATGGAACATCCTCGCCCAAGCTCTTGGA 661
```

```
Qy        122 AspGlyLeuAsp----------------------LeuAsnAsnLeuSerGluArgAla 133
              ::: |||  |||                      ||||||
Db        662 GAAGGCAAAGACAACTTTGTACAGTGCCCTGTTGAAGCACTCAAATGGGAAGAAAGGAAA 721

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
                  :::      :::   |||  ||||||::::::   ||||||||||||
Db        722 TGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTCCAAGAGGTG 781

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
                  |||:::      :::          |||                      |||
Db        782 ---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA------------CTAGGC 826

Qy        174 HisGluGluGlyTyrPhe----------------------------------------- 179
              ::::::      :::|||
Db        827 TATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAACACAACAAT 886

Qy        180 ------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
                       |||:::      :::::::|||   ||||||   :::
Db        887 GGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTCAACAGTGCC 946

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal------------ 211
              |||          :::      :::       |||   :::   :::
Db        947 AATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCACAGACCCTG 1006

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
                    ||||||   :::   |||:::      :::||||||::::::
Db       1007 GAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAAGCACGCACT 1066

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
              |||       |||  :::      |||      :::|||:::   :::|||
Db       1067 GGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTGCAAAACATC 1126

Qy        252 ProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
                 :::  |||  :::      |||      ||||||  |||             |||
Db       1127 ACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAGCCAACAGAA 1186

Qy        270 ------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
                         |||:::          ::::::   |||
Db       1187 GAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTACAAGCTGCTG 1246

Qy        288 ---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304
                    |||:::   :::          ||||||
Db       1247 AGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG-----------------AAG 1288

Qy        305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 324
              :::      :::      |||:::      |||  |||::::::          |||
Db       1289 ATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCTAAACATGCT 1348

Qy        325 GluGluGlyHisIleIleProArgSerLeuAspLeuLeu--------------GlyLeu 339
                       :::      :::||||||||||||                  |||
Db       1349 CTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAGATTGGACCC 1399

Qy        340 GluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsn 357
                ::|||          :::||||||||||||      |||:::|||:::
Db       1400 AACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGTGAC 1453

RESULT 21
US-10-540-898-788
; Sequence 788, Application US/10540898
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: CHIR0056-101 (PP023367.0003)
;  CURRENT APPLICATION NUMBER: US/10/540,898
;  CURRENT FILING DATE:  2005-06-27
;  PRIOR APPLICATION NUMBER: US 10/330,773
;  PRIOR FILING DATE: 2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 788
;   LENGTH: 1769
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-540-898-788

Alignment Scores:
Pred. No.:                0.00024       Length:          1769
Score:                    128.00        Matches:         83
Percent Similarity:       33.0%         Conservative:    55
Best Local Similarity:    19.9%         Mismatches:      164
Query Match:              6.7%          Indels:          116
DB:                       6             Gaps:            18

US-10-757-745-2 (1-362) x US-10-540-898-788 (1-1769)

Qy         20 ProGluValLysLysArgArgLeuLeuCysValGluPheAla---------SerValAla 36
              |||                 |||||||||             |||         :::
Db        308 CCCAGGCCCGCATCCCCCCGGCTGCTGGCGGCGGCCTCGGCGGCCTCGGGCGCCGCGAGG 367

Qy         37 SerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArg 56
              ||||||            |||      |||   :::                  :::
Db        368 TCGTGTTCCCGAACAGTG------TGTTCCATGGGAACCGGTACAAGCAGACTCTATAGT 421

Qy         57 AlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGlu 76
              ||||||                  :::          ||||||      :::   ||||||
Db        422 GCTCTC------------GCCAAGACACTGAACAGCAGCGCTGCCTCCCAGCACCCAGAG 469

Qy         77 ThrIleSerGluPro---------------------------------------------- 81
                 :::      |||
Db        470 TATTTGGTGTCACCTGACCCAGAGCATCTGGAGCCCATTGATCCTAAAGAGCTTCTTGAG 529

Qy         82 ---------------------------------------------LysThrTyrValAspLeu 87
                                                          :::   :::||||||||||
Db        530 GAATGCAGGGCCGTCCTGCACACCCGACCTCCCCGGTTCCAGAGGGATTTTGTGGATCTG 589

Qy         88 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 107
                    :::          ||||||      |||
Db        590 AGGACAGAT------TGCCCTAGTACCCACCCACCTATC---------------------- 622

Qy        108 GlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle------------------ 121
                                      ::::::   |||||||||
Db        623 --------------------AGGGTTATGCAATGGAACATCCTCGCCCAAGCTCTTGGA 661

Qy        122 AspGlyLeuAsp----------------------LeuAsnAsnLeuSerGluArgAla 133
              :::|||  |||                      ||||||
Db        662 GAAGGCAAAGACAACTTTGTACAGTGCCCTGTTGAAGCACTCAAATGGGAAGAAAGGAAA 721

Qy        134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
                  :::      :::   |||  ||||||::::::   ||||||||||||
Db        722 TGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTCCAAGAGGTG 781

Qy        154 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 173
                  |||:::      :::          |||                      |||
Db        782 ---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA------------CTAGGC 826

Qy        174 HisGluGluGlyTyrPhe----------------------------------------- 179
              ::::::      :::|||
Db        827 TATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAACACAACAAT 886

Qy        180 ------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
                       |||:::      :::::::|||   ||||||   :::
Db        887 GGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTCAACAGTGCC 946

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal------------ 211
              |||          :::      :::       |||   :::   :::
Db        947 AATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCACAGACCCTG 1006

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
                    ||||||   :::   |||:::      :::||||||::::::
Db       1007 GAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAAGCACGCACT 1066

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
              |||       |||  :::      |||      :::|||:::   :::|||
Db       1067 GGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTGCAAAACATC 1126
```

```
Qy        252 ProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArgAspArgGlu 269
Db       1127 ACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAGCCAACAGAA 1186

Qy        270 ------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 287
Db       1187 GAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTACAAGCTGCTG 1246

Qy        288 ---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 304
Db       1247 AGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG-----------------AAG 1288

Qy        305 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 324
Db       1289 ATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCTAAACATGCT 1348

Qy        325 GluGluGlyHisIleIleProArgSerLeuAspLeuLeu--------------GlyLeu 339
Db       1349 CTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAGATTGGACCC 1399

Qy        340 GluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsn 357
Db       1400 AACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGTGAC 1453
```

RESULT 22
US-10-540-898-784
; Sequence 784, Application US/10540898
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)
; CURRENT APPLICATION NUMBER: US/10/540,898
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: US 10/330,773
; PRIOR FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 784
; LENGTH: 47040
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(47040)
; OTHER INFORMATION: n = A,T,C or G
US-10-540-898-784

Alignment Scores:

Pred. No.:	0.0739	Length:	47040
Score:	126.00	Matches:	57
Percent Similarity:	37.9%	Conservative:	42
Best Local Similarity:	21.8%	Mismatches:	114
Query Match:	6.6%	Indels:	48
DB:	6	Gaps:	10

US-10-757-745-2 (1-362) x US-10-540-898-784 (1-47040)

```
Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db      35153 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 35212

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db      35213 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 35260

Qy        171 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 179
Db      35261 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 35317

Qy        180 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
Db      35318 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 35377

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 211
Db      35378 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 35437

Qy        212 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
Db      35438 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 35497

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
Db      35498 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 35557

Qy        249 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 266
Db      35558 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 35617

Qy        267 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
Db      35618 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACGCGCCTAC 35677

Qy        285 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
Db      35678 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 35725

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321
Db      35726 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 35779

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu------------ 337
Db      35780 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 35830

Qy        338 ---GlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356
Db      35831 ATTGGGCCCAACCGGCTCCCATCCTTCCATTACCCCTCGGACCACCTGTCCCTGGTGTGT 35890

Qy        357 Asn 357
Db      35891 GAC 35893
```

RESULT 23
US-10-540-898-787
; Sequence 787, Application US/10540898
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)
; CURRENT APPLICATION NUMBER: US/10/540,898
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: US 10/330,773
; PRIOR FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 787
; LENGTH: 49355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(49355)
; OTHER INFORMATION: n = A,T,C or G
US-10-540-898-787

Alignment Scores:

Pred. No.:	0.183	Length:	49355
Score:	123.00	Matches:	56
Percent Similarity:	37.9%	Conservative:	43
Best Local Similarity:	21.5%	Mismatches:	114
Query Match:	6.5%	Indels:	48
DB:	6	Gaps:	10

```
US-10-757-745-2 (1-362) x US-10-540-898-787 (1-49355)

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ||||||          :::         :::      |||   ||||||::::::   |||
Db      38299 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 38358

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              |||||||||         |||:::         :::         |||
Db      38359 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 38406

Qy        171 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 179
                    |||::::::      :::|||
Db      38407 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 38463

Qy        180 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
                                  |||:::         ::::::|||   ||||||
Db      38464 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 38523

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 211
              :::         |||         :::      :::      |||   :::   :::
Db      38524 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 38583

Qy        212 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
                      ||||||   :::    |||:::      :::||||||:::
Db      38584 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 38643

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
              :::      |||         |||      |||         :::|||:::   :::
Db      38644 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 38703

Qy        249 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 266
              |||         :::   |||   :::      |||      ||||||   |||
Db      38704 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 38763

Qy        267 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
                  |||                                 |||:::         :::
Db      38764 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 38823

Qy        285 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
              :::   |||         |||:::   :::         ||||||
Db      38824 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 38871

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321
                    :::      :::      |||:::         |||   |||::::::
Db      38872 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 38925

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu------------ 337
                  |||               :::      :::|||||||||||||
Db      38926 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 38976

Qy        338 ---GlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356
                 |||      :::|||            :::||||||||||||      |||:::|||
Db      38977 ATTGGACCCAACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGT 39036

Qy        357 Asn 357
              :::
Db      39037 GAC 39039

RESULT 24
US-11-266-748A-283521
; Sequence 283521, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 283521
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-283521

Alignment Scores:
Pred. No.:                0.000512        Length:          1000
Score:                    122.00          Matches:         56
Percent Similarity:       37.9%           Conservative:    43
Best Local Similarity:    21.5%           Mismatches:      114
Query Match:              6.4%            Indels:          48
DB:                       9               Gaps:            10

US-10-757-745-2 (1-362) x US-11-266-748A-283521 (1-1000)

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              ||||||          :::         :::      |||   ||||||::::::   |||
Db        173 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 232

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              |||||||||         |||:::         :::         |||
Db        233 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 280

Qy        171 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 179
                    |||::::::      :::|||
Db        281 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 337

Qy        180 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
                                  |||:::         ::::::|||   ||||||
Db        338 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 397

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 211
              :::         |||         :::      :::      |||   :::   :::
Db        398 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGACCATTGCA 457

Qy        212 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228
                      ||||||   :::    |||:::      :::||||||:::
Db        458 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 517

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248
              :::      |||         |||      |||         :::|||:::   :::
Db        518 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 577

Qy        249 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 266
              |||         :::   |||   :::      |||      ||||||   |||
Db        578 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 637

Qy        267 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284
                  |||                                 |||:::         :::
Db        638 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 697

Qy        285 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301
              :::   |||         |||:::   :::         ||||||
Db        698 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 745

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321
                    :::      :::      |||:::         |||   |||::::::
```

```
Db        746 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 799

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu------------ 337

Db        800 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 850

Qy        338 ---GlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356

Db        851 ATTGGACCCAACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGT 910

Qy        357 Asn 357

Db        911 GAC 913
```

RESULT 25
US-11-266-748A-310161/c
; Sequence 310161, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 310161
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-310161

Alignment Scores:

Pred. No.:	0.000512	Length:	1000
Score:	122.00	Matches:	56
Percent Similarity:	37.9%	Conservative:	43
Best Local Similarity:	21.5%	Mismatches:	114
Query Match:	6.4%	Indels:	48
DB:	9	Gaps:	10

US-10-757-745-2 (1-362) x US-11-266-748A-310161 (1-1000)

```
Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db        828 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 769

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db        768 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 721

Qy        171 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 179

Db        720 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 664

Qy        180 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
```

```
Db        663 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 604

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 211

Db        603 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGACCATTGCA 544

Qy        212 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 228

Db        543 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 484

Qy        229 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 248

Db        483 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 424

Qy        249 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 266

Db        423 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 364

Qy        267 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 284

Db        363 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 304

Qy        285 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 301

Db        303 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 256

Qy        302 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 321

Db        255 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 202

Qy        322 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu------------ 337

Db        201 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 151

Qy        338 ---GlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCys 356

Db        150 ATTGGACCCAACAGGTTACCTTCCTTCAATTATCCTTCAGACCACCTGTCTCTAGTGTGT 91

Qy        357 Asn 357

Db         90 GAC 88
```

RESULT 26
US-10-868-184C-10580/c
; Sequence 10580, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.

```
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10580
;   LENGTH: 6198
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10580

Alignment Scores:
Pred. No.:                0.0239          Length:         6198
Score:                    118.50          Matches:        75
Percent Similarity:       42.3%           Conservative:   52
Best Local Similarity:    25.0%           Mismatches:     100
Query Match:              6.2%            Indels:         74
DB:                       6               Gaps:           14

US-10-757-745-2 (1-362) x US-10-868-184C-10580 (1-6198)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::      |||:::                  :::
Db        4134 GAAGAAACTGCATCAACTAACAAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 4076

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   |||::::::||||||         |||   |||
Db        4075 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTTCA 4025

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                :::         :::   :::::::::             |||   |||      :::
Db        4024 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3965

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||      |||   ::::::
Db        3964 CAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 3923

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||     :::   |||                |||||:::   |||
Db        3922 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3863

Qy         191 LeuLysSerGlnGluIleIleProProSerThrLysMetMetArg------------ 206
               :::   :::   :::          |||      ||||||:::   |||
Db        3862 ATCCTAGCCTCTGATAAAACAGATTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3803

Qy         207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 226
               :::   :::   |||   |||::::::          ||||||   ::::::   :::
Db        3802 CATTACATAATGGTAAAGGTATCAATTCAACAAGAAGAATTAACTATCCTAAATACATGT 3743

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::||||||          |||   |||          :::|||   ||||||
Db        3742 GCACCCAATACAAGA---------GCATCCAGATTC------ATAAAGCAAGTCCTTAGT 3698

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp------------ 262
                  :::|||          :::|||   |||:::|||      ||||||
Db        3697 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCCCA 3641

Qy         263 -------------ThrAsnLeuArgAspArgGluValThrArgCys-GlyGlyLeu---- 276
                          |||:::   |||||||||:::::::||||||       |||:::
Db        3640 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTTAACAAGGATACCCAGGAATTCAAC 3581

Qy         277 -----------ProAsn-AsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                       |||::: :::::::::|||::::::      |||      ||||||
Db        3580 TCAGCTCTGCACCAAGTGGACCTAATAGACATCTACAGAATTCTC---CACCCCAAATCA 3524

Qy         293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
                  :::||||||
Db        3523 ACAGAATATACA---------------------------------------------- 3512

Qy         313 LeuArgPheAspArgIlePhePheArgAlaAla-----------AlaGluGluGlyHis 328
                        ||||||   |||                ::::::      |||
Db        3511 -----------------TTTTTTTCAGCACCACACCACACCTATTCCAAAATTGACCAC 3470

Qy         329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg------ 346
               |||:::   ||||||      ||||||                  |||   |||
Db        3469 ATAGTT---CGAAGTAAAGCTCTCCTCAGC-----------ACATGTAAAAGAACAGAA 3425

Qy         347 -----------PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
                          :::   |||||||||      :::         |||   |||
Db        3424 ATTATAACAAACTATCTCTCAGACCACAGTGCAATCAAACTAGAACTCAGTATT 3371

RESULT 27
US-10-868-184D-10580/c
; Sequence 10580, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10580
;   LENGTH: 6198
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10580

Alignment Scores:
Pred. No.:                0.0239          Length:         6198
Score:                    118.50          Matches:        75
Percent Similarity:       42.3%           Conservative:   52
Best Local Similarity:    25.0%           Mismatches:     100
Query Match:              6.2%            Indels:         74
DB:                       6               Gaps:           14

US-10-757-745-2 (1-362) x US-10-868-184D-10580 (1-6198)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::      |||:::                  :::
Db        4134 GAAGAAACTGCATCAACTAACAAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 4076

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   |||::::::||||||         |||   |||
Db        4075 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTTCA 4025

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                :::         :::   :::::::::             |||   |||      :::
Db        4024 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3965

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||      |||   ::::::
Db        3964 CAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 3923

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||     :::   |||                |||||:::   |||
Db        3922 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3863
```

```
Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg------------ 206
              :::   :::   :::          |||       ||||||:::   |||
Db       3862 ATCCTAGCCTCTGATAAAACAGATTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3803

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 226
              :::   :::   |||   |||::::::          ||||||   ::::::   :::
Db       3802 CATTACATAATGGTAAAGGTATCAATTCAACAAGAAGAATTAACTATCCTAAATACATGT 3743

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                 :::||||||          |||   |||         :::|||   ||||||
Db       3742 GCACCCAATACAAGA---------GCATCCAGATTC------ATAAAGCAAGTCCTTAGT 3698

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp------------ 262
                 :::|||       :::|||   |||:::|||      ||||||
Db       3697 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCCCA 3641

Qy        263 ------------ThrAsnLeuArgAspArgGluValThrArgCys-GlyGlyLeu---- 276
                       |||:::   |||||||||:::::|||||       |||:::
Db       3640 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTTAACAAGGATACCCAGGAATTCAAC 3581

Qy        277 -----------ProAsn-AsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                      |||::: :::::::::|||:::::      |||      ||||||
Db       3580 TCAGCTCTGCACCAAGTGGACCTAATAGACATCTACAGAATTCTC---CACCCCAAATCA 3524

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
                 :::||||||
Db       3523 ACAGAATATACA---------------------------------------------- 3512

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAla------------AlaGluGluGlyHis 328
                       ||||||   |||                 ::::::      |||
Db       3511 -----------------TTTTTTTCAGCACCACACCACACCTATTCCAAAATTGACCAC 3470

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg------ 346
              |||:::   ||||||      ||||||              |||   |||
Db       3469 ATAGTT---CGAAGTAAAGCTCTCCTCAGC------------ACATGTAAAAGAACAGAA 3425

Qy        347 ------------PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
                         :::   |||||||||       :::         |||   |||
Db       3424 ATTATAACAAACTATCTCTCAGACCACAGTGCAATCAAACTAGAACTCAGTATT 3371
```

RESULT 28
US-10-868-184C-10303/c
; Sequence 10303, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10303
; LENGTH: 6198
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-868-184C-10303

Alignment Scores:

Pred. No.:	0.0953	Length:	6198
Score:	113.50	Matches:	74
Percent Similarity:	42.0%	Conservative:	52
Best Local Similarity:	24.7%	Mismatches:	101
Query Match:	6.0%	Indels:	74
DB:	6	Gaps:	14

US-10-757-745-2 (1-362) x US-10-868-184C-10303 (1-6198)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::      |||:::                  :::
Db       4134 GAAGAAACTGCATCAACTAACAAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 4076

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||      :::::::::|||   |||::::::||||||          |||   |||
Db       4075 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTTCA 4025

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                 :::   :::   :::::::::          |||  |||   :::
Db       4024 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3965

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                ::::::|||      |||   ::::::
Db       3964 CAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 3923

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                ||||||:::   |||
Db       3922 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3863

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg------------ 206
              :::   :::   :::          |||       ||||||:::   |||
Db       3862 ATCCTAGCCTCTGATAAAACAGATTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3803

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 226
              :::   :::   |||   |||::::::          ||||||   ::::::   :::
Db       3802 CATTACATAATGGTAAAGGTATCAATTCAACAAGAAGAATTAACTATCCTAAATACATGT 3743

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                 :::|||          |||   |||         :::|||   ||||||
Db       3742 GCACCCAATACA---------GGAGCATCCAGATTC------ATAAAGCAAGTCCTTAGT 3698

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp------------ 262
                 :::|||       :::|||   |||:::|||      ||||||
Db       3697 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCCCA 3641

Qy        263 ------------ThrAsnLeuArgAspArgGluValThrArgCys-GlyGlyLeu---- 276
                       |||:::   |||||||||:::::|||||       |||:::
Db       3640 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTTAACAAGGATACCCAGGAATTCAAC 3581

Qy        277 -----------ProAsn-AsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                      |||::: :::::::::|||:::::      |||      ||||||
Db       3580 TCAGCTCTGCACCAAGTGGACCTAATAGACATCTACAGAATTCTC---CACCCCAAATCA 3524

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
                 :::||||||
Db       3523 ACAGAATATACA---------------------------------------------- 3512

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAla------------AlaGluGluGlyHis 328
                       ||||||   |||                 ::::::      |||
Db       3511 -----------------TTTTTTTCAGCACCACACCACACCTATTCCAAAATTGACCAC 3470

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg------ 346
              |||:::   ||||||      ||||||              |||   |||
Db       3469 ATAGTT---CGAAGTAAAGCTCTCCTCAGC------------ACATGTAAAAGAACAGAA 3425

Qy        347 ------------PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
                         :::   |||||||||       :::         |||   |||
```

```
Db       3424 ATTATAACAAACTATCTCTCAGACCACAGTGCAATCAAACTAGAACTCAGTATT 3371

RESULT 29
US-10-868-184D-10303/c
; Sequence 10303, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10303
;   LENGTH: 6198
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10303

Alignment Scores:
Pred. No.:                0.0953          Length:          6198
Score:                    113.50          Matches:         74
Percent Similarity:       42.0%           Conservative:    52
Best Local Similarity:    24.7%           Mismatches:      101
Query Match:              6.0%            Indels:          74
DB:                       6               Gaps:            14

US-10-757-745-2 (1-362) x US-10-868-184D-10303 (1-6198)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::      |||:::                 :::
Db       4134 GAAGAAACTGCATCAACTAACAAGCAAAATAA-CCAGCTAACATCATAATGACAGGATCA 4076

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                 |||      :::::::::|||  |||::::::||||||         |||   |||
Db       4075 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTTCA 4025

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                 :::        :::   :::::::::            |||   |||      :::
Db       4024 ATTAAAAGACACAGACTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3965

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  ::::::|||      |||   ::::::
Db       3964 CAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 3923

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                 ||||||:::   |||
Db       3922 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3863

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg------------ 206
              :::   :::   :::          |||      ||||||:::   |||
Db       3862 ATCCTAGCCTCTGATAAAACAGATTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3803

Qy        207 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 226
              :::   :::   |||   |||::::::         ||||||   ::::::   :::
Db       3802 CATTACATAATGGTAAAGGTATCAATTCAACAAGAAGAATTAACTATCCTAAATACATGT 3743

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                 :::|||            |||   |||          :::|||   ||||||
Db       3742 GCACCCAATACA---------GGAGCATCCAGATTC------ATAAAGCAAGTCCTTAGT 3698

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp------------ 262
                 :::|||         :::|||   |||:::|||       ||||||
Db       3697 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCCCA 3641

Qy        263 -------------ThrAsnLeuArgAspArgGluValThrArgCys-GlyGlyLeu---- 276
                          |||:::   |||||||||:::::::||||||       |||:::
Db       3640 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTTAACAAGGATACCCAGGAATTCAAC 3581

Qy        277 -----------ProAsn-AsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          |||::: ::::::::::|||::::::       |||       ||||||
Db       3580 TCAGCTCTGCACCAAGTGGACCTAATAGACATCTACAGAATTCTC---CACCCCAAATCA 3524

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
                 :::||||||
Db       3523 ACAGAATATACA---------------------------------------------- 3512

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAla------------AlaGluGluGlyHis 328
                          ||||||   |||                 ::::::      |||
Db       3511 -----------------TTTTTTTCAGCACCACACCACACCTATTCCAAAATTGACCAC 3470

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg------ 346
              |||:::   ||||||      ||||||                 |||   |||
Db       3469 ATAGTT---CGAAGTAAAGCTCTCCTCAGC------------ACATGTAAAAGAACAGAA 3425

Qy        347 ------------PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
                         :::   |||||||||      :::          |||   |||
Db       3424 ATTATAACAAACTATCTCTCAGACCACAGTGCAATCAAACTAGAACTCAGTATT 3371

RESULT 30
US-10-868-184C-10659/c
; Sequence 10659, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10659
;   LENGTH: 6075
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
US-10-868-184C-10659

Alignment Scores:
Pred. No.:                0.106          Length:          6075
Score:                    113.00         Matches:         62
Percent Similarity:       45.9%          Conservative:    43
Best Local Similarity:    27.1%          Mismatches:      83
Query Match:              5.9%           Indels:          42
DB:                       6              Gaps:            11

US-10-757-745-2 (1-362) x US-10-868-184C-10659 (1-6075)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||  ||||||      :::   ||||||:::                 :::
Db        4137 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4079

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||:::|||      ||||||:::
Db        4078 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4025

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||  |||   :::   |||::::::              |||   |||      :::
Db        4024 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3968

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  ::::::|||      |||   ::::::
Db        3967 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3926

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                 ||||||:::   |||
Db        3925 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3866

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::      :::        |||      ||||||:::   |||:::
Db        3865 ATCCTAGTCTCTGACAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3806

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                  ::::::   |||   |||:::            ||||||   ::::::      :::
Db        3805 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3746

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||          |||   |||         :::|||   ||||||
Db        3745 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 3701

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp----------- 262
                  :::|||      :::|||   |||:::|||      ||||||
Db        3700 CACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3644

Qy         263 -------------ThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro-- 277
                           |||:::   ||||||||||:::   ||||||         :::|||
Db        3643 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGG---------ATACCCAG 3593

Qy         278 ------------------------AsnAsnIleValAspValTrpGluPheLeuGlyLy 289
                                       |||:::::::|||::::::      |||
Db        3592 GAATTGAACTCAGCTCTGCACCAAGCAGAACCTAATAGACATCTACAGAACTCTC---CA 3536

Qy         289 sProLysHisCysGlnTyrThrTrp 297
               ||||||      :::||||||:::
Db        3535 CCCCAAATCAACAGAATATACATTT 3511

RESULT 31
US-10-868-184D-10659/c
; Sequence 10659, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10659
;   LENGTH: 6075
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10659

Alignment Scores:
Pred. No.:                0.106          Length:          6075
Score:                    113.00         Matches:         62
Percent Similarity:       45.9%          Conservative:    43
Best Local Similarity:    27.1%          Mismatches:      83
Query Match:              5.9%           Indels:          42
DB:                       6              Gaps:            11

US-10-757-745-2 (1-362) x US-10-868-184D-10659 (1-6075)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||  ||||||      :::   ||||||:::                 :::
Db        4137 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4079

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||:::|||      ||||||:::
Db        4078 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4025

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||  |||   :::   |||::::::              |||   |||      :::
Db        4024 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3968

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  ::::::|||      |||   ::::::
Db        3967 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3926

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                 ||||||:::   |||
Db        3925 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3866

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::      :::        |||      ||||||:::   |||:::
Db        3865 ATCCTAGTCTCTGACAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3806

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                  ::::::   |||   |||:::            ||||||   ::::::      :::
Db        3805 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3746

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||          |||   |||         :::|||   ||||||
Db        3745 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 3701

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp----------- 262
                  :::|||      :::|||   |||:::|||      ||||||
Db        3700 CACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3644

Qy         263 -------------ThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro-- 277
                           |||:::   ||||||||||:::   ||||||         :::|||
```

```
Db       3643 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGG---------ATACCCAG 3593

Qy        278 ------------------------AsnAsnIleValAspValTrpGluPheLeuGlyLy 289
                                      |||:::::|||::::::      |||
Db       3592 GAATTGAACTCAGCTCTGCACCAAGCAGAACCTAATAGACATCTACAGAACTCTC---CA 3536

Qy        289 sProLysHisCysGlnTyrThrTrp 297
              ||||||      :::||||||:::
Db       3535 CCCCAAATCAACAGAATATACATTT 3511

RESULT 32
US-10-953-349-28325
; Sequence 28325, Application US/10953349
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCONDED THERBY
;  FILE REFERENCE: 2750-1579PUS2
;  CURRENT APPLICATION NUMBER: US/10/953,349
;  CURRENT FILING DATE:  2004-09-30
;  NUMBER OF SEQ ID NOS: 40252
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28325
;   LENGTH: 1421
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
US-10-953-349-28325

Alignment Scores:
Pred. No.:              0.0123        Length:          1421
Score:                  112.50        Matches:           85
Percent Similarity:     37.6%         Conservative:      59
Best Local Similarity:  22.2%         Mismatches:       136
Query Match:            5.9%          Indels:           104
DB:                     6             Gaps:              20

US-10-757-745-2 (1-362) x US-10-953-349-28325 (1-1421)

Qy          57 AlaLeuAsnSerTyrPheGluProProValGluGlu---SerAlaLeuGluArgArgPro 75
               |||:::      :::|||:::|||   :::::::::   |||      ::::::||||||
Db          93 GCCATGAAGCGCTTCTTCCAGCCCGTACTCAAGGACGGATCCCCCGCCAAGAAGAGGCCC 152

Qy          76 ------GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 93
                     ||||||:::          ||||||              |||
Db         153 GCCGGAGCCGCGATCTCTGAC---------TGCGTTGAT---GGCCGCGCCCCCACCGCG 200

Qy          94 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 113
                           |||   |||:::   ::::::
Db         201 GCCGACGCCGGCGGGGACGGCCCTCCCGGCGAGGAGCCACGCAGG--------------- 245

Qy         114 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 133
               ||||||||||||   :::   |||   |||          |||:::
Db         246 ------TTCATCACGTGGAACGCCAACAGCCTTCTGCTCCGGATGAAGAGCGACTGGCCC 299

Qy         134 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 153
                  |||   ::::::|||          ||||||||||||   :::||||||||||
Db         300 ---GCGTTCTGTCAGTTCGTCGCCCGCCTCGACCCCGACGTCATCTGCGTCCAGGAAGTG 356

Qy         154 --------------------------IleProProTyrTyrSerTyrLeuLys------ 162
                                        |||         |||   ||||||
Db         357 CGGATGCCAGCAGCTGGTTCCAAAGGGGCACCTAAAAATCCTAGTGAGCTAAAAGACGAC 416

Qy         163 ---------LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 179
                        |||                 ::::::                    |||
Db         417 ACAACGTCGTCACGCGAAGAAAAACAGACAGTGCTGCGTGCTTTGTCGAGCTCACCTTTC 476

Qy         180 ------------------------------------------------ThrAlaIleMet 183
                                                              ||||||:::
Db         477 AAAGACTACCGTGTGTGGTGGTCCCTTTCAGATTCAAAATATGCAGGCACAGCTGTGTTT 536

Qy         184 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 203
               :::||||||           |||           :::         |||||||||
Db         537 ATGAAGAAGAAGTTTGAACCAAAGAAAGTGTCATTCAACTTGGATAGAACATCTACTAAG 596

Qy         204 -MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMe 223
                ||||||:::   ::::::           |||:::      |||      |||
Db         597 CATGATGCAGACGGTCGTGTCATCATTGTGGAATTTG------AATCATTCTTGTTGCTA 650

Qy         223 tThrSerHisLeuGluSerThrArgGly--------------HisAlaAlaGluArgMe 238
                |||       |||::::::|||   |||               |||      |||
Db         651 AACACATATGCTCCAAACAACGGGTGGAAAGAGGAGGAAAATTCATTTCAAAGAAGAAGA 710

Qy         238 tAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIl 258
                :::   :::|||   :::   :::   :::|||:::      :::         :::||
Db         711 AAGTGGGATAAAAGA-ATGTCGGAGTTTGTTCAACAAGTAGACAAG------CCCTTAAT 763

Qy         258 ePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys-------------- 273
               |:::   ||||||   |||:::         |||:::
Db         764 CTGGTGCGGAGACCTGAACGTTAGTCATGACGAAATCGATGTCAGCCATCCTGATTTTTT 823

Qy         274 -------------------GlyGlyLeuProAsnAsn---------------------- 279
                                  |||  :::|||  |||
Db         824 TAGCGATGCTAAGCTCAATGGATATATCCCTCCAAATAAAGAGGACTGTGGACAGCCAGG 883

Qy         280 ------------------------------------------------------IleValAs 282
                                                                    :::||||
Db         884 ATTCACCCTTGCAGAGAGACAGCGCTTTGGCAATATATTGTCCCAGGGAAAGCTGGTAGA 943

Qy         282 pValTrpGluPheLeuGlyLysProLys------HisCysGlnTyrThrTrpAspThrGl 300
               |   :::   :::|||   |||  |||          |||   ::::::|||
Db         944 TGCTTACAGATACCTGCACAAAGAAAAAGACATGGAGTGTGGCTTCTCTTGGTCT----- 998

Qy         300 nMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePh 320
                     :::   :::|||                ::::::|||   |||      |||
Db         999 ----GGGAATCCAGTTGGGAAGTACCGAGGAAAGAGAATGCGGATTGAC---TATTTTGT 1051

Qy         320 eArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGl 340
                     :::          :::      |||:::          :::         |||:::||
Db        1052 TATTTCAGAAGGGTTGAAAGACAGAATTGTTTCATGTGAAATGCATGGTCATGGAATTGA 1111

Qy         340 uLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIl 360
               |   |||:::   |||   :::   |||||||||              |||   :::   ::
Db        1112 A---CTGGAA---GGATTTTATGGAAGTGATCAC------------TGTCCTGTTTCACT 1153

Qy         360 eIleLeu 362
               ::::|||
Db        1154 TGTGCTC 1160

RESULT 33
US-10-276-817B-96
; Sequence 96, Application US/10276817B
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-104
;  CURRENT APPLICATION NUMBER: US/10/276,817B
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: PCT/US01/14827
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 16102
;  SOFTWARE: Custom
; SEQ ID NO 96
;   LENGTH: 5130
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (2614)..(4473)
;   OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
;   OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-276-817B-96

Alignment Scores:
Pred. No.:              0.162         Length:          5130
```

Score:	110.50	Matches:	73
Percent Similarity:	43.0%	Conservative:	55
Best Local Similarity:	24.5%	Mismatches:	112
Query Match:	5.8%	Indels:	59
DB:	6	Gaps:	15

US-10-757-745-2 (1-362) x US-10-276-817B-96 (1-5130)

```
Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
              :::::::::::   ||||||       |||   ::::::      |||   ||||||
Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
                 |||             |||         |||
Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
               ||||||        :::    |||:::   |||       |||:::
Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
                           ||||||::::::                 :::   |||
Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
                 |||||||||      |||:::      |||   |||      :::::::::|||
Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
               |||:::::||||         |||       :::       :::       :::::
Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
               ::::        |||   |||      :::||||||                ::
Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
               ::::|||      |||   ::::::       |||   |||      |||   |||
Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
                        ||||||:::   |||   :::       :::       ||
Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy        199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215
               |      ||||||:::   |||           :::   :::   |||      :::::
Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
               :      ||||||      :::      :::      :::|||              ||
Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
               |  |||      :::|||  ||||||:::  :::|||          :::|||
Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267
               |||:::|||      ||||||                      |||:::   ||||
Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
               |||||::::::||||||        :::   |||                :::::::
Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
               ::|||::::::      |||      ||||||      :::|||||:::
Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728
```

RESULT 34
US-10-276-817B-3643
; Sequence 3643, Application US/10276817B
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: US/10/276,817B
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US01/14827
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 3643
; LENGTH: 5130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2614)..(4473)
; OTHER INFORMATION: 94% homologous to Homo sapiens putative p150,accession number
; OTHER INFORMATION: U93572,Smith-Waterman Score=3044.
US-10-276-817B-3643

Alignment Scores:			
Pred. No.:	0.162	Length:	5130
Score:	110.50	Matches:	73
Percent Similarity:	43.0%	Conservative:	55
Best Local Similarity:	24.5%	Mismatches:	112
Query Match:	5.8%	Indels:	59
DB:	6	Gaps:	15

US-10-757-745-2 (1-362) x US-10-276-817B-3643 (1-5130)

```
Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
              :::::::::::   ||||||       |||   ::::::      |||   ||||||
Db        948 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 995

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
                 |||             |||         |||
Db        996 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 1031

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
               ||||||        :::    |||:::   |||       |||:::
Db       1032 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 1088

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
                           ||||||::::::                 :::   |||
Db       1089 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 1130

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
                 |||||||||      |||:::      |||   |||      :::::::::|||
Db       1131 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 1187

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
               |||:::::||||         |||       :::       :::       :::::
Db       1188 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 1238

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
               ::::        |||   |||      :::||||||                ::
Db       1239 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 1280

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
               ::::|||      |||   ::::::       |||   |||      |||   |||
Db       1281 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 1340

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
                        ||||||:::   |||   :::       :::       ||
Db       1341 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 1400

Qy        199 eProSerThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnVa 215
               |      ||||||:::   |||           :::   :::   |||      :::::
Db       1401 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 1460

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
               :      ||||||      :::      :::      :::|||              ||
Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
               |  |||      :::|||  ||||||:::  :::|||          :::|||
Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267
               |||:::|||      ||||||                      |||:::   ||||
Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
               |||||::::::||||||        :::   |||                :::::::
Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
               ::|||::::::      |||      ||||||      :::|||||:::
Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728
```

```
Db       1461 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 1511

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255

Db       1512 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 1565

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267

Db       1566 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 1622

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281

Db       1623 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 1681

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297

Db       1682 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 1728

RESULT 35
US-10-461-673-4287
; Sequence 4287, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 4287
;   LENGTH: 8065
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-461-673-4287

Alignment Scores:
Pred. No.:                0.331          Length:          8065
Score:                    110.50         Matches:         73
Percent Similarity:       43.0%          Conservative:    55
Best Local Similarity:    24.5%          Mismatches:      112
Query Match:              5.8%           Indels:          59
DB:                       6              Gaps:            15

US-10-757-745-2 (1-362) x US-10-461-673-4287 (1-8065)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41

Db       4302 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 4349

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61

Db       4350 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 4385

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81

Db       4386 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 4442

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100

Db       4443 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 4484

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119

Db       4485 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 4541

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139

Db       4542 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 4592

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159

Db       4593 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 4634

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179

Db       4635 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 4694

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199

Db       4695 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 4754

Qy        199 eProSerThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnVa 215

Db       4755 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 4814

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235

Db       4815 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 4865

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255

Db       4866 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 4919

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267

Db       4920 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 4976

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281

Db       4977 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 5035

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297

Db       5036 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 5082
```

```
RESULT 36
US-10-461-673-7266
; Sequence 7266, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 7266
;   LENGTH: 10136
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-461-673-7266

Alignment Scores:
Pred. No.:              0.474          Length:          10136
Score:                  110.50         Matches:         73
Percent Similarity:     43.0%          Conservative:    55
Best Local Similarity:  24.5%          Mismatches:      112
Query Match:            5.8%           Indels:          59
DB:                     6              Gaps:            15

US-10-757-745-2 (1-362) x US-10-461-673-7266 (1-10136)

Qy         22 ValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAlaAla 41
              :::::::::::   ||||||          |||   ::::::      |||   ||||||
Db       5954 ATCAGGCAGCAACATCTGCTG---------TTCACCAATATC---CGCTGTTCTGCAGCC 6001

Qy         42 ValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSerTyr 61
              |||           |||           |||
Db       6002 TCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGT---------------------- 6037

Qy         62 PheGluProProValGluGluSerAla-LeuGluArgArgProGluThrIleSerGluPr 81
              ||||||      :::   |||:::  |||      |||:::
Db       6038 ---GGACCTCCAGCAAACTCCAACAGACCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAA 6094

Qy         81 oLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle-- 100
                          ||||||::::::                  :::   |||
Db       6095 A-----------------ACTAACAAACAGAAAGGACATCCACACCAAAACCCCATCTG 6136

Qy        101 ----SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTr 119
                  |||||||||      |||:::      |||   |||      :::::::::|||
Db       6137 TACGTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTT 6193

Qy        119 pAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTy 139
               |||::::::||||||         |||         :::         :::   :::::
Db       6194 AAATGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTG 6244

Qy        139 rLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSe 159
               ::::           |||   |||      :::||||||                  ::
Db       6245 GATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------AC 6286

Qy        159 rTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPh 179
               ::::|||      |||   ::::::          |||   |||      |||   |||
Db       6287 CCATCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCA 6346

Qy        179 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPh 199
                          ||||||:::   |||   :::         :::          ||
Db       6347 AGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTT 6406

Qy        199 eProSerThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnVa 215
               |      ||||||:::   |||             :::   :::   |||      :::::
Db       6407 TAAACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAAT 6466

Qy        215 lSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAl 235
               :          ||||||      :::       :::       :::|||           ||
Db       6467 TCAACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGC 6517

Qy        235 aGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAl 255
               |  |||          :::|||   ||||||:::   :::|||          :::|||
Db       6518 ACCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCA 6571

Qy        255 aThrValIlePheAlaGlyAsp------------------------ThrAsnLeuArgA 267
               |||:::|||      ||||||                        |||:::   ||||
Db       6572 CACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAG 6628

Qy        267 spArgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleV 281
               |||||::::::||||||          :::   |||                :::::::
Db       6629 ACAGAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAA 6687

Qy        281 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
               ::|||::::::      |||      ||||||    :::||||||:::
Db       6688 TAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 6734

RESULT 37
US-10-868-184C-10371/c
; Sequence 10371, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
```

```
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10371
;   LENGTH: 6063
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10371

Alignment Scores:
Pred. No.:              0.242          Length:          6063
Score:                  110.00         Matches:         56
Percent Similarity:     44.0%          Conservative:    43
Best Local Similarity:  24.9%          Mismatches:      92
Query Match:            5.8%           Indels:          35
DB:                     6              Gaps:            9

US-10-757-745-2 (1-362) x US-10-868-184C-10371 (1-6063)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||     :::   ||||||:::                   :::
Db       4121 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4063

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   |||:::::|||      ||||||:::
Db       4062 AATTCACACATAACAATATTAACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 4009

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||   :::          |||   |||   ::::::
Db       4008 AAAAGACACAGA---CTGGCAAGTTCGATAAAGAGTCAAGACCCATCAGTGTGCTATATT 3952

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::|||      |||   ::::::
Db       3951 CAGGAA----------------ACCCATCTTACGTGCAGAGACACACATAGGCTCAAA 3910

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                  ||||||:::   |||
Db       3909 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3850

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::       :::         |||      ||||||:::   |||:::
Db       3849 ATCCTAGTCTCTGATAAAGCAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3790

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::             ||||||   ::::::      |||
Db       3789 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATACAT 3730

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||   |||          :::|||   ||||||
Db       3729 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3685

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db       3684 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3628

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db       3627 CTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3568

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
```

```
                                  :::::::::|||:::::      |||      ||||||
Db       3567 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3511

Qy        293 CysGlnTyrThrTrp 297
                 :::|||||:::
Db       3510 ACAGAATATACATTT 3496

RESULT 38
US-10-868-184D-10371/c
; Sequence 10371, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10371
;   LENGTH: 6063
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10371

Alignment Scores:
Pred. No.:              0.242          Length:          6063
Score:                  110.00         Matches:         56
Percent Similarity:     44.0%          Conservative:    43
Best Local Similarity:  24.9%          Mismatches:      92
Query Match:            5.8%           Indels:          35
DB:                     6              Gaps:            9

US-10-757-745-2 (1-362) x US-10-868-184D-10371 (1-6063)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||     :::   ||||||:::                   :::
Db       4121 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4063

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   |||:::::|||      ||||||:::
Db       4062 AATTCACACATAACAATATTAACTTTAAATGTAAATGGA------CTAAATTCTCCAATT 4009

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||   :::          |||   |||   ::::::
Db       4008 AAAAGACACAGA---CTGGCAAGTTCGATAAAGAGTCAAGACCCATCAGTGTGCTATATT 3952

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::|||      |||   ::::::
Db       3951 CAGGAA----------------ACCCATCTTACGTGCAGAGACACACATAGGCTCAAA 3910

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                  ||||||:::   |||
```

```
Db       3909 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3850

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||         |||||:::    |||:::
Db       3849 ATCCTAGTCTCTGATAAAGCAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3790

Qy        211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
                 ::::::   |||   |||:::             ||||||   ::::::      |||
Db       3789 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATACAT 3730

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||             |||   |||          :::|||   ||||||
Db       3729 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3685

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||         :::|||   |||:::|||       ||||||   |||
Db       3684 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3628

Qy        266 ---------------------------------------ArgAspArgGluValThrArg 272
                                                      :::|||   :::
Db       3627 CTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3568

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||     ||||||
Db       3567 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3511

Qy        293 CysGlnTyrThrTrp 297
                 :::||||||:::
Db       3510 ACAGAATATACATTT 3496

RESULT 39
US-10-868-184C-10476/c
; Sequence 10476, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10476
;   LENGTH: 6093
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10476

Alignment Scores:
Pred. No.:              0.244          Length:         6093
Score:                  110.00         Matches:        56
Percent Similarity:     44.4%          Conservative:   44
Best Local Similarity:  24.9%          Mismatches:     91

Query Match:            5.8%           Indels:         35
DB:                     6              Gaps:           9

US-10-757-745-2 (1-362) x US-10-868-184C-10476 (1-6093)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   |||||||:::                  :::
Db       4129 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4071

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||       :::::::::|||   ||||||:::|||      ||||||:::
Db       4070 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4017

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||::::::              |||  |||      :::
Db       4016 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3960

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::|||     |||   :::::
Db       3959 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3918

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||       :::   |||                  ||||||:::||||||
Db       3917 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAAGGGTTGCA 3858

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||         ||||||:::   ::::::
Db       3857 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAAAGACAAAGAAGGC 3798

Qy        211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
                 ::::::   |||   |||:::             ||||||   ::::::      |||
Db       3797 CATTACATAATGGTAAAAGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3738

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||             |||   |||          :::|||   ||||||
Db       3737 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3693

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||         :::|||   |||:::|||       ||||||   |||
Db       3692 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3636

Qy        266 ---------------------------------------ArgAspArgGluValThrArg 272
                                                      :::|||   :::
Db       3635 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3576

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||     ||||||
Db       3575 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3519

Qy        293 CysGlnTyrThrTrp 297
                 :::||||||:::
Db       3518 ACAGAATATACATTT 3504

RESULT 40
US-10-868-184D-10476/c
; Sequence 10476, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
```

```
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10476
;   LENGTH: 6093
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10476

Alignment Scores:
Pred. No.:                0.244          Length:          6093
Score:                    110.00         Matches:         56
Percent Similarity:       44.4%          Conservative:    44
Best Local Similarity:    24.9%          Mismatches:      91
Query Match:              5.8%           Indels:          35
DB:                       6              Gaps:            9

US-10-757-745-2 (1-362) x US-10-868-184D-10476 (1-6093)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||  ||||||      :::   ||||||:::                  :::
Db       4129 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4071

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      ::::::::::|||   ||||||:::|||      ||||||:::
Db       4070 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4017

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::            |||   |||     :::
Db       4016 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3960

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                 :::::::|||      |||   ::::::
Db       3959 CAGGAA----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3918

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                 ||||||:::|||||||
Db       3917 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAAGGGTTGCA 3858

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||      ||||||:::   ::::::
Db       3857 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAAAGACAAAGAAGGC 3798

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::            ||||||   ::::::      :::
Db       3797 CATTACATAATGGTAAAAGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3738

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||   |||         :::|||   ||||||
Db       3737 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3693

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db       3692 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3636

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                  :::|||   :::
Db       3635 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3576

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                         :::::::::|||::::::      |||      ||||||
Db       3575 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3519

Qy        293 CysGlnTyrThrTrp 297
```

```
                 :::||||||:::
Db       3518 ACAGAATATACATTT 3504

RESULT 41
US-10-868-184C-10470/c
; Sequence 10470, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10470
;   LENGTH: 5381
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10470

Alignment Scores:
Pred. No.:                0.264          Length:          5381
Score:                    109.00         Matches:         54
Percent Similarity:       43.1%          Conservative:    43
Best Local Similarity:    24.0%          Mismatches:      94
Query Match:              5.7%           Indels:          35
DB:                       6              Gaps:            8

US-10-757-745-2 (1-362) x US-10-868-184C-10470 (1-5381)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||  ||||||      :::   ||||||:::                  :::
Db       4113 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4055

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      ::::::::::|||   |||::::::||||||          |||
Db       4054 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTCCA 4004

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::   :::  |||::::::            |||   |||     :::
Db       4003 ATTAAAAGACACAGACTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3944

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                 ::::::|||      |||   ::::::
Db       3943 CAGGAA----------------ACCCATCTCACGTGCAGAGATACACATAGGCTCAAA 3902

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||   |||             ||||||:::   |||
Db       3901 ATAAAAGGATGGAGGAAGATCTACCAAACAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3842

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||      ||||||:::   |||:::
```

```
Db       3841 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3782

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::  |||   |||:::          |||||  ::::::      :::
Db       3781 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATATAT 3722

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                :::|||          |||   |||         :::|||   ||||||
Db       3721 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3677

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               ::|||        :::|||   |||:::|||      ||||||  |||
Db       3676 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3620

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                  :::|||   :::
Db       3619 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3560

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                         :::::::::|||::::::      |||     ||||||
Db       3559 TCAGCTCTGTACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3503

Qy        293 CysGlnTyrThrTrp 297
                ::::||||||:::
Db       3502 ACAGAATATACATTT 3488

RESULT 42
US-10-868-184C-10513/c
; Sequence 10513, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10513
;    LENGTH: 5381
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-868-184C-10513

Alignment Scores:
Pred. No.:                0.264          Length:          5381
Score:                    109.00         Matches:         54
Percent Similarity:       43.1%          Conservative:    43
Best Local Similarity:    24.0%          Mismatches:      94
Query Match:              5.7%           Indels:          35
DB:                       6              Gaps:            8

US-10-757-745-2 (1-362) x US-10-868-184C-10513 (1-5381)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                 :::
Db       4113 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4055

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      :::::::::|||   |||::::::||||||          |||
Db       4054 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTCCA 4004

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::         :::   |||::::::              |||   |||      :::
Db       4003 ATTAAAAGACACAGACTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3944

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                ::::::|||      |||   ::::::
Db       3943 CAGGAA----------------ACCCATCTCACGTGCAGAGATACACATAGGCTCAAA 3902

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||  |||      :::   |||  |||              ||||||:::   |||
Db       3901 ATAAAAGGATGGAGGAAGATCTACCAAACAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3842

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::         :::         |||      ||||||:::   |||:::
Db       3841 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3782

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::  |||   |||:::          |||||  ::::::      :::
Db       3781 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATATAT 3722

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                :::|||          |||   |||         :::|||   ||||||
Db       3721 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3677

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               ::::|||        :::|||   |||:::|||      ||||||  |||
Db       3676 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3620

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                  :::|||   :::
Db       3619 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3560

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                         :::::::::|||::::::      |||     ||||||
Db       3559 TCAGCTCTGTACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3503

Qy        293 CysGlnTyrThrTrp 297
                ::::||||||:::
Db       3502 ACAGAATATACATTT 3488

RESULT 43
US-10-868-184D-10470/c
; Sequence 10470, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
```

```
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10470
;   LENGTH: 5381
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10470

Alignment Scores:
Pred. No.:                0.264          Length:          5381
Score:                    109.00         Matches:         54
Percent Similarity:       43.1%          Conservative:    43
Best Local Similarity:    24.0%          Mismatches:      94
Query Match:              5.7%           Indels:          35
DB:                       6              Gaps:            8

US-10-757-745-2 (1-362) x US-10-868-184D-10470 (1-5381)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::                  :::
Db        4113 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4055

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      ::::::::::|||   |||::::::||||||         |||
Db        4054 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTCCA 4004

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::         :::   |||::::::             |||   |||      :::
Db        4003 ATTAAAAGACACAGACTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3944

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  ::::::|||      |||   ::::::
Db        3943 CAGGAA-----------------ACCCATCTCACGTGCAGAGATACACATAGGCTCAAA 3902

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||     :::   |||   |||           ||||||:::   |||
Db        3901 ATAAAAGGATGGAGGAAGATCTACCAAACAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3842

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::         :::          |||       |||||:::   |||:::
Db        3841 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3782

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::            ||||||   ::::::      :::
Db        3781 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATATAT 3722

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||           |||   |||          :::|||   ||||||
Db        3721 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3677

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||          :::|||   |||:::||       ||||||   |||
Db        3676 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3620

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db        3619 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3560

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                         :::::::::|||::::::      |||      ||||||
Db        3559 TCAGCTCTGTACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3503

Qy         293 CysGlnTyrThrTrp 297
               :::||||||:::
Db        3502 ACAGAATATACATTT 3488

RESULT 44
```

```
US-10-868-184D-10513/c
; Sequence 10513, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10513
;   LENGTH: 5381
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10513

Alignment Scores:
Pred. No.:                0.264          Length:          5381
Score:                    109.00         Matches:         54
Percent Similarity:       43.1%          Conservative:    43
Best Local Similarity:    24.0%          Mismatches:      94
Query Match:              5.7%           Indels:          35
DB:                       6              Gaps:            8

US-10-757-745-2 (1-362) x US-10-868-184D-10513 (1-5381)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::                  :::
Db        4113 GAAGAAACTGCATCAACTAACGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4055

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      ::::::::::|||   |||::::::||||||         |||
Db        4054 AATTCACACATAACAATATTAACTTTAAATGTAAATGGACTA---------AATGCTCCA 4004

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::         :::   |||::::::             |||   |||      :::
Db        4003 ATTAAAAGACACAGACTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3944

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  ::::::|||      |||   ::::::
Db        3943 CAGGAA-----------------ACCCATCTCACGTGCAGAGATACACATAGGCTCAAA 3902

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||     :::   |||   |||           ||||||:::   |||
Db        3901 ATAAAAGGATGGAGGAAGATCTACCAAACAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3842

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::         :::          |||       |||||:::   |||:::
Db        3841 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3782

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::            ||||||   ::::::      :::
```

```
Db       3781 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATATAT 3722

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db       3721 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3677

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db       3676 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3620

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db       3619 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3560

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db       3559 TCAGCTCTGTACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3503

Qy        293 CysGlnTyrThrTrp 297

Db       3502 ACAGAATATACATTT 3488

RESULT 45
PCT-US05-40306-111
; Sequence 111, Application PC/TUS0540306
; GENERAL INFORMATION:
;  APPLICANT: E.I. duPont de Nemours & Co., Inc.
;  TITLE OF INVENTION: HIGH ARACHIDONIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA
;  FILE REFERENCE: CL3136 PCT
;  CURRENT APPLICATION NUMBER: PCT/US05/40306
;  CURRENT FILING DATE:  2005-11-14
;  NUMBER OF SEQ ID NOS: 375
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 111
;   LENGTH: 1434
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
PCT-US05-40306-111

Alignment Scores:
Pred. No.:              0.0432          Length:         1434
Score:                  108.00          Matches:        78
Percent Similarity:     35.2%           Conservative:   47
Best Local Similarity:  22.0%           Mismatches:     124
Query Match:            5.7%            Indels:         106
DB:                     1               Gaps:           20

US-10-757-745-2 (1-362) x PCT-US05-40306-111 (1-1434)

Qy         90 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 109

Db         37 AAGGAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCAATTATGTCCGAT 96

Qy        110 AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGlyLeuAspLeu--- 126

Db         97 AATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGGTTAAAATACGTC 156

Qy        127 ---AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145

Db        157 TCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCGGGCCACTCAATG 216

Qy        145 ---------------------------------------------------------- 145

Db        217 CTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAATGAAAACGAAGAT 276

Qy        146 ---AspValIlePheLeuGlnGluVal---IleProProTyrTyrSerTyrLeuLysLys 163

Db        277 TACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAGTATCTAGCTTCT 336

Qy        164 Arg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr------ 180

Db        337 GCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATTCTGACGGGGCCT 396

Qy        181 ------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIlePro 198

Db        397 GGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACCTTTCTTTACCGG 447

Qy        199 PhePro----------------------------------SerThrLysMetMetArgAsn 207

Db        448 TTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTACGTAGGGAAATCT 507

Qy        208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227

Db        508 ATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATGAACAGTCACATG 567

Qy        228 Glu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln------------ 240

Db        568 CACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGATCTTGTCAGGCC 627

Qy        241 ------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258

Db        628 TGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGTTATGCGGTGATT 681

Qy        259 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 278

Db        682 GTGGTGGGTGACTTAAAC-----------------TCCAGACCGGGCTCACTGCCCCAC 723

Qy        279 ---------------------AsnIleValAspValTrpGluPheLeu---GlyLysPro 290

Db        724 AAATTTCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTGCATGGGAAGCAA 783

Qy        291 ---------------------------------------------LysHisCysGln--- 294

Db        784 GACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTTAAGGGCTGTACCACG 843

Qy        295 --------------TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 309

Db        844 TGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAACCCGAT------------GAG 891

Qy        310 AlaCysLysLeuArgPhe---------AspArgIlePhePheArgAlaAlaAlaGluGlu 326

Db        892 GCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACAGTAGACGCAGGT 951

Qy        327 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 346

Db        952 GTCAGGTTCACTGAACGG-----------------ATCCCTCACCTGGACTGCAGT--- 990

Qy        347 PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361

Db        991 ---GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 1032

RESULT 46
PCT-US05-40255-124
; Sequence 124, Application PC/TUS0540255
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: HIGH EICOSAPENTAENOIC ACID PRODUCING STRAINS OF YARROWIA
;  TITLE OF INVENTION:  LIPOLYTICA
;  FILE REFERENCE: CL2698 PCT
;  CURRENT APPLICATION NUMBER: PCT/US05/40255
;  CURRENT FILING DATE:  2005-11-15
;  PRIOR APPLICATION NUMBER: US 60/624812
;  PRIOR FILING DATE: 2004-11-04
;  NUMBER OF SEQ ID NOS: 414
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 124
;   LENGTH: 1434
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
PCT-US05-40255-124

Alignment Scores:
Pred. No.:              0.0432          Length:         1434
Score:                  108.00          Matches:        78
Percent Similarity:     35.2%           Conservative:   47
Best Local Similarity:  22.0%           Mismatches:     124
```

```
Query Match:          5.7%      Indels:       106
DB:                   1         Gaps:          20

US-10-757-745-2 (1-362) x PCT-US05-40255-124 (1-1434)

Qy         90 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 109
              :::|||        |||             ::::::   :::                :::
Db         37 AAGGAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCAATTATGTCCGAT 96

Qy        110 AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGlyLeuAspLeu--- 126
              |||         :::|||:::      :::|||:::|||      ||||||
Db         97 AATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGGTTAAAATACGTC 156

Qy        127 ---AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
                 :::      ||||||  |||  :::       ||||||   :::|||
Db        157 TCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCGGGCCACTCAATG 216

Qy        145 ------------------------------------------------------------ 145

Db        217 CTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAATGAAAACGAAGAT 276

Qy        146 ---AspValIlePheLeuGlnGluVal---IleProProTyrTyrSerTyrLeuLysLys 163
                 |||||||||  |||||||||:::              :::   ||||||
Db        277 TACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAGTATCTAGCTTCT 336

Qy        164 Arg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr------ 180
                 :::|||   |||                          |||       |||
Db        337 GCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATTCTGACGGGGCCT 396

Qy        181 ------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198
                    ||||||:::         |||:::|||   ::::::|||      :::
Db        397 GGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACCTTTCTTTACCGG 447

Qy        199 PhePro---------------------------------SerThrLysMetMetArgAsn 207
                 ||||||                                    :::   ::::::
Db        448 TTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTACGTAGGGAAATCT 507

Qy        208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
                 :::        :::|||            :::   :::|||   ||||||:::
Db        508 ATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATGAACAGTCACATG 567

Qy        228 Glu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln------------ 240
                          :::   :::|||   ||||||         ::::::
Db        568 CACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGATCTTGTCAGGCC 627

Qy        241 ------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258
                       ::::::|||      ::::::|||                  ||||||
Db        628 TGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGTTATGCGGTGATT 681

Qy        259 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 278
                 ||||||   |||                 :::|||   |||   ||||||:::
Db        682 GTGGTGGGTGACTTAAAC----------------TCCAGACCGGGCTCACTGCCCCAC 723

Qy        279 --------------------AsnIleValAspValTrpGluPheLeu---GlyLysPro 290
                              ::::||||||   ||||||   |||   ||||||
Db        724 AAATTTCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTGCATGGGAAGCAA 783

Qy        291 ----------------------------------------------LysHisCysGln--- 294
                                                           |||   |||
Db        784 GACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTTAAGGGCTGTACCACG 843

Qy        295 --------------TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 309
                           ||||||     |||      :::
Db        844 TGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAACCCGAT------------GAG 891

Qy        310 AlaCysLysLeuArgPhe---------AspArgIlePhePheArgAlaAlaAlaGluGlu 326
              ||||||:::|||   :::         |||      |||   :::
Db        892 GCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACAGTAGACGCAGGT 951

Qy        327 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 346
                    |||                  :::      |||||||||
Db        952 GTCAGGTTCACTGAACGG----------------ATCCCTCACCTGGACTGCAGT--- 990
```

```
Qy        347 PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
                 |||||||||:::          |||   |||:::|||:::
Db        991 ---GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 1032

RESULT 47
PCT-US05-40256-152
; Sequence 152, Application PC/TUS0540256
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: DOCOSAHEXAENOIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA
;  FILE REFERENCE: CL3160 PCT
;  CURRENT APPLICATION NUMBER: PCT/US05/40256
;  CURRENT FILING DATE:  2005-11-15
;  NUMBER OF SEQ ID NOS: 467
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 152
;   LENGTH: 1434
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
PCT-US05-40256-152

Alignment Scores:
Pred. No.:              0.0432      Length:        1434
Score:                  108.00      Matches:       78
Percent Similarity:     35.2%       Conservative:  47
Best Local Similarity:  22.0%       Mismatches:    124
Query Match:            5.7%        Indels:        106
DB:                     1           Gaps:          20

US-10-757-745-2 (1-362) x PCT-US05-40256-152 (1-1434)

Qy         90 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 109
              :::|||        |||             ::::::   :::                :::
Db         37 AAGGAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCAATTATGTCCGAT 96

Qy        110 AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGlyLeuAspLeu--- 126
              |||         :::|||:::      :::|||:::|||      ||||||
Db         97 AATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGGTTAAAATACGTC 156

Qy        127 ---AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
                 :::      ||||||  |||  :::       ||||||   :::|||
Db        157 TCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCGGGCCACTCAATG 216

Qy        145 ------------------------------------------------------------ 145

Db        217 CTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAATGAAAACGAAGAT 276

Qy        146 ---AspValIlePheLeuGlnGluVal---IleProProTyrTyrSerTyrLeuLysLys 163
                 |||||||||  |||||||||:::              :::   ||||||
Db        277 TACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAGTATCTAGCTTCT 336

Qy        164 Arg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr------ 180
                 :::|||   |||                          |||       |||
Db        337 GCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATTCTGACGGGGCCT 396

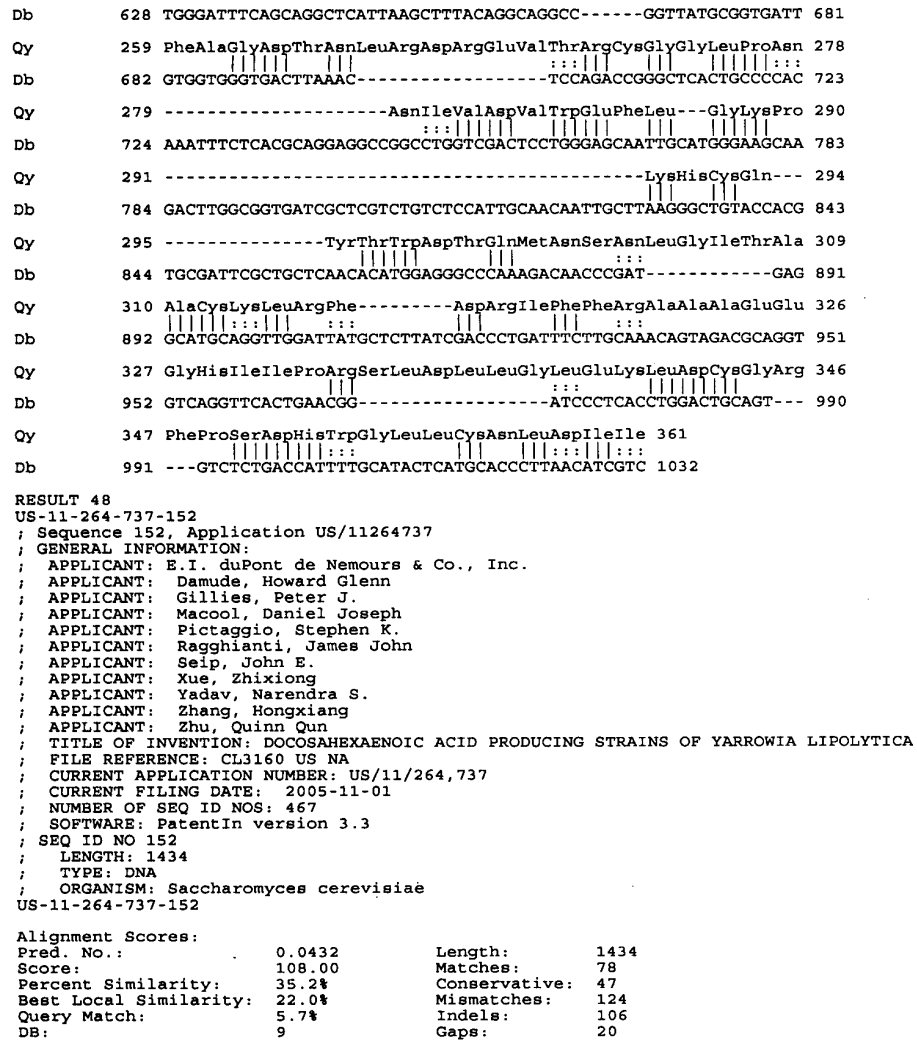
Qy        181 ------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198
                    ||||||:::         |||:::|||   ::::::|||      :::
Db        397 GGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACCTTTCTTTACCGG 447

Qy        199 PhePro---------------------------------SerThrLysMetMetArgAsn 207
                 ||||||                                    :::   ::::::
Db        448 TTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTACGTAGGGAAATCT 507

Qy        208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
                 :::        :::|||            :::   :::|||   ||||||:::
Db        508 ATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATGAACAGTCACATG 567

Qy        228 Glu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln------------ 240
                          :::   :::|||   ||||||         ::::::
Db        568 CACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGATCTTGTCAGGCC 627

Qy        241 ------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258
                       ::::::|||      ::::::|||                  ||||||
```

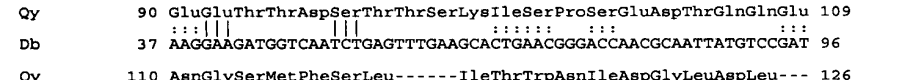


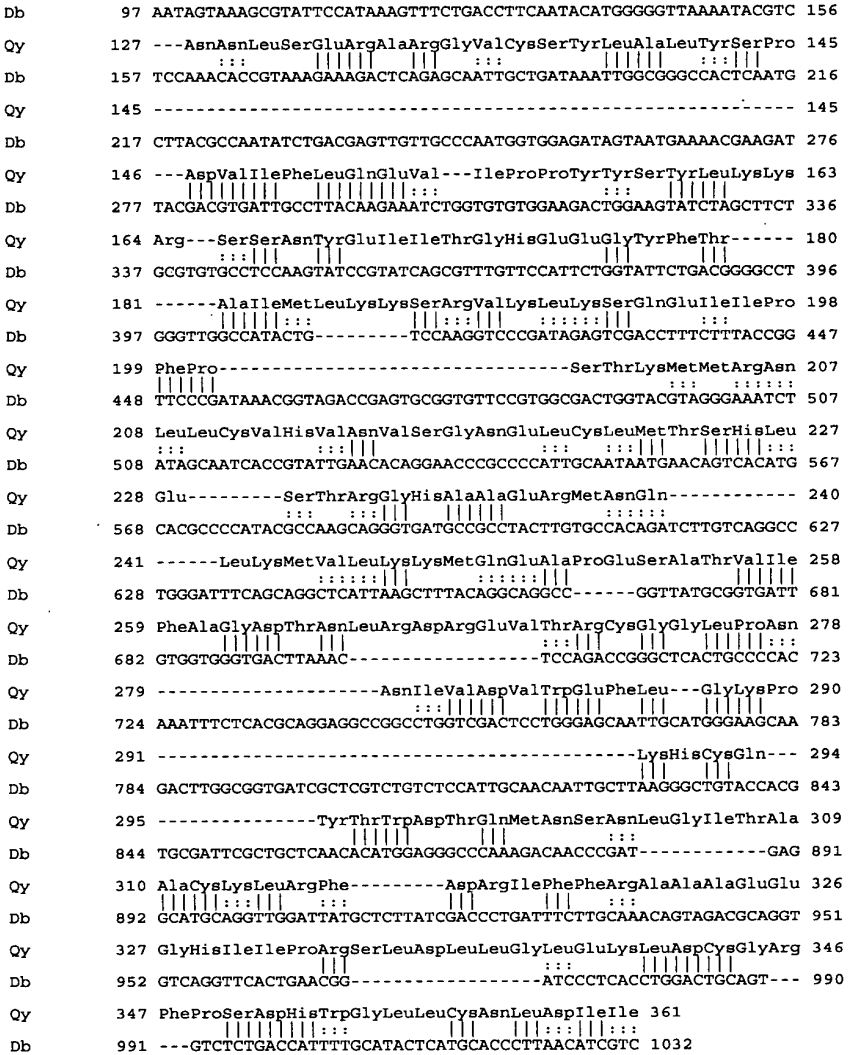
RESULT 48
US-11-264-737-152
; Sequence 152, Application US/11264737
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours & Co., Inc.
; APPLICANT: Damude, Howard Glenn
; APPLICANT: Gillies, Peter J.
; APPLICANT: Macool, Daniel Joseph
; APPLICANT: Pictaggio, Stephen K.
; APPLICANT: Ragghianti, James John
; APPLICANT: Seip, John E.
; APPLICANT: Xue, Zhixiong
; APPLICANT: Yadav, Narendra S.
; APPLICANT: Zhang, Hongxiang
; APPLICANT: Zhu, Quinn Qun
; TITLE OF INVENTION: DOCOSAHEXAENOIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA
; FILE REFERENCE: CL3160 US NA
; CURRENT APPLICATION NUMBER: US/11/264,737
; CURRENT FILING DATE: 2005-11-01
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 152
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-11-264-737-152

Alignment Scores:

Pred. No.:	0.0432	Length:	1434
Score:	108.00	Matches:	78
Percent Similarity:	35.2%	Conservative:	47
Best Local Similarity:	22.0%	Mismatches:	124
Query Match:	5.7%	Indels:	106
DB:	9	Gaps:	20

US-10-757-745-2 (1-362) x US-11-264-737-152 (1-1434)





RESULT 49
US-11-265-761-124
; Sequence 124, Application US/11265761
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Co., Inc.
; APPLICANT: Damude, Howard Glenn
; APPLICANT: Gillies, Peter John

```
;  APPLICANT:  Macool, Daniel Joseph
;  APPLICANT:  Picataggio, Stephen K.
;  APPLICANT:  Pollak, Dana M. Walters
;  APPLICANT:  Ragghianti, James John
;  APPLICANT:  Xue, Zhixiong
;  APPLICANT:  Yadav, Narendra S.
;  APPLICANT:  Zhang, Hongxiang
;  APPLICANT:  Zhu, Quinn
;  TITLE OF INVENTION: HIGH EICOSAPENTAENOIC ACID PRODUCING STRAINS OF YARROWIA
;  TITLE OF INVENTION:  LIPOLYTICA
;  FILE REFERENCE: CL2698 USNA
;  CURRENT APPLICATION NUMBER: US/11/265,761
;  CURRENT FILING DATE:  2005-11-02
;  PRIOR APPLICATION NUMBER: US 60/624812
;  PRIOR FILING DATE: 2004-11-04
;  NUMBER OF SEQ ID NOS: 414
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 124
;   LENGTH: 1434
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
US-11-265-761-124

Alignment Scores:
Pred. No.:                0.0432          Length:          1434
Score:                    108.00          Matches:         78
Percent Similarity:       35.2%           Conservative:    47
Best Local Similarity:    22.0%           Mismatches:      124
Query Match:              5.7%            Indels:          106
DB:                       9               Gaps:            20

US-10-757-745-2 (1-362) x US-11-265-761-124 (1-1434)

Qy         90 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 109
              :::|||          |||            ::::::   :::             :::
Db         37 AAGGAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCAATTATGTCCGAT 96

Qy        110 AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGlyLeuAspLeu--- 126
              |||        :::|||:::      :::|||:::|||      ||||||
Db         97 AATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGGTTAAAATACGTC 156

Qy        127 ---AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
                 :::      ||||||   |||  :::         ||||||   :::|||
Db        157 TCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCGGGCCACTCAATG 216

Qy        145 ------------------------------------------------------------ 145

Db        217 CTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAATGAAAACGAAGAT 276

Qy        146 ---AspValIlePheLeuGlnGluVal---IleProProTyrTyrSerTyrLeuLysLys 163
                 |||||||||   |||||||||:::            :::   ||||||
Db        277 TACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAGTATCTAGCTTCT 336

Qy        164 Arg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr------ 180
                 :::|||   |||                        |||      |||
Db        337 GCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATTCTGACGGGGCCT 396

Qy        181 ------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198
                    ||||||:::         |||:::|||   ::::::|||      :::
Db        397 GGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACCTTTCTTTACCGG 447

Qy        199 PhePro--------------------------------SerThrLysMetMetArgAsn 207
                 ||||||                                        :::   ::::::
Db        448 TTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTACGTAGGGAAATCT 507

Qy        208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
              :::            :::|||            :::   :::|||   ||||||:::
Db        508 ATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATGAACAGTCACATG 567

Qy        228 Glu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln------------ 240
                       :::   :::|||   ||||||          ::::::
Db        568 CACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGATCTTGTCAGGCC 627

Qy        241 ------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258
                    ::::::|||      ::::::|||                      ||||||
Db        628 TGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGTTATGCGGTGATT 681

Qy        259 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 278
              ||||||  |||                    :::|||   |||   ||||||:::
Db        682 GTGGTGGGTGACTTAAAC----------------TCCAGACCGGGCTCACTGCCCCAC 723

Qy        279 --------------------AsnIleValAspValTrpGluPheLeu---GlyLysPro 290
                                  :::||||||   ||||||   |||   ||||||
Db        724 AAATTTCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTGCATGGGAAGCAA 783

Qy        291 ------------------------------------------------LysHisCysGln--- 294
                                                             |||   |||
Db        784 GACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTTAAGGGCTGTACCACG 843

Qy        295 --------------TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 309
                            ||||||     |||          :::
Db        844 TGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAACCCGAT------------GAG 891

Qy        310 AlaCysLysLeuArgPhe---------AspArgIlePhePheArgAlaAlaAlaGluGlu 326
              ||||||:::|||  :::          |||     |||  :::
Db        892 GCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACAGTAGACGCAGGT 951

Qy        327 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 346
                       |||                  :::      |||||||||
Db        952 GTCAGGTTCACTGAACGG----------------ATCCCTCACCTGGACTGCAGT--- 990

Qy        347 PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
                    |||||||||:::          |||   |||:::|||:::
Db        991 ---GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 1032

RESULT 50
US-11-264-784-111
; Sequence 111, Application US/11264784
; GENERAL INFORMATION:
;  APPLICANT: E.I. duPont de Nemours & Co., Inc.
;  APPLICANT:  Damude, Howard Glenn
;  APPLICANT:  Gillies, Peter John
;  APPLICANT:  Macool, Daniel Joseph
;  APPLICANT:  Picataggio, Stephen K.
;  APPLICANT:  Pollak, Dana M. Walters
;  APPLICANT:  Ragghianti, James John
;  APPLICANT:  Xue, Zhixiong
;  APPLICANT:  Yadav, Narendra S.
;  APPLICANT:  Zhang, Hongxiang
;  APPLICANT:  Zhu, Quinn
;  TITLE OF INVENTION: HIGH ARACHIDONIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA
;  FILE REFERENCE: CL3136 USNA
;  CURRENT APPLICATION NUMBER: US/11/264,784
;  CURRENT FILING DATE:  2005-11-01
;  NUMBER OF SEQ ID NOS: 375
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 111
;   LENGTH: 1434
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
US-11-264-784-111

Alignment Scores:
Pred. No.:                0.0432          Length:          1434
Score:                    108.00          Matches:         78
Percent Similarity:       35.2%           Conservative:    47
Best Local Similarity:    22.0%           Mismatches:      124
Query Match:              5.7%            Indels:          106
DB:                       9               Gaps:            20

US-10-757-745-2 (1-362) x US-11-264-784-111 (1-1434)

Qy         90 GluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlu 109
              :::|||          |||            ::::::   :::             :::
Db         37 AAGGAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCAATTATGTCCGAT 96
```

```
Qy        110 AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGlyLeuAspLeu--- 126
Db         97 AATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGGTTAAAATACGTC 156

Qy        127 ---AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 145
Db        157 TCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCGGGCCACTCAATG 216

Qy        145 ------------------------------------------------------------ 145
Db        217 CTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAATGAAAACGAAGAT 276

Qy        146 ---AspValIlePheLeuGlnGluVal---IleProProTyrTyrSerTyrLeuLysLys 163
Db        277 TACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAGTATCTAGCTTCT 336

Qy        164 Arg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr------ 180
Db        337 GCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATTCTGACGGGGCCT 396

Qy        181 ------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 198
Db        397 GGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACCTTTCTTTACCGG 447

Qy        199 PhePro--------------------------------SerThrLysMetMetArgAsn 207
Db        448 TTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTACGTAGGGAAATCT 507

Qy        208 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
Db        508 ATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATGAACAGTCACATG 567

Qy        228 Glu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln------------ 240
Db        568 CACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGATCTTGTCAGGCC 627

Qy        241 ------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 258
Db        628 TGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGTTATGCGGTGATT 681

Qy        259 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 278
Db        682 GTGGTGGGTGACTTAAAC-----------------TCCAGACCGGGCTCACTGCCCCAC 723

Qy        279 ---------------------AsnIleValAspValTrpGluPheLeu---GlyLysPro 290
Db        724 AAATTTCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTGCATGGGAAGCAA 783

Qy        291 ----------------------------------------------LysHisCysGln--- 294
Db        784 GACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTTAAGGGCTGTACCACG 843

Qy        295 --------------TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 309
Db        844 TGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAACCCGAT------------GAG 891

Qy        310 AlaCysLysLeuArgPhe---------AspArgIlePhePheArgAlaAlaAlaGluGlu 326
Db        892 GCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACAGTAGACGCAGGT 951

Qy        327 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 346
Db        952 GTCAGGTTCACTGAACGG-----------------ATCCCTCACCTGGACTGCAGT--- 990

Qy        347 PheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
Db        991 ---GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 1032
```

RESULT 51
US-10-868-184C-10381/c
; Sequence 10381, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10381
; LENGTH: 6044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-868-184C-10381

Alignment Scores:

Pred. No.:	0.418	Length:	6044
Score:	108.00	Matches:	56
Percent Similarity:	44.0%	Conservative:	43
Best Local Similarity:	24.9%	Mismatches:	92
Query Match:	5.7%	Indels:	35
DB:	6	Gaps:	9

US-10-757-745-2 (1-362) x US-10-868-184C-10381 (1-6044)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db       4101 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4043

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db       4042 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 3989

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db       3988 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3932

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db       3931 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3890

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db       3889 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3830

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db       3829 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3770

Qy        211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
Db       3769 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3710

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
```

```
Db       3709 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 3665

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||          :::|||   |||:::|||      ||||||   |||
Db       3664 GACCTACAAAGAGTCTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3608

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db       3607 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3548

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db       3547 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3491

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db       3490 ACAGAATATACATTT 3476

RESULT 52
US-10-868-184D-10381/c
; Sequence 10381, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10381
;   LENGTH: 6044
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10381

Alignment Scores:
Pred. No.:              0.418          Length:         6044
Score:                  108.00         Matches:        56
Percent Similarity:     44.0%          Conservative:   43
Best Local Similarity:  24.9%          Mismatches:     92
Query Match:            5.7%           Indels:         35
DB:                     6              Gaps:           9

US-10-757-745-2 (1-362) x US-10-868-184D-10381 (1-6044)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||  ||||||    :::   ||||||:::                  :::
Db       4101 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4043

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||      :::::::::|||   ||||||:::|||      ||||||:::
Db       4042 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 3989

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||  :::   |||::::::              |||   |||      :::
Db       3988 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3932

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                 ::::::|||      |||   ::::::
Db       3931 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3890

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                |||||:::   |||
Db       3889 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3830

Qy        191 LeuLysSerGlnGluIleIleProProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::       :::       |||      ||||||:::   |||:::
Db       3829 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3770

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::             ||||||   ::::::      :::
Db       3769 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3710

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||          |||   |||          :::|||   ||||||
Db       3709 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 3665

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||          :::|||   |||:::|||      ||||||   |||
Db       3664 GACCTACAAAGAGTCTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3608

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db       3607 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3548

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db       3547 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3491

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db       3490 ACAGAATATACATTT 3476

RESULT 53
PCT-US05-40306-155
; Sequence 155, Application PC/TUS0540306
; GENERAL INFORMATION:
;  APPLICANT: E.I. duPont de Nemours & Co., Inc.
;  TITLE OF INVENTION: HIGH ARACHIDONIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA
;  FILE REFERENCE: CL3136 PCT
;  CURRENT APPLICATION NUMBER: PCT/US05/40306
;  CURRENT FILING DATE:  2005-11-14
;  NUMBER OF SEQ ID NOS: 375
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 155
;   LENGTH: 8179
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Plasmid pTEF::ISC1
PCT-US05-40306-155

Alignment Scores:
Pred. No.:              0.674          Length:         8179
Score:                  108.00         Matches:        88
Percent Similarity:     34.8%          Conservative:   44
Best Local Similarity:  23.2%          Mismatches:     124
Query Match:            5.7%           Indels:         124
DB:                     1              Gaps:           22

US-10-757-745-2 (1-362) x PCT-US05-40306-155 (1-8179)

Qy         79 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 98
              ||||||   |||  :::       |||   |||:::|||   |||
Db       6723 TCTGAGTATAAGAATCATTCACCATGGACAACAGAAAAGACAGA-GATGTTCACGAGAGG 6781
```

```
Qy         99 LysIleSerProSerGluAspThrGlnGlnGlu-------------------------- 109
Db       6782 AAG------------GAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCA 6829

Qy        110 ------------AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGly 123
Db       6830 ATTATGTCCGATAATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGG 6889

Qy        124 LeuAspLeu------AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 141
Db       6890 TTAAAATACGTCTCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCG 6949

Qy        142 LeuTyrSerPro------------------------------------------------ 145
Db       6950 GGCCACTCAATGCTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAAT 7009

Qy        146 --------------AspValIlePheLeuGlnGluVal---IleProProTyrTyrSer 159
Db       7010 GAAAACGAAGATTACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAG 7069

Qy        160 TyrLeuLysLysArg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178
Db       7070 TATCTAGCTTCTGCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATT 7129

Qy        179 PheThr------------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194
Db       7130 CTGACGGGGCCTGGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACC 7180

Qy        195 GluIleIleProPhePro---------------------------------SerThrLys 203
Db       7181 TTTCTTTACCGGTTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTAC 7240

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMet 223
Db       7241 GTAGGGAAATCTATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATG 7300

Qy        224 ThrSerHisLeuGlu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
Db       7301 AACAGTCACATGCACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGA 7360

Qy        241 ------------------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
Db       7361 TCTTGTCAGGCCTGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGT 7414

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
Db       7415 TATGCGGTGATTGTGGTGGGTGACTTAAAC-----------------TCCAGACCGGGC 7456

Qy        275 GlyLeuProAsn---------------------AsnIleValAspValTrpGluPheLeu 287
Db       7457 TCACTGCCCCACAAATTTCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTG 7516

Qy        288 ---GlyLysPro-------------------------------------------------Lys 291
Db       7517 CATGGGAAGCAAGACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTTAAG 7576

Qy        292 HisCysGln------------------TyrThrTrpAspThrGlnMetAsnSerAsnLeu 305
Db       7577 GGCTGTACCACGTGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAACCCGAT--- 7633

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPhe---------AspArgIlePhePheArgAla 322
Db       7634 ---------GAGGCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACA 7684

Qy        323 AlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 342
Db       7685 GTAGACGCAGGTGTCAGGTTCACTGAACGG-----------------ATCCCTCACCTG 7726

Qy        343 AspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
Db       7727 GACTGCAGT------GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 7777
```

```
RESULT 54
PCT-US05-40255-170
; Sequence 170, Application PC/TUS0540255
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: HIGH EICOSAPENTAENOIC ACID PRODUCING STRAINS OF YARROWIA
;  TITLE OF INVENTION:  LIPOLYTICA
;  FILE REFERENCE: CL2698 PCT
;  CURRENT APPLICATION NUMBER: PCT/US05/40255
;  CURRENT FILING DATE:  2005-11-15
;  PRIOR APPLICATION NUMBER: US 60/624812
;  PRIOR FILING DATE: 2004-11-04
;  NUMBER OF SEQ ID NOS: 414
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 170
;   LENGTH: 8179
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Plasmid pTEF::ISC1
PCT-US05-40255-170

Alignment Scores:
Pred. No.:              0.674          Length:          8179
Score:                  108.00         Matches:         88
Percent Similarity:     34.8%          Conservative:    44
Best Local Similarity:  23.2%          Mismatches:      124
Query Match:            5.7%           Indels:          124
DB:                     1              Gaps:            22

US-10-757-745-2 (1-362) x PCT-US05-40255-170 (1-8179)

Qy         79 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 98
Db       6723 TCTGAGTATAAGAATCATTCACCATGGACAACAGAAAAGACAGA-GATGTTCACGAGAGG 6781

Qy         99 LysIleSerProSerGluAspThrGlnGlnGlu-------------------------- 109
Db       6782 AAG------------GAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCA 6829

Qy        110 ------------AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGly 123
Db       6830 ATTATGTCCGATAATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGG 6889

Qy        124 LeuAspLeu------AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 141
Db       6890 TTAAAATACGTCTCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCG 6949

Qy        142 LeuTyrSerPro------------------------------------------------ 145
Db       6950 GGCCACTCAATGCTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAAT 7009

Qy        146 --------------AspValIlePheLeuGlnGluVal---IleProProTyrTyrSer 159
Db       7010 GAAAACGAAGATTACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAG 7069

Qy        160 TyrLeuLysLysArg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178
Db       7070 TATCTAGCTTCTGCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATT 7129

Qy        179 PheThr------------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194
Db       7130 CTGACGGGGCCTGGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACC 7180

Qy        195 GluIleIleProPhePro---------------------------------SerThrLys 203
Db       7181 TTTCTTTACCGGTTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTAC 7240

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMet 223
Db       7241 GTAGGGAAATCTATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATG 7300

Qy        224 ThrSerHisLeuGlu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
```

```
Db       7301 AACAGTCACATGCACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGA 7360

Qy        241 ------------------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254

Db       7361 TCTTGTCAGGCCTGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGT 7414

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274

Db       7415 TATGCGGTGATTGTGGTGGGTGACTTAAAC-----------------TCCAGACCGGGC 7456

Qy        275 GlyLeuProAsn--------------------AsnIleValAspValTrpGluPheLeu 287

Db       7457 TCACTGCCCCACAAATTTCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTG 7516

Qy        288 ---GlyLysPro-------------------------------------------------Lys 291

Db       7517 CATGGGAAGCAAGACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTTAAG 7576

Qy        292 HisCysGln------------------TyrThrTrpAspThrGlnMetAsnSerAsnLeu 305

Db       7577 GGCTGTACCACGTGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAACCCGAT--- 7633

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPhe---------AspArgIlePhePheArgAla 322

Db       7634 ---------GAGGCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACA 7684

Qy        323 AlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 342

Db       7685 GTAGACGCAGGTGTCAGGTTCACTGAACGG-----------------ATCCCTCACCTG 7726

Qy        343 AspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361

Db       7727 GACTGCAGT------GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 7777
```

RESULT 55
PCT-US05-40256-207
; Sequence 207, Application PC/TUS0540256
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: DOCOSAHEXAENOIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA
; FILE REFERENCE: CL3160 PCT
; CURRENT APPLICATION NUMBER: PCT/US05/40256
; CURRENT FILING DATE: 2005-11-15
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 207
; LENGTH: 8179
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pTEF::ISC1
PCT-US05-40256-207

Alignment Scores:

Pred. No.:	0.674	Length:	8179
Score:	108.00	Matches:	88
Percent Similarity:	34.8%	Conservative:	44
Best Local Similarity:	23.2%	Mismatches:	124
Query Match:	5.7%	Indels:	124
DB:	1	Gaps:	22

US-10-757-745-2 (1-362) x PCT-US05-40256-207 (1-8179)

```
Qy         79 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 98

Db       6723 TCTGAGTATAAGAATCATTCACCATGGACAACAGAAAAGACAGA-GATGTTCACGAGAGG 6781

Qy         99 LysIleSerProSerGluAspThrGlnGlnGlu-------------------------- 109

Db       6782 AAG-----------GAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCA 6829

Qy        110 -----------AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGly 123

Db       6830 ATTATGTCCGATAATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGG 6889

Qy        124 LeuAspLeu------AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 141

Db       6890 TTAAAATACGTCTCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCG 6949

Qy        142 LeuTyrSerPro------------------------------------------------ 145

Db       6950 GGCCACTCAATGCTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAAT 7009

Qy        146 --------------AspValIlePheLeuGlnGluVal---IleProProTyrTyrSer 159

Db       7010 GAAAACGAAGATTACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAG 7069

Qy        160 TyrLeuLysLysArg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178

Db       7070 TATCTAGCTTCTGCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATT 7129

Qy        179 PheThr------------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194

Db       7130 CTGACGGGGCCTGGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACC 7180

Qy        195 GluIleIleProPhePro---------------------------------SerThrLys 203

Db       7181 TTTCTTTACCGGTTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTAC 7240

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMet 223

Db       7241 GTAGGGAAATCTATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATG 7300

Qy        224 ThrSerHisLeuGlu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln 240

Db       7301 AACAGTCACATGCACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGA 7360

Qy        241 ------------------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254

Db       7361 TCTTGTCAGGCCTGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGT 7414

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274

Db       7415 TATGCGGTGATTGTGGTGGGTGACTTAAAC-----------------TCCAGACCGGGC 7456

Qy        275 GlyLeuProAsn--------------------AsnIleValAspValTrpGluPheLeu 287

Db       7457 TCACTGCCCCACAAATTTCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTG 7516

Qy        288 ---GlyLysPro-------------------------------------------------Lys 291

Db       7517 CATGGGAAGCAAGACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTTAAG 7576

Qy        292 HisCysGln------------------TyrThrTrpAspThrGlnMetAsnSerAsnLeu 305

Db       7577 GGCTGTACCACGTGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAACCCGAT--- 7633

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPhe---------AspArgIlePhePheArgAla 322

Db       7634 ---------GAGGCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACA 7684

Qy        323 AlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 342

Db       7685 GTAGACGCAGGTGTCAGGTTCACTGAACGG-----------------ATCCCTCACCTG 7726

Qy        343 AspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361

Db       7727 GACTGCAGT------GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 7777
```

RESULT 56
US-11-264-737-207
; Sequence 207, Application US/11264737
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours & Co., Inc.
; APPLICANT: Damude, Howard Glenn

```
;  APPLICANT:  Gillies, Peter J.
;  APPLICANT:  Macool, Daniel Joseph
;  APPLICANT:  Pictaggio, Stephen K.
;  APPLICANT:  Ragghianti, James John
;  APPLICANT:  Seip, John E.
;  APPLICANT:  Xue, Zhixiong
;  APPLICANT:  Yadav, Narendra S.
;  APPLICANT:  Zhang, Hongxiang
;  APPLICANT:  Zhu, Quinn Qun
;  TITLE OF INVENTION: DOCOSAHEXAENOIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA
;  FILE REFERENCE: CL3160 US NA
;  CURRENT APPLICATION NUMBER: US/11/264,737
;  CURRENT FILING DATE:  2005-11-01
;  NUMBER OF SEQ ID NOS: 467
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 207
;   LENGTH: 8179
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Plasmid pTEF::ISC1
US-11-264-737-207

Alignment Scores:
Pred. No.:                0.674          Length:          8179
Score:                    108.00         Matches:         88
Percent Similarity:       34.8%          Conservative:    44
Best Local Similarity:    23.2%          Mismatches:      124
Query Match:              5.7%           Indels:          124
DB:                       9              Gaps:            22

US-10-757-745-2 (1-362) x US-11-264-737-207 (1-8179)

Qy         79 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 98
              ||||||   |||   :::         |||   |||:::|||   |||
Db       6723 TCTGAGTATAAGAATCATTCACCATGGACAACAGAAAAGACAGA-GATGTTCACGAGAGG 6781

Qy         99 LysIleSerProSerGluAspThrGlnGlnGlu-------------------------- 109
              |||           ||||||   |||   |||
Db       6782 AAG------------GAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCA 6829

Qy        110 ------------AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGly 123
                         |||         :::|||:::      :::|||:::|||      |||
Db       6830 ATTATGTCCGATAATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGG 6889

Qy        124 LeuAspLeu------AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 141
              |||           :::      ||||||   |||   :::         ||||||
Db       6890 TTAAAATACGTCTCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCG 6949

Qy        142 LeuTyrSerPro------------------------------------------------ 145
                 :::|||
Db       6950 GGCCACTCAATGCTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAAT 7009

Qy        146 ---------------AspValIlePheLeuGlnGluVal---IleProProTyrTyrSer 159
                            |||||||||   |||||||||:::               :::
Db       7010 GAAAACGAAGATTACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAG 7069

Qy        160 TyrLeuLysLysArg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178
              ||||||          :::|||   |||                          |||
Db       7070 TATCTAGCTTCTGCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATT 7129

Qy        179 PheThr------------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194
                 |||             ||||||:::         |||:::|||   :::::::|||
Db       7130 CTGACGGGGCCTGGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACC 7180

Qy        195 GluIleIleProPhePro--------------------------------SerThrLys 203
                 :::      ||||||
Db       7181 TTTCTTTACCGGTTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTAC 7240

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMet 223
              :::   :::::::::           :::|||           :::   :::|||
Db       7241 GTAGGGAAATCTATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATG 7300

Qy        224 ThrSerHisLeuGlu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
                 ||||||:::            :::   :::|||   ||||||         ::::::
Db       7301 AACAGTCACATGCACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGA 7360

Qy        241 ------------------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
                               :::::::|||         :::::::|||
Db       7361 TCTTGTCAGGCCTGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGT 7414

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
                    ||||||       ||||||   |||                  :::|||   |||
Db       7415 TATGCGGTGATTGTGGTGGGTGACTTAAAC-----------------TCCAGACCGGGC 7456

Qy        275 GlyLeuProAsn--------------------AsnIleValAspValTrpGluPheLeu 287
                 ||||||:::                  :::||||||   ||||||   |||
Db       7457 TCACTGCCCCACAAATTTCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTG 7516

Qy        288 ---GlyLysPro-------------------------------------------------Lys 291
                 |||||                                                   |||
Db       7517 CATGGGAAGCAAGACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTTAAG 7576

Qy        292 HisCysGln------------------TyrThrTrpAspThrGlnMetAsnSerAsnLeu 305
                 |||                     ||||||   |||          :::
Db       7577 GGCTGTACCACGTGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAACCCGAT--- 7633

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPhe---------AspArgIlePhePheArgAla 322
                       ||||||:::|||   :::         |||       |||   :::
Db       7634 ---------GAGGCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACA 7684

Qy        323 AlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 342
                                     |||                 :::         |||
Db       7685 GTAGACGCAGGTGTCAGGTTCACTGAACGG-----------------ATCCCTCACCTG 7726

Qy        343 AspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
              ||||||           |||||||||:::         |||   |||:::|||:::
Db       7727 GACTGCAGT------GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 7777

RESULT 57
US-11-265-761-170
; Sequence 170, Application US/11265761
; GENERAL INFORMATION:
;  APPLICANT: E.I. duPont de Nemours and Co., Inc.
;  APPLICANT:  Damude, Howard Glenn
;  APPLICANT:  Gillies, Peter John
;  APPLICANT:  Macool, Daniel Joseph
;  APPLICANT:  Picataggio, Stephen K.
;  APPLICANT:  Pollak, Dana M. Walters
;  APPLICANT:  Ragghianti, James John
;  APPLICANT:  Xue, Zhixiong
;  APPLICANT:  Yadav, Narendra S.
;  APPLICANT:  Zhang, Hongxiang
;  APPLICANT:  Zhu, Quinn
;  TITLE OF INVENTION: HIGH EICOSAPENTAENOIC ACID PRODUCING STRAINS OF YARROWIA
;  TITLE OF INVENTION:  LIPOLYTICA
;  FILE REFERENCE: CL2698 USNA
;  CURRENT APPLICATION NUMBER: US/11/265,761
;  CURRENT FILING DATE:  2005-11-02
;  PRIOR APPLICATION NUMBER: US 60/624812
;  PRIOR FILING DATE: 2004-11-04
;  NUMBER OF SEQ ID NOS: 414
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 170
;   LENGTH: 8179
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Plasmid pTEF::ISC1
US-11-265-761-170

Alignment Scores:
Pred. No.:                0.674          Length:          8179
Score:                    108.00         Matches:         88
Percent Similarity:       34.8%          Conservative:    44
Best Local Similarity:    23.2%          Mismatches:      124
```

```
Query Match:          5.7%      Indels:       124
DB:                   9         Gaps:         22

US-10-757-745-2 (1-362) x US-11-265-761-170 (1-8179)

Qy         79 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 98
Db       6723 TCTGAGTATAAGAATCATTCACCATGGACAACAGAAAAGACAGA-GATGTTCACGAGAGG 6781

Qy         99 LysIleSerProSerGluAspThrGlnGlnGlu------------------------- 109
Db       6782 AAG------------GAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCA 6829

Qy        110 ------------AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGly 123
Db       6830 ATTATGTCCGATAATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGG 6889

Qy        124 LeuAspLeu------AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 141
Db       6890 TTAAAATACGTCTCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCG 6949

Qy        142 LeuTyrSerPro------------------------------------------------ 145
Db       6950 GGCCACTCAATGCTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAAT 7009

Qy        146 --------------AspValIlePheLeuGlnGluVal---IleProProTyrTyrSer 159
Db       7010 GAAAACGAAGATTACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAG 7069

Qy        160 TyrLeuLysLysArg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178
Db       7070 TATCTAGCTTCTGCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATT 7129

Qy        179 PheThr------------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194
Db       7130 CTGACGGGGCCTGGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACC 7180

Qy        195 GluIleIleProPhePro--------------------------------SerThrLys 203
Db       7181 TTTCTTTACCGGTTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTAC 7240

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMet 223
Db       7241 GTAGGGAAATCTATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATG 7300

Qy        224 ThrSerHisLeuGlu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
Db       7301 AACAGTCACATGCACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGA 7360

Qy        241 -----------------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
Db       7361 TCTTGTCAGGCCTGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGT 7414

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
Db       7415 TATGCGGTGATTGTGGTGGGTGACTTAAAC-----------------TCCAGACCGGGC 7456

Qy        275 GlyLeuProAsn--------------------AsnIleValAspValTrpGluPheLeu 287
Db       7457 TCACTGCCCCACAAATTTCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTG 7516

Qy        288 ---GlyLysPro-------------------------------------------------Lys 291
Db       7517 CATGGGAAGCAAGACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTTAAG 7576

Qy        292 HisCysGln-----------------TyrThrTrpAspThrGlnMetAsnSerAsnLeu 305
Db       7577 GGCTGTACCACGTGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAACCCGAT--- 7633

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPhe---------AspArgIlePhePheArgAla 322
Db       7634 ---------GAGGCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACA 7684

Qy        323 AlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 342
Db       7685 GTAGACGCAGGTGTCAGGTTCACTGAACGG-----------------ATCCCTCACCTG 7726

Qy        343 AspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
Db       7727 GACTGCAGT------GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 7777

RESULT 58
US-11-264-784-155
; Sequence 155, Application US/11264784
; GENERAL INFORMATION:
;  APPLICANT: E.I. duPont de Nemours & Co., Inc.
;  APPLICANT:  Damude, Howard Glenn
;  APPLICANT:  Gillies, Peter John
;  APPLICANT:  Macool, Daniel Joseph
;  APPLICANT:  Picataggio, Stephen K.
;  APPLICANT:  Pollak, Dana M. Walters
;  APPLICANT:  Ragghianti, James John
;  APPLICANT:  Xue, Zhixiong
;  APPLICANT:  Yadav, Narendra S.
;  APPLICANT:  Zhang, Hongxiang
;  APPLICANT:  Zhu, Quinn
;  TITLE OF INVENTION: HIGH ARACHIDONIC ACID PRODUCING STRAINS OF YARROWIA LIPOLYTICA
;  FILE REFERENCE: CL3136 USNA
;  CURRENT APPLICATION NUMBER: US/11/264,784
;  CURRENT FILING DATE:  2005-11-01
;  NUMBER OF SEQ ID NOS: 375
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 155
;   LENGTH: 8179
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Plasmid pTEF::ISC1
US-11-264-784-155

Alignment Scores:
Pred. No.:            0.674     Length:       8179
Score:                108.00    Matches:      88
Percent Similarity:   34.8%     Conservative: 44
Best Local Similarity: 23.2%    Mismatches:   124
Query Match:          5.7%      Indels:       124
DB:                   9         Gaps:         22

US-10-757-745-2 (1-362) x US-11-264-784-155 (1-8179)

Qy         79 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 98
Db       6723 TCTGAGTATAAGAATCATTCACCATGGACAACAGAAAAGACAGA-GATGTTCACGAGAGG 6781

Qy         99 LysIleSerProSerGluAspThrGlnGlnGlu------------------------- 109
Db       6782 AAG------------GAAGATGGTCAATCTGAGTTTGAAGCACTGAACGGGACCAACGCA 6829

Qy        110 ------------AsnGlySerMetPheSerLeu------IleThrTrpAsnIleAspGly 123
Db       6830 ATTATGTCCGATAATAGTAAAGCGTATTCCATAAAGTTTCTGACCTTCAATACATGGGGG 6889

Qy        124 LeuAspLeu------AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 141
Db       6890 TTAAAATACGTCTCCAAACACCGTAAAGAAAGACTCAGAGCAATTGCTGATAAATTGGCG 6949

Qy        142 LeuTyrSerPro------------------------------------------------ 145
Db       6950 GGCCACTCAATGCTTACGCCAATATCTGACGAGTTGTTGCCCAATGGTGGAGATAGTAAT 7009

Qy        146 --------------AspValIlePheLeuGlnGluVal---IleProProTyrTyrSer 159
Db       7010 GAAAACGAAGATTACGACGTGATTGCCTTACAAGAAATCTGGTGTGTGGAAGACTGGAAG 7069

Qy        160 TyrLeuLysLysArg---SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 178
```

```
Db       7070 TATCTAGCTTCTGCGTGTGCCTCCAAGTATCCGTATCAGCGTTTGTTCCATTCTGGTATT 7129

Qy        179 PheThr------------AlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 194
Db       7130 CTGACGGGGCCTGGGTTGGCCATACTG---------TCCAAGGTCCCGATAGAGTCGACC 7180

Qy        195 GluIleIleProPhePro--------------------------------SerThrLys 203
Db       7181 TTTCTTTACCGGTTCCCGATAAACGGTAGACCGAGTGCGGTGTTCCGTGGCGACTGGTAC 7240

Qy        204 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMet 223
Db       7241 GTAGGGAAATCTATAGCAATCACCGTATTGAACACAGGAACCCGCCCCATTGCAATAATG 7300

Qy        224 ThrSerHisLeuGlu---------SerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
Db       7301 AACAGTCACATGCACGCCCCATACGCCAAGCAGGGTGATGCCGCCTACTTGTGCCACAGA 7360

Qy        241 ------------------LeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 254
Db       7361 TCTTGTCAGGCCTGGGATTTCAGCAGGCTCATTAAGCTTTACAGGCAGGCC------GGT 7414

Qy        255 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 274
Db       7415 TATGCGGTGATTGTGGTGGGTGACTTAAAC------------------TCCAGACCGGGC 7456

Qy        275 GlyLeuProAsn--------------------AsnIleValAspValTrpGluPheLeu 287
Db       7457 TCACTGCCCCACAAATTTCTCACGCAGGAGGCCGGCCTGGTCGACTCCTGGGAGCAATTG 7516

Qy        288 ---GlyLysPro-----------------------------------------------Lys 291
Db       7517 CATGGGAAGCAAGACTTGGCGGTGATCGCTCGTCTGTCTCCATTGCAACAATTGCTTAAG 7576

Qy        292 HisCysGln------------------TyrThrTrpAspThrGlnMetAsnSerAsnLeu 305
Db       7577 GGCTGTACCACGTGCGATTCGCTGCTCAACACATGGAGGGCCCAAAGACAACCCGAT--- 7633

Qy        306 GlyIleThrAlaAlaCysLysLeuArgPhe---------AspArgIlePhePheArgAla 322
Db       7634 ---------GAGGCATGCAGGTTGGATTATGCTCTTATCGACCCTGATTTCTTGCAAACA 7684

Qy        323 AlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuGluGlyLeuGluLysLeu 342
Db       7685 GTAGACGCAGGTGTCAGGTTCACTGAACGG-----------------ATCCCTCACCTG 7726

Qy        343 AspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIleIle 361
Db       7727 GACTGCAGT------GTCTCTGACCATTTTGCATACTCATGCACCCTTAACATCGTC 7777
```

RESULT 59
US-10-868-184C-10688/c
; Sequence 10688, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10688
; LENGTH: 14411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-868-184C-10688

Alignment Scores:

Pred. No.:	1.65	Length:	14411
Score:	108.00	Matches:	56
Percent Similarity:	44.0%	Conservative:	43
Best Local Similarity:	24.9%	Mismatches:	92
Query Match:	5.7%	Indels:	35
DB:	6	Gaps:	9

US-10-757-745-2 (1-362) x US-10-868-184C-10688 (1-14411)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
Db       4114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4056

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
Db       4055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4002

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
Db       4001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3945

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
Db       3944 CAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 3903

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
Db       3902 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3843

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
Db       3842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3783

Qy        211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
Db       3782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATGTAT 3723

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
Db       3722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3678

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
Db       3677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3621

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
Db       3620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3561

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
Db       3560 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3504

Qy        293 CysGlnTyrThrTrp 297
Db       3503 GCAGAATATACATTT 3489
```

```
RESULT 60
US-10-868-184D-10688/c
; Sequence 10688, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10688
;   LENGTH: 14411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10688

Alignment Scores:
Pred. No.:                1.65          Length:          14411
Score:                    108.00        Matches:         56
Percent Similarity:       44.0%         Conservative:    43
Best Local Similarity:    24.9%         Mismatches:      92
Query Match:              5.7%          Indels:          35
DB:                       6             Gaps:            9

US-10-757-745-2 (1-362) x US-10-868-184D-10688 (1-14411)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::                 :::
Db        4114 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4056

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||       :::::::::|||   ||||||:::|||       ||||||:::
Db        4055 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4002

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||::::::             |||   |||     :::
Db        4001 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3945

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  ::::::|||      |||   ::::::
Db        3944 CAGGAA----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 3903

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||       :::   |||                  ||||||:::   |||
Db        3902 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3843

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::       :::       |||       ||||||:::   |||:::
Db        3842 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3783

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                      ::::::   |||   |||:::             ||||||   ::::::      :::
Db        3782 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATGTAT 3723

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::|||           |||   |||         :::|||   ||||||
Db        3722 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3678

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                  :::|||       :::|||   |||:::|||       ||||||   |||
Db        3677 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3621

Qy         266 -------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db        3620 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3561

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                           :::::::::|||::::::       |||       ||||||
Db        3560 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3504

Qy         293 CysGlnTyrThrTrp 297
                  :::|||||||:::
Db        3503 GCAGAATATACATTT 3489

RESULT 61
US-11-045-004-1/c
; Sequence 1, Application US/11045004
; GENERAL INFORMATION:
;  APPLICANT: BUCHRIESER, CARMEN
;  APPLICANT:  FRANGEUL, LIONEL
;  APPLICANT:  COUVE, ELISABETH
;  APPLICANT:  RUSNIOK, CHRISTOPHE
;  APPLICANT:  FSIHI, HAFIDA
;  APPLICANT:  DEHOUX, PIERRE
;  APPLICANT:  DUSSURGET, OLIVIER
;  APPLICANT:  CHETOUANI, FARID
;  APPLICANT:  NEDJARI, HAFED
;  APPLICANT:  GLASER, PHILIPPE
;  APPLICANT:  KUNST, FRANCK
;  APPLICANT:  COSSART, PASCALE
;  APPLICANT:  DANIELS, JUSTIN
;  APPLICANT:  GOEBEI, WERNER
;  APPLICANT:  KREFT, JURGEN
;  APPLICANT:  KUHN, MICHAEL
;  APPLICANT:  NG, EVA
;  APPLICANT:  VAZQUEZ-BOLAND, ANTONIO
;  APPLICANT:  DOMINGUEZ-BERNAL, GUSTAVO
;  APPLICANT:  GARRIDO-GARCIA, PATRICIA
;  APPLICANT:  TIERREZ-MARTINEZ, ALBERTO
;  APPLICANT:  AMEND, ALEXANDRA
;  APPLICANT:  CHAKRABORTY, TRINAD
;  APPLICANT:  DOMANN, EUGEN
;  APPLICANT:  HAIN, THORSTEN
;  APPLICANT:  BERCHE, PATRICK
;  APPLICANT:  CHARBIT, ALAIN
;  APPLICANT:  DURANT, LIONEL
;  APPLICANT:  PEREZ-DIAZ, JOSE-CLAUDIO
;  APPLICANT:  BAQUERO, FERNANDO
;  APPLICANT:  GARCIA DEL PORTILLO, FRANCISCO
;  APPLICANT:  GOMEZ-LOPEZ, NURIA
;  APPLICANT:  MADUENIO, ENCARNA
;  APPLICANT:  PABLOS, BETRIZ DE
;  APPLICANT:  WEHLAND, JURGEN
;  APPLICANT:  KARST, UWE
;  APPLICANT:  ENTIAN, KARL-DIETER
;  APPLICANT:  HAUF, JORG
;  APPLICANT:  ROSE, MATTHIAS
;  APPLICANT:  VOSS, HAMUT
;  TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
;  FILE REFERENCE: 05394.0018-02
;  CURRENT APPLICATION NUMBER: US/11/045,004
;  CURRENT FILING DATE:  2005-01-28
;  PRIOR APPLICATION NUMBER: 10/637,657
;  PRIOR FILING DATE: 2003-08-11
```

```
;  PRIOR APPLICATION NUMBER: 10/257,023
;  PRIOR FILING DATE: 2002-10-08
;  PRIOR APPLICATION NUMBER: PCT/FR01/01118
;  PRIOR FILING DATE: 2001-04-11
;  PRIOR APPLICATION NUMBER: FR 00/04,629
;  PRIOR FILING DATE: 2000-04-11
;  NUMBER OF SEQ ID NOS: 2854
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
;   LENGTH: 2944528
;   TYPE: DNA
;   ORGANISM: Listeria monocytogenes
US-11-045-004-1

Alignment Scores:
Pred. No.:              8.37e+03        Length:         2944528
Score:                  107.50          Matches:        71
Percent Similarity:     37.8%           Conservative:   44
Best Local Similarity:  23.4%           Mismatches:     87
Query Match:            5.7%            Indels:         103
DB:                     9               Gaps:           16

US-10-757-745-2 (1-362) x US-11-045-004-1 (1-2944528)

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||   |||||||||:::   |||:::||||||                 ::::::   |||
Db     348434 GGAATCATGTTCAGTGTAACTACTTTTAACATT---------CGTTTTGATGATACTTCT 348384

Qy        131 GluArgAlaArgGly--------------------ValCysSerTyrLeuAlaLeuTyr 143
              ||||||   :::                              |||   |||      |||
Db     348383 GAGCGTAAAAAGAGCTGGGAATTGCGTAAAACACTAACAAAATCATTACTCGATAAATAT 348324

Qy        144 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 163
                  |||   :::   ::::::|||   :::   |||            :::|||:::
Db     348323 CAATGGGATTTCATGGGTGTGGAGGAACCTCTTCTGCCACAAATGCGCGACATGAAAGAA 348264

Qy        164 ArgSerSer---AsnTyrGluIleIle---------ThrGlyHisGluGluGlyTyrPhe 179
                  :::|||   :::             |||   |||:::|||   |||
Db     348263 ATGCAGGATTGGGATTATTTTGGGGTAGGACGAGATGACGGTTTTGAAAAGGGAGAATTT 348204

Qy        180 ThrAlaIleMetLeuLysLysSerArgValLysLeu---------------------- 191
              ||||||:::           :::|||     |||
Db     348203 ACTGCAGTTTTTTACAATTCTACTCGTTTTACCTTGTTGCAAGAAGGTCATTTTTGGCTT 348144

Qy        192 ---LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
                 ::::::   ::::::          ||||||   |||          :::|||
Db     348143 TCCGAAACGCCTGATGTTCCTTCCATTCACTCAACAGCCATGTTTCCGCGA---ATTTGT 348087

Qy        211 Val---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 227
              |||             :::      |||   :::      :::      :::||||||
Db     348086 GTTTGGGGGAAATTCGCCGATTCGGACGGAAAACAGTTCTATATTTTTAACACTCATTTG 348027

Qy        228 GluSerThrArgGlyHisAlaAlaGlu------ArgMetAsnGlnLeuLysMetValLeu 245
              :::           |||   :::|||      |||:::         :::::::::|||
Db     348026 GAT-----------CACATCTCTGAGGAAGCTCGTCTATTTGCAAGCCAACTTTTACTA 347979

Qy        246 LysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu 265
              ||||||          |||::::::   ||||||    ||||||   |||
Db     347978 AAAAAAGCTGCCACCATTGCAGAAAATTCACCCGTGATTATATTAGGCGATTTTAACAC- 347920

Qy        266 Arg--------------------------AspArgGluValThrArgCysGlyGlyLeu 276
              |||                         ::::::::::   ||||||
Db     347919 CGAGCCTGATACACCGACTTATAACTTCATAACAAAAAAATACCAAGATGC--------- 347869

Qy        277 ProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 296

Db     347869 ------------------------------------------------------------ 347869

Qy        297 TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAsp 316
                    :::||||||      |||      ||||||   |||
Db     347868 --------------ACAACTAATCTCTCAAAAACA------TGCAAA---AGGACCCAT 347833

Qy        317 ArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSer-LeuAspLe 336
                  |||   |||   |||                          :::|||   :::       ::
Db     347832 CGGAAGTTTTCACGA----------------------CTTCCGCCCACTCCGCCCTAT 347797

Qy        336 uLeuGlyLeuGluLysLeuAsp-----------------------------Cys----- 344
              :      |||||||||:::|||                               |||
Db     347796 AAATGAGCTAGAAAAAATTGACTACATCTTCGTTTCTGAAGAATTCCAGGTATGTACTTA 347737

Qy        345 -------------------------GlyArgPheProSerAspHisTrpGlyLeuLeuCy 356
                                        |||         |||||||||:::   :::
Db     347736 CGAAACTATTGTAGACGAAGTGGATGGTTTCTCAGCTTCCGATCATTTTCCTGTTACAGC 347677

Qy        356 sAsnLeuAsp 359
               |||||||||
Db     347676 TAATTTAGAC 347667

RESULT 62
US-10-868-184C-10429/c
; Sequence 10429, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10429
;   LENGTH: 5691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10429

Alignment Scores:
Pred. No.:              0.501           Length:         5691
Score:                  107.00          Matches:        51
Percent Similarity:     43.6%           Conservative:   47
Best Local Similarity:  22.7%           Mismatches:     93
Query Match:            5.6%            Indels:         35
DB:                     6               Gaps:           8

US-10-757-745-2 (1-362) x US-10-868-184C-10429 (1-5691)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::     |||:::                  :::
Db      4113 GAAGAAACTGCATCAACTAATGAGCAAAATAA-CCAGCTAACATCATAATGACAAGATCA 4055

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                  |||       :::::::::|||   |||:::::|||||:::       |||
Db      4054 AATTCACACATAACCATATTAACCTTAAATGTAAATGGGCTAAATGCTCCAAATTAAAGA 3995

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                 |||          :::   :::::::::             |||   |||   ::::::
```

```
Db       3994 CACAGA---------CTGGCAAATTGGATAAGGAGTCAAGACCCATCAGTGTGCTATATT 3944

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::::|||      |||   ::::::
Db       3943 CAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 3902

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                  ||||||:::   |||
Db       3901 ATAAAGGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGAGGTTGCA 3842

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||      ||||||:::   |||:::
Db       3841 ATCCTAGTCTCTGATAAAACAGACTTTAAGCCAACAAAGATCAAAAGAGACAAAGAAGGC 3782

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::             |||:::   ::::::      :::
Db       3781 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGATAACTATCCTAAATATATAT 3722

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||            |||   |||          :::|||   ||||||:::
Db       3721 GCACTCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTTAGA 3677

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db       3676 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCCTA 3620

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db       3619 CTGTCAACATTAGACAGATCAACAAGACAGAAAGTTAACAAGGATATCCAGGAATTGAAC 3560

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db       3559 TCGGCTCTGCACCAAGTGGACCTAATAGACATCTACAGAATTCTC---CACCCCAAATCA 3503

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db       3502 ACAGAATATACATTC 3488
```

RESULT 63
US-10-868-184D-10429/c
; Sequence 10429, Application US/10868184D
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184D
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10429
; LENGTH: 5691
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-868-184D-10429

Alignment Scores:

Pred. No.:	0.501	Length:	5691
Score:	107.00	Matches:	51
Percent Similarity:	43.6%	Conservative:	47
Best Local Similarity:	22.7%	Mismatches:	93
Query Match:	5.6%	Indels:	35
DB:	6	Gaps:	8

US-10-757-745-2 (1-362) x US-10-868-184D-10429 (1-5691)

```
Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::      |||:::                  :::
Db       4113 GAAGAAACTGCATCAACTAATGAGCAAAATAA-CCAGCTAACATCATAATGACAAGATCA 4055

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||      :::::::::|||   |||::::::||||||:::      |||
Db       4054 AATTCACACATAACCATATTAACCTTAAATGTAAATGGGCTAAATGCTCCAAATTAAAGA 3995

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              |||         :::   :::::::::                |||   |||   ::::::
Db       3994 CACAGA---------CTGGCAAATTGGATAAGGAGTCAAGACCCATCAGTGTGCTATATT 3944

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                  :::::::|||      |||   ::::::
Db       3943 CAGGAA-----------------ACCCATCTCATGTGCAGAGACACACATAGGCTCAAA 3902

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                  ||||||:::   |||
Db       3901 ATAAAGGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGAGGTTGCA 3842

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::         |||      ||||||:::   |||:::
Db       3841 ATCCTAGTCTCTGATAAAACAGACTTTAAGCCAACAAAGATCAAAAGAGACAAAGAAGGC 3782

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::             |||:::   ::::::      :::
Db       3781 CATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGATAACTATCCTAAATATATAT 3722

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||            |||   |||          :::|||   ||||||:::
Db       3721 GCACTCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTTAGA 3677

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||         :::|||   |||:::|||      ||||||   |||
Db       3676 GACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGAGACTTTAACACCCTA 3620

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db       3619 CTGTCAACATTAGACAGATCAACAAGACAGAAAGTTAACAAGGATATCCAGGAATTGAAC 3560

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          :::::::::|||::::::      |||      ||||||
Db       3559 TCGGCTCTGCACCAAGTGGACCTAATAGACATCTACAGAATTCTC---CACCCCAAATCA 3503

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db       3502 ACAGAATATACATTC 3488
```

RESULT 64
US-10-868-184C-10456/c
; Sequence 10456, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245

```
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10456
;   LENGTH: 10789
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10456

Alignment Scores:
Pred. No.:                1.38          Length:          10789
Score:                    107.00        Matches:         55
Percent Similarity:       44.0%         Conservative:    44
Best Local Similarity:    24.4%         Mismatches:      92
Query Match:              5.6%          Indels:          35
DB:                       6             Gaps:            9

US-10-757-745-2 (1-362) x US-10-868-184C-10456 (1-10789)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||     :::   ||||||:::                  :::
Db        4136 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4078

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||:::|||      ||||||:::
Db        4077 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4024

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||:::::::            |||   |||      :::
Db        4023 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3967

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||      |||   ::::::
Db        3966 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3925

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||       :::   |||                   ||||||:::   |||
Db        3924 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGAGGTTGCA 3865

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::       :::       |||       ||||||:::   |||:::
Db        3864 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3805

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
                ::::::   :::   |||:::           ||||||   ::::::   :::
Db        3804 CATTACATAATGATAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3745

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                    :::|||                |||   |||          :::|||   ||||||
Db        3744 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 3700

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||       :::|||   |||:::|||      ||||||   |||
Db        3699 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3643

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272
```

```
                                                        :::|||   :::
Db        3642 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3583

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                           :::::::::|||::::::      |||      ||||||
Db        3582 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3526

Qy         293 CysGlnTyrThrTrp 297
                  :::|||||:::
Db        3525 ACAGAATATACATTT 3511

RESULT 65
US-10-868-184D-10456/c
; Sequence 10456, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10456
;   LENGTH: 10789
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10456

Alignment Scores:
Pred. No.:                1.38          Length:          10789
Score:                    107.00        Matches:         55
Percent Similarity:       44.0%         Conservative:    44
Best Local Similarity:    24.4%         Mismatches:      92
Query Match:              5.6%          Indels:          35
DB:                       6             Gaps:            9

US-10-757-745-2 (1-362) x US-10-868-184D-10456 (1-10789)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||     :::   ||||||:::                  :::
Db        4136 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4078

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                |||      :::::::::|||   ||||||:::|||      ||||||:::
Db        4077 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4024

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||:::::::            |||   |||      :::
Db        4023 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3967

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                 :::::::|||      |||   ::::::
```

```
Db       3966 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3925

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db       3924 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGAGGTTGCA 3865

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db       3864 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3805

Qy        211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226

Db       3804 CATTACATAATGATAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3745

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db       3744 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 3700

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db       3699 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3643

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272

Db       3642 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3583

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db       3582 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3526

Qy        293 CysGlnTyrThrTrp 297

Db       3525 ACAGAATATACATTT 3511

RESULT 66
US-11-266-748A-80889
; Sequence 80889, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 80889
;   LENGTH: 1207
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-80889

Alignment Scores:
Pred. No.:              0.0498          Length:         1207
Score:                  106.50          Matches:        87

Percent Similarity:     36.4%           Conservative:   41
Best Local Similarity:  24.7%           Mismatches:     143
Query Match:            5.6%            Indels:         81
DB:                     9               Gaps:           19

US-10-757-745-2 (1-362) x US-11-266-748A-80889 (1-1207)

Qy         52 TrpGluMetGluArg---AlaLeuAsnSerTyrPheGlu---------------------- 63

Db        154 TGGGACGCGGAGCGTCCCGGGCTGCGTTACTACGCGGAGCGTGGGAACGATCGGTTAACA 213

Qy         64 ---ProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 82

Db        214 GCACCGCCTGTAGGTCACTCGAAACTTTCC------GTCAACACCCGCAGTCAGCCAATA 267

Qy         83 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 102

Db        268 ACATCAGCCAATAAAACAGGCTTCACCCTTTACTTATCCACAGGAGCGATCATGGCTGCG 327

Qy        103 SerGluAsp------------ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 118

Db        328 ATTGCCGACTGGGTGCCCGTTACGCCCAGCACCGCCATCACCCGCTTCACCGTGCTGACG 387

Qy        119 TrpAsnIleAsp-----------------------GlyLeuAspLeuAsnAsnLeuSer 130

Db        388 GTCAATATCCATAAGGGTTTCACCGCGTTGAACCGACGCTTCATCCTGCCCGAGTTGCGT 447

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db        448 GAGGCGGTGCGCAACGTG--------------------GGCGCCGACATGGTGTTCCTG 486

Qy        151 GlnGluVal------------------IleProProTyrTyrSerTyrLeuLysLysArg 164

Db        487 CAGGAAATCCACGGCACCCATGAGCGGCACCCGCAGCGTTACAGCGACTGGCCGAAG--- 543

Qy        165 SerSerAsnTyrGluIleIleThrGly------------------------------- 173

Db        544 ATGCCGCAGTATGAGTTCCTCGCCGACAGCATCTGGCCGCAGTTCGCCTACGGCCGCAAC 603

Qy        174 ------HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191

Db        604 GCGGTCTACCCGCATGGCGACCACGGCAACGCGCTGCTGTCCAAATTCCAGATC---ATC 660

Qy        192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu---Cys 210

Db        661 CGCTACGACAACCTCGACATCTCCCAGAGCGGCCATGAAAGCCGTGGCCTGCTGCATTGC 720

Qy        211 Val---HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSer 229

Db        721 GTGCTGCGCTTGCCGGGCAGCGGCCAAGAGGTGCATGCGATCTGCATGCACCTGGGCCTG 780

Qy        230 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 249

Db        781 CGCGAGGTGCAT---------CGCCAGCAGCAATTGCGTTTGCTGGAGCAGCGCATTGCC 831

Qy        250 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn---LeuArgAspArg 268

Db        832 GAGATTCCGGCTGACGCGGCGCTGATCGTCGCCGGTGATTTCAATGACTGGCGCCACAAG 891

Qy        269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288

Db        892 GCCGACCTGAGCCAAAGCGGGCTCAGGGAAGTGTTTATC---------GAAGCCCATGGC 942

Qy        289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308

Db        943 AAACCGGCGCGC---------ACCTTTCCGGCGCGCCTGCCA------------------ 975

Qy        309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328

Db        976 ---------TTGCTGCCTTTGGATCGCATCTATGTGCGTAACCTGAAGGTGCATAAC--- 1023

Qy        329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348
```

```
Db       1024 ------CCCAGGGCGCTGACGACGCGGCCCTGGTCTCACCTG------------------ 1059

Qy        349 SerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              |||||||||       |||         ::::::::
Db       1060 TCAGACCACGTGCCGCTGTCGGTGGAGATCGAACTA 1095

RESULT 67
US-11-266-748A-133700/c
; Sequence 133700, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 133700
;   LENGTH: 1207
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-133700

Alignment Scores:
Pred. No.:                0.0498          Length:          1207
Score:                    106.50          Matches:         87
Percent Similarity:       36.4%           Conservative:    41
Best Local Similarity:    24.7%           Mismatches:      143
Query Match:              5.6%            Indels:          81
DB:                       9               Gaps:            19

US-10-757-745-2 (1-362) x US-11-266-748A-133700 (1-1207)

Qy          52 TrpGluMetGluArg---AlaLeuAsnSerTyrPheGlu---------------------- 63
               |||:::   ||||||      |||      |||  .|||
Db        1054 TGGGACGCGGAGCGTCCCGGGCTGCGTTACTACGCGGAGCGTGGGAACGATCGGTTAACA 995

Qy          64 ---ProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 82
                |||||||||       |||  |||           |||    |||:::|||
Db         994 GCACCGCCTGTAGGTCACTCGAAACTTTCC------GTCAACACCCGCAGTCAGCCAATA 941

Qy          83 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 102
               |||       :::   |||                ||||||   :::   ::::::
Db         940 ACATCAGCCAATAAAACAGGCTTCACCCTTTACTTATCCACAGGAGCGATCATGGCTGCG 881

Qy         103 SerGluAsp------------ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 118
                  |||               |||           :::   |||::::::::::|||
Db         880 ATTGCCGACTGGGTGCCCGTTACGCCCAGCACCGCCATCACCCGCTTCACCGTGCTGACG 821

Qy         119 TrpAsnIleAsp-----------------------GlyLeuAspLeuAsnAsnLeuSer 130
                  ||||||                          |||       |||
Db         820 GTCAATATCCATAAGGGTTTCACCGCGTTGAACCGACGCTTCATCCTGCCCGAGTTGCGT 761

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
                |||       |||   |||                              |||::::::||||||
Db         760 GAGGCGGTGCGCAACGTG---------------------GGCGCCGACATGGTGTTCCTG 722

Qy         151 GlnGluVal------------------IleProProTyrTyrSerTyrLeuLysLysArg 164
               ||||||:::                |||      ||||||           |||
Db         721 CAGGAAATCCACGGCACCCATGAGCGGCACCCGCAGCGTTACAGCGACTGGCCGAAG--- 665

Qy         165 SerSerAsnTyrGluIleIleThrGly-------------------------------- 173
                     ||||||   :::
Db         664 ATGCCGCAGTATGAGTTCCTCGCCGACAGCATCTGGCCGCAGTTCGCCTACGGCCGCAAC 605

Qy         174 ------HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191
                   :::      |||          :::|||   |||   ::::::   :::
Db         604 GCGGTCTACCCGCATGGCGACCACGGCAACGCGCTGCTGTCCAAATTCCAGATC---ATC 548

Qy         192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu---Cys 210
               :::         :::          |||          |||   ||||||   |||
Db         547 CGCTACGACAACCTCGACATCTCCCAGAGCGGCCATGAAAGCCGTGGCCTGCTGCATTGC 488

Qy         211 Val---HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSer 229
               |||      :::      ||||||   |||:::      :::      ||||||
Db         487 GTGCTGCGCTTGCCGGGCAGCGGCCAAGAGGTGCATGCGATCTGCATGCACCTGGGCCTG 428

Qy         230 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 249
                     |||          |||      ||||||:::::::::   :::::::::
Db         427 CGCGAGGTGCAT---------CGCCAGCAGCAATTGCGTTTGCTGGAGCAGCGCATTGCC 377

Qy         250 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn---LeuArgAspArg 268
               |||   |||       |||   :::|||   |||||||||   |||       |||   :::
Db         376 GAGATTCCGGCTGACGCGGCGCTGATCGTCGCCGGTGATTTCAATGACTGGCGCCACAAG 317

Qy         269 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 288
                        ||||||          :::       |||       |||
Db         316 GCCGACCTGAGCCAAAGCGGGCTCAGGGAAGTGTTTATC---------GAAGCCCATGGC 266

Qy         289 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 308
               ||||||               |||:::      ::::::
Db         265 AAACCGGCGCGC---------ACCTTTCCGGCGCGCCTGCCA------------------ 233

Qy         309 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 328
                        |||       ||||||||||:::   |||
Db         232 ---------TTGCTGCCTTTGGATCGCATCTATGTGCGTAACCTGAAGGTGCATAAC--- 185

Qy         329 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhePro 348
                     ||||||:::|||                |||
Db         184 ------CCCAGGGCGCTGACGACGCGGCCCTGGTCTCACCTG------------------ 149

Qy         349 SerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
               |||||||||       |||         ::::::::
Db         148 TCAGACCACGTGCCGCTGTCGGTGGAGATCGAACTA 113

RESULT 68
US-60-742-219-1685
; Sequence 1685, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
```

```
; SEQ ID NO 1685
;   LENGTH: 3958
;   TYPE: DNA
;   ORGANISM: Homo sapiens.
US-60-742-219-1685

Alignment Scores:
Pred. No.:              0.324          Length:          3958
Score:                  106.50         Matches:         80
Percent Similarity:     36.4%          Conservative:    44
Best Local Similarity:  23.5%          Mismatches:      108
Query Match:            5.6%           Indels:          110
DB:                     11             Gaps:            20

US-10-757-745-2 (1-362) x US-60-742-219-1685 (1-3958)

Qy         64 ProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThr 83
Db        670 CCACCTACAGAGGAAACTGCCTATGGAAGA-----------------TCACTAAAAGTG 711

Qy         84 TyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSer 103
Db        712 GCACTGGGTATTCTGAATGAGAGAACA---------AATTTGAGTCAAGCAAGCACTTCA 762

Qy        104 GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle------ 121
Db        763 GATGAAGAGGAG-----------------ATCACTATGCTGTCTCAAAATGTACCACAA 804

Qy        122 ------------------------------------AspGlyLeuAspLeuAsnAsn 128
Db        805 AAACAGTCCGATTCACCCCCTCATAAAAAATACCGGAAGGATGAAGGTGACTTACCAGGG 864

Qy        129 ---LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspVal 147
Db        865 TGTCTTGAGGAAAGGGAAAACTCAGCATGCTTGTTAGCATCTTCAGAGAGTGATGATTCC 924

Qy        148 IlePhe-LeuGlnGluValIleProProTyrTyrSerTyrLeu---------LysLysAr 164
Db        925 CTGTATGATGATAAATCACAAGCACCCACAATGGTCGATACTATTCCAAGTGAAGTGGAA 984

Qy        164 gSerSerAsnTyrGluIleIleThrGlyHisGluGluGly---TyrPheThrAlaIleMe 183
Db        985 ACAAAGTCATTACAAAACTCTAGCTGGTGCGAGACTTTCCCTTCACTTTCGGAAGATAAT 1044

Qy        183 tLeuLysLys---SerArgValLysLeuLysSerGlnGluIleIleProPheProSerTh 202
Db       1045 GATGAAAAAGAGAACAAGAATAAGATTGATATCTCAGCAG-------------------- 1084

Qy        202 rLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLe 222
Db       1085 ------------------TTATGT-------------------------CTGTGCAT 1098

Qy        222 uMetThrSer------------HisLeuGluSerThrArgGly------HisAlaAlaGl 236
Db       1099 TCTGCAGTCAAAGAGGAAAGTGCATGTGTTAAAGATGAAAAGTTTGCTCCACCTTTGTCA 1158

Qy        236 uArgMetAsnGlnLeuLysMetValLeu---LysLysMetGlnGluAlaProGluSerAl 255
Db       1159 CCTTTGTCATCAGATATG-CTCATTATGCCCAAAGCTTTGAAAGAAGAGAGCGAG----- 1212

Qy        255 aThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluVal---ThrArgCysGl 274
Db       1213 ------------------GATACCTGCCTAGAGACCCTGGCTGTTCCCTCTGAATGCTC 1253

Qy        274 yGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGl 294
Db       1254 TGCTTTCTCAGAGAATATTGAGGATCCTGGAGAG------GGTCCCTCAAATCCATGCTT 1307

Qy        294 nTyrThrTrpAspThrGln------MetAsnSerAsnLeuGlyIleThrAlaAlaCys-- 311
Db       1308 AGATACCAGCCAGAATCAACCTTCCATGGAATCAGAGATGGGG---GCTGCAGCATGCCC 1364

Qy        311 ------------------------------------------------------------ 311
```

```
Db       1365 TGGGAGTTGTTCAAGGGAATGCGAGGTTTCATTTAGTGCCTCTAACCCTGTCTGGGATTA 1424

Qy        312 ------------------LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGl 325
Db       1425 TTCACATCTTATGAGTAGTGAAAGAAATTTTCAGAGACTGGATTTTGAAGAACTTGAGGA 1484

Qy        325 uGluGlyHisIleIleProArgSerLeu------AspLeuLeuGlyLeuGluLysLeuAs 343
Db       1485 AGAAGGTCAAGCCTCTGACAAGTCATTGCTTCCAAGTCGCATTAATCTTTCTCTATTAGA 1544

Qy        343 p 343
Db       1545 T 1545

RESULT 69
US-10-498-451-2320
; Sequence 2320, Application US/10498451
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Wang, Jian-rui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Xu, Chongjun
;  APPLICANT:  Boyle, Bryan J
;  APPLICANT:  Drmanac, Radoje T
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 820/PCT
;  CURRENT APPLICATION NUMBER: US/10/498,451
;  CURRENT FILING DATE:  2004-06-09
;  PRIOR APPLICATION NUMBER: US 60/339,739
;  PRIOR FILING DATE: 2001-12-10
;  PRIOR APPLICATION NUMBER: US 10/128,558
;  PRIOR FILING DATE: 2002-04-22
;  PRIOR APPLICATION NUMBER: US 60/339,453
;  PRIOR FILING DATE: 2001-12-11
;  PRIOR APPLICATION NUMBER: US 60/365,384
;  PRIOR FILING DATE: 2002-03-14
;  PRIOR APPLICATION NUMBER: US 60/365,091
;  PRIOR FILING DATE: 2002-03-14
;  PRIOR APPLICATION NUMBER: US 60/372,615
;  PRIOR FILING DATE: 2002-04-12
;  PRIOR APPLICATION NUMBER: US 60/376,045
;  PRIOR FILING DATE: 2002-04-24
;  PRIOR APPLICATION NUMBER: US 60/372,381
;  PRIOR FILING DATE: 2002-04-12
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 3134
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 2320
;   LENGTH: 6028
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-498-451-2320

Alignment Scores:
```

```
Pred. No.:              0.63          Length:         6028
Score:                  106.50        Matches:        76
Percent Similarity:     40.7%         Conservative:   53
Best Local Similarity:  24.0%         Mismatches:     106
Query Match:            5.6%          Indels:         83
DB:                     6             Gaps:           16

US-10-757-745-2 (1-362) x US-10-498-451-2320 (1-6028)

Qy          15 GluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGluPheAlaSer 34
               :::|||           :::::::::::   ||||||          |||:::::
Db         144 GACGAAACTTCCAGAGGAACGATCAGGCAGCAACATTTGCTG---------TTCAGCAAT 194

Qy          35 ValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMet 54
               :::      |||  ||||||  |||          |||          |||
Db         195 ATC---CGCTGTTCTGCAGCCTCTGCTGCTGATACCCAGGCAAACAGGGTCTGGAGTGGA 251

Qy          55 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 74
                  |||   ||||||                                    :::|||
Db         252 ACTCCAGCA---AACTCC-----------------------------------AAAAGA 272

Qy          75 Pro-----------------------------------------GluThrIleSerGlu 80
               |||                                       |||      |||
Db         273 CCTGCAGCTGAGGGTCCTGACTGTTAGAAGGAAAACTAACAAACAGAAAGGACATCCACA 332

Qy          81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
               ||||||   |||:::
Db         333 CCAAAACCCTATCTG-----------------------------------------TAC 350

Qy         101 -SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAs 120
                |||||||||      |||:::     |||   |||     :::::::::|||   ||
Db         351 GTCACCATCATCAAAGACCAAAGGATCAAAT---TCACACATAACAATATTAACCTTAAA 407

Qy         120 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLe 140
               |:::::::||||||         |||          :::        :::   ::::::::
Db         408 TGTAAATGGGCTA---------AATGCTCCAATTAAAAGACACAGACTGGCAAATTGGAT 458

Qy         140 uAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTy 160
               :          |||  |||      :::||||||                    :::::
Db         459 AAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------ACCCA 500

Qy         160 rLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheTh 180
               :|||     |||   ::::::          |||   |||       |||   |||
Db         501 TCTCACGTGCAGAGACACACACAGGATCAAAATAAAGGGATGGAGGGAGATCTACCAAGC 560

Qy         180 rAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePr 200
                        ||||||:::   |||  :::          :::          |||
Db         561 AAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACGGACTTTAA 620

Qy         200 oSerThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnValSe 216
                  ||||||:::   |||             :::   :::   |||      ::::::
Db         621 ACCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATGATGGTAAAGGGATCAATTCA 680

Qy         216 rGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGl 236
                  ||||||      :::      :::      :::|||            |||
Db         681 ACAAGAAGAGCTAACTACCCTAAATATATATGCACCCAATACA---------GGAGCACC 731

Qy         236 uArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaTh 256
                |||      :::|||   ||||||:::  :::|||          :::|||   ||
Db         732 CAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACAC 785

Qy         256 rValIlePheAlaGlyAsp------------------------ThrAsnLeuArgAspA 268
               |:::|||      ||||||                      |||:::   |||||||
Db         786 ATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACA 842

Qy         268 rgGluValThrArgCysGlyGlyLeuProAsn-----------------AsnIleValA 282
               ||:::::||||||          :::   |||               :::::::::|
Db         843 GAAAGTTAACAAGGATATCCAGGATTTG-AACTCAGCTCTGCACCAAGTGGACCTAATAG 901

Qy         282 spValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 297
               ||::::::      |||     ||||||     :::|||||:::
Db         902 ACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 945
```

```
RESULT 70
US-60-762-056-10029/c
; Sequence 10029, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 10029
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-10029

Alignment Scores:
Pred. No.:              0.158         Length:         2300
Score:                  106.00        Matches:        56
Percent Similarity:     44.0%         Conservative:   43
Best Local Similarity:  24.9%         Mismatches:     92
Query Match:            5.6%          Indels:         35
DB:                     11            Gaps:           9

US-10-757-745-2 (1-362) x US-60-762-056-10029 (1-2300)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||  ||||||      :::   ||||||:::                   :::
Db         634 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 576

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
                  |||      :::::::::|||   ||||||:::|||      ||||||:::
Db         575 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 522

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||  |||   :::  |||::::::           |||  |||       :::
Db         521 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 465

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  :::::::|||      |||   ::::::
Db         464 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 423

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                ||||||:::  . |||
Db         422 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 363

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::          :::          |||      ||||||:::   |||:::
Db         362 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 303

Qy         211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
                  ::::::   |||   |||:::            ||||||   ::::::       :::
Db         302 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 243

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::|||          |||   |||          :::|||   ||||||
Db         242 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 198

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
                  :::|||          :::|||   |||:::|||      ||||||   |||
Db         197 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 141

Qy         266 -------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db         140 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 81

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                           :::::::::|||::::::      |||      ||||||
Db          80 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 24
```

```
Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db         23 ACAGAATATACATTT 9

RESULT 71
US-10-868-184C-10658/c
; Sequence 10658, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10658
;   LENGTH: 5446
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10658

Alignment Scores:
Pred. No.:                0.616          Length:          5446
Score:                    106.00         Matches:         56
Percent Similarity:       44.0%          Conservative:    43
Best Local Similarity:    24.9%          Mismatches:      92
Query Match:              5.6%           Indels:          35
DB:                       6              Gaps:            9

US-10-757-745-2 (1-362) x US-10-868-184C-10658 (1-5446)

Qy         91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
              |||   |||   ||||||      :::   ||||||:::                  :::
Db       4138 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4080

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
              |||       :::::::::|||   ||||||:::|||      ||||||:::
Db       4079 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4026

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
              :::|||   |||   :::   |||::::::             |||   |||      :::
Db       4025 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3969

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
              ||||||                 :::::::|||      |||   ::::::
Db       3968 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3927

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
              |||   |||      :::   |||                ||||||:::   |||
Db       3926 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3867

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
              :::         :::          |||       ||||||:::   |||:::
Db       3866 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3807

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
              ::::::   |||   |||:::             ||||||   ::::::      :::
Db       3806 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3747

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
              :::|||           |||   |||          :::|||   ||||||
Db       3746 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 3702

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
              :::|||     :::|||   |||:::|||      ||||||   |||
Db       3701 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3645

Qy        266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db       3644 CTGTCAATATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3585

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                          ::::::::::||:::::      |||      ||||||
Db       3584 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3528

Qy        293 CysGlnTyrThrTrp 297
              :::||||||:::
Db       3527 ACAGAATATACATTT 3513

RESULT 72
US-10-868-184D-10658/c
; Sequence 10658, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10658
;   LENGTH: 5446
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10658

Alignment Scores:
Pred. No.:                0.616          Length:          5446
Score:                    106.00         Matches:         56
Percent Similarity:       44.0%          Conservative:    43
Best Local Similarity:    24.9%          Mismatches:      92
Query Match:              5.6%           Indels:          35
DB:                       6              Gaps:            9
```

```
US-10-757-745-2 (1-362) x US-10-868-184D-10658 (1-5446)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::                  :::
Db        4138 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4080

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      ::::::::::|||   ||||||:::|||      ||||||:::
Db        4079 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4026

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||::::::                |||   |||      :::
Db        4025 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3969

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                    :::::::|||     |||   ::::::
Db        3968 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3927

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                    ||||||:::   |||
Db        3926 ATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3867

Qy         191 LeuLysSerGlnGluIleIleProProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::         :::         |||       ||||||:::   |||:::
Db        3866 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3807

Qy         211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::              ||||||   ::::::      :::
Db        3806 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3747

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||              |||   |||          :::|||   ||||||
Db        3746 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 3702

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||          :::|||   |||:::|||      ||||||   |||
Db        3701 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3645

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db        3644 CTGTCAATATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3585

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                        ::::::::::|||:::::      |||      ||||||
Db        3584 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3528

Qy         293 CysGlnTyrThrTrp 297
               :::||||||:::
Db        3527 ACAGAATATACATTT 3513

RESULT 73
US-10-868-184C-10684/c
; Sequence 10684, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10684
;   LENGTH: 5453
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10684

Alignment Scores:
Pred. No.:                0.618          Length:          5453
Score:                    106.00         Matches:         56
Percent Similarity:       44.0%          Conservative:    43
Best Local Similarity:    24.9%          Mismatches:      92
Query Match:              5.6%           Indels:          35
DB:                       6              Gaps:            9

US-10-757-745-2 (1-362) x US-10-868-184C-10684 (1-5453)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::                  :::
Db        4134 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4076

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      ::::::::::|||   ||||||:::|||      ||||||:::
Db        4075 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4022

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||::::::                |||   |||      :::
Db        4021 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3965

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                    :::::::|||     |||   ::::::
Db        3964 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3923

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                    ||||||:::   |||
Db        3922 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3863

Qy         191 LeuLysSerGlnGluIleIleProProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::         :::         |||       ||||||:::   |||:::
Db        3862 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3803

Qy         211 ValHisValAsnValSerGlyAsn------------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::              ||||||   ::::::      :::
Db        3802 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3743

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
               :::|||              |||   |||          :::|||   ||||||
Db        3742 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3698

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||          :::|||   |||:::|||      ||||||   |||
Db        3697 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3641

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272
                                                   :::|||   :::
Db        3640 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3581

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                        ::::::::::|||:::::      |||      ||||||
Db        3580 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3524

Qy         293 CysGlnTyrThrTrp 297
               :::||||||:::
Db        3523 ACAGAATATACATTT 3509
```

```
RESULT 74
US-10-868-184D-10684/c
; Sequence 10684, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10684
;   LENGTH: 5453
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10684

Alignment Scores:
Pred. No.:                0.618          Length:          5453
Score:                    106.00         Matches:         56
Percent Similarity:       44.0%          Conservative:    43
Best Local Similarity:    24.9%          Mismatches:      92
Query Match:              5.6%           Indels:          35
DB:                       6              Gaps:            9

US-10-757-745-2 (1-362) x US-10-868-184D-10684 (1-5453)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
               |||   |||   ||||||      :::   ||||||:::                  :::
Db        4134 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 4076

Qy         111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130
               |||      ::::::::::|||   |||||||:::|||      |||||||:::
Db        4075 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 4022

Qy         131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150
               :::|||   |||   :::   |||::::::           |||   |||      :::
Db        4021 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3965

Qy         151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170
               ||||||                  ::::::|||      |||   ::::::
Db        3964 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3923

Qy         171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190
               |||   |||      :::   |||                  ||||||:::   |||
Db        3922 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3863

Qy         191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210
               :::         :::         |||      ||||||:::   |||:::
Db        3862 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3803

Qy         211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226
               ::::::   |||   |||:::             ||||||   ::::::      :::
Db        3802 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3743

Qy         227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246
                  :::|||           |||   |||         :::|||   ||||||
Db        3742 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGT 3698

Qy         247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265
               :::|||          :::|||   |||:::|||      ||||||   |||
Db        3697 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3641

Qy         266 ------------------------------------ArgAspArgGluValThrArg 272
                                                    :::|||   :::
Db        3640 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3581

Qy         273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
                         ::::::::::|||::::::      |||      ||||||
Db        3580 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3524

Qy         293 CysGlnTyrThrTrp 297
               :::||||||:::
Db        3523 ACAGAATATACATTT 3509

RESULT 75
US-10-868-184C-10535/c
; Sequence 10535, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10535
;   LENGTH: 5916
;   TYPE: DNA
; ~ ORGANISM: Homo sapiens
US-10-868-184C-10535

Alignment Scores:
Pred. No.:                0.702          Length:          5916
Score:                    106.00         Matches:         56
Percent Similarity:       44.0%          Conservative:    43
Best Local Similarity:    24.9%          Mismatches:      92
Query Match:              5.6%           Indels:          35
DB:                       6              Gaps:            9

US-10-757-745-2 (1-362) x US-10-868-184C-10535 (1-5916)

Qy          91 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 110
```

```
Db       3973 GAAGAAACTGCATCAACTAATGAGCAAAA-TCACCAGCTAACATCATAATGACAGGATCA 3915

Qy        111 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 130

Db       3914 AATTCACACATAACAATATTAACTTTAAATATAAATGGA------CTAAATTCTGCAATT 3861

Qy        131 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 150

Db       3860 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATT 3804

Qy        151 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 170

Db       3803 CAGGAA-----------------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAA 3762

Qy        171 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 190

Db       3761 ATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAAAAGGCAGGGGTTGCA 3702

Qy        191 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 210

Db       3701 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAGGC 3642

Qy        211 ValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMetThrSerHis 226

Db       3641 CATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACTATCCTAAATATTTAT 3582

Qy        227 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 246

Db       3581 GCACCCAATACA---------GGAGCACCCAGATTC------ATAAAGCAAGTCCTCAGT 3537

Qy        247 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--- 265

Db       3536 GACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGAGACTTTAACACCCCA 3480

Qy        266 -------------------------------------ArgAspArgGluValThrArg 272

Db       3479 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGATACCCAGGAATTGAAC 3420

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292

Db       3419 TCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCA 3363

Qy        293 CysGlnTyrThrTrp 297

Db       3362 ACAGAATATACCTTT 3348
```

Search completed: March 11, 2006, 00:16:07
Job time : 1358.96 secs

Liu, S.
10/757745
Seq IDs 2, 3 & 6
w/Inters

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:19:50 ; Search time 14878.9 Seconds
(without alignments)
1345.232 Million cell updates/sec

Title: US-10-757-745-2
Perfect score: 1902
Sequence: 1 MELGSCLEGGREAAEEEGEP.....DCGRFPSDHWGLLCNLDIIL 362

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
MODEL=frame+_p2n.model -DEV=xlh
Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170333_4165/app_query.fasta_1
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=rnpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss02h -USER=US10757745_@CGN_1_1_20983_@runat_09032006_170333_4165
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUSB_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUSC_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US075_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US076_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US077_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US078_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US079_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US101_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US104_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US105_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US112_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US117_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US600_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US601_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US603_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US605_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US606_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US607_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1902	100.0	1919	42	US-10-170-235-27345	Sequence 27345, A
2	1902	100.0	1920	61	US-10-757-745-1	Sequence 1, Appli
3	1902	100.0	1940	3	PCT-US04-24424-921	Sequence 921, App
4	1902	100.0	1940	61	US-10-783-271-28	Sequence 28, Appl
5	1902	100.0	1940	80	US-60-490-890-921	Sequence 921, App
6	1902	100.0	1940	81	US-60-568-073-582	Sequence 582, App
7	1902	100.0	1940	82	US-60-671-501-156	Sequence 156, App
8	1902	100.0	1940	82	US-60-685-372-523	Sequence 523, App

9	1902	100.0	1940	83	US-60-704-740-28	Sequence 28, Appl
10	1902	100.0	1940	83	US-60-710-726-741	Sequence 741, App
11	1902	100.0	1948	24	US-09-488-725A-1200	Sequence 1200, Ap
12	1902	100.0	1948	40	US-10-037-270-889	Sequence 889, App
13	1902	100.0	1948	42	US-10-105-837-889	Sequence 889, App
14	1902	100.0	1948	42	US-10-117-722-889	Sequence 889, App
15	1902	100.0	1948	42	US-10-122-851-889	Sequence 889, App
16	1902	100.0	1948	43	US-10-258-898A-1200	Sequence 1200, Ap
17	1902	100.0	1948	43	US-10-286-897-1200	Sequence 1200, Ap
18	1902	100.0	1958	24	US-09-471-275-1314	Sequence 1314, Ap
19	1902	100.0	1958	24	US-09-488-725A-4772	Sequence 4772, Ap
20	1902	100.0	1958	24	US-09-488-725B-8295	Sequence 8295, Ap
21	1902	100.0	1958	27	US-09-552-317-8295	Sequence 8295, Ap
22	1902	100.0	1958	43	US-10-258-898A-4772	Sequence 4772, Ap
23	1902	100.0	1958	43	US-10-286-897-4772	Sequence 4772, Ap
24	1902	100.0	1958	48	US-10-302-689A-59957	Sequence 59957, A
25	1902	100.0	1958	48	US-10-302-689A-77389	Sequence 77389, A
26	1902	100.0	2262	28	US-09-606-680-2868	Sequence 2868, Ap
27	1902	100.0	2262	28	US-09-606-776-4085	Sequence 4085, Ap
28	1902	100.0	2262	29	US-09-644-869-9498	Sequence 9498, Ap
29	1902	100.0	2262	29	US-09-652-121-6794	Sequence 6794, Ap
30	1902	100.0	2262	29	US-09-652-122-3756	Sequence 3756, Ap
31	1902	100.0	2262	29	US-09-652-128-9329	Sequence 9329, Ap
32	1902	100.0	2262	29	US-09-652-816-9572	Sequence 9572, Ap
33	1902	100.0	2262	31	US-09-700-000-6419	Sequence 6419, Ap
34	1902	100.0	2262	31	US-09-710-280-2045	Sequence 2045, Ap
35	1902	100.0	2262	31	US-09-710-282-3926	Sequence 3926, Ap
36	1902	100.0	2262	31	US-09-726-802-2141	Sequence 2141, Ap
37	1902	100.0	3152	38	US-09-981-353-88	Sequence 88, Appl
38	1902	100.0	3152	77	US-60-239-841-88	Sequence 88, Appl
39	1894	99.6	1296	1	PCT-US00-05883-170	Sequence 170, App
40	1894	99.6	1296	36	US-09-925-299-170	Sequence 170, App
41	1893	99.5	1898	28	US-09-629-469A-13209	Sequence 13209, A
42	1893	99.5	1898	63	US-10-917-503-13209	Sequence 13209, A
43	1889	99.3	1930	31	US-09-705-256A-7708	Sequence 7708, Ap
44	1889	99.3	1930	75	US-60-164-285-7708	Sequence 7708, Ap
45	1889	99.3	3080	27	US-09-585-944-2184	Sequence 2184, Ap
c 46	1889	99.3	3082	75	US-60-172-360-956	Sequence 956, App
47	1864	98.0	1919	79	US-60-443-566-1795	Sequence 1795, Ap
48	1864	98.0	1919	79	US-60-452-680-10426	Sequence 10426, A
49	1864	98.0	1919	79	US-60-455-444-3324	Sequence 3324, Ap
50	1864	98.0	1919	79	US-60-465-241-3324	Sequence 3324, Ap
c 51	1664	87.5	1849	42	US-10-170-235-26834	Sequence 26834, A
c 52	1626	85.5	1849	79	US-60-443-566-1789	Sequence 1789, Ap
c 53	1626	85.5	1849	79	US-60-452-680-10422	Sequence 10422, A
c 54	1626	85.5	1849	79	US-60-455-444-3323	Sequence 3323, Ap
c 55	1626	85.5	1849	79	US-60-465-241-3323	Sequence 3323, Ap
56	1549	81.4	1852	29	US-09-644-868-7772	Sequence 7772, Ap
57	1549	81.4	1852	29	US-09-649-161-8751	Sequence 8751, Ap
58	1549	81.4	1852	29	US-09-652-109-8165	Sequence 8165, Ap
59	1549	81.4	1852	29	US-09-652-128-7441	Sequence 7441, Ap
60	1549	81.4	1852	29	US-09-652-816-8491	Sequence 8491, Ap
61	1549	81.4	1852	29	US-09-652-914-8450	Sequence 8450, Ap
62	1258	66.1	1312	61	US-10-757-745-3	Sequence 3, Appli
63	1258	66.1	1944	42	US-10-144-771-826	Sequence 826, App
64	1258	66.1	1944	78	US-60-360-207-826	Sequence 826, App
65	1199	63.0	1088	1	PCT-US01-08631-22058	Sequence 22058, A
66	1199	63.0	1088	48	US-10-302-689A-136240	Sequence 136240,
67	1199	63.0	1088	51	US-10-450-763-22058	Sequence 22058, A
68	1182	62.1	752	28	US-09-629-469A-4908	Sequence 4908, Ap
69	1182	62.1	752	63	US-10-917-503-4908	Sequence 4908, Ap
c 70	1177	61.9	1079	22	US-09-248-178-63	Sequence 63, Appl
c 71	1177	61.9	1079	22	US-09-288-950-63	Sequence 63, Appl
c 72	1177	61.9	1079	23	US-09-346-327-63	Sequence 63, Appl
c 73	1177	61.9	1079	29	US-09-687-507-63	Sequence 63, Appl
c 74	1177	61.9	1079	29	US-09-687-507A-63	Sequence 63, Appl
c 75	1177	61.9	1079	31	US-09-745-288-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-10-170-235-27345

; Sequence 27345, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 27345
; LENGTH: 1919
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-27345

Alignment Scores:

Pred. No.:	2.31e-291	Length:	1919
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	42	Gaps:	0

US-10-757-745-2 (1-362) x US-10-170-235-27345 (1-1919)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 77

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 137

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 197

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 257

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 317

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 377

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 617

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 677

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 737
```

```
Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 797

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        798 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 857

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 917

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 977

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1037

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1038 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1097

Qy        361 IleLeu 362
              ||||||
Db       1098 ATATTG 1103
```

RESULT 2
US-10-757-745-1
; Sequence 1, Application US/10757745
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/10/757,745
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/697,863A
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1627)..(1627)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1108)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

Alignment Scores:

Pred. No.:	2.31e-291	Length:	1920
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	61	Gaps:	0

US-10-757-745-2 (1-362) x US-10-757-745-1 (1-1920)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 79

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 139

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 199

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 259

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 319

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 379

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1099

Qy        361 IleLeu 362
```

```
                ||||||
Db       1100 ATATTG 1105

RESULT 3
PCT-US04-24424-921
; Sequence 921, Application PC/TUS0424424
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PCT
;  CURRENT APPLICATION NUMBER: PCT/US04/24424
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US 60/490,890
;  PRIOR FILING DATE: 2003-07-29
;  NUMBER OF SEQ ID NOS: 2786
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US04-24424-921

Alignment Scores:
Pred. No.:                2.36e-291      Length:          1940
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       3              Gaps:            0

US-10-757-745-2 (1-362) x PCT-US04-24424-921 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
              ||||||
Db       1106 ATATTG 1111

RESULT 4
US-10-783-271-28
; Sequence 28, Application US/10783271
; GENERAL INFORMATION:
;  APPLICANT: Veridex, LLC
;  APPLICANT:  Wang, Yixin
;  TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
;  FILE REFERENCE: VDX-5003 USNP
;  CURRENT APPLICATION NUMBER: US/10/783,271
;  CURRENT FILING DATE:  2004-02-20
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: human
US-10-783-271-28

Alignment Scores:
Pred. No.:                2.36e-291      Length:          1940
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       61             Gaps:            0

US-10-757-745-2 (1-362) x US-10-783-271-28 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
              ||||||
Db       1106 ATATTG 1111

RESULT 5
US-60-490-890-921
; Sequence 921, Application US/60490890

; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PSP
;  CURRENT APPLICATION NUMBER: US/60/490,890
;  CURRENT FILING DATE:  2003-07-29
;  NUMBER OF SEQ ID NOS: 2779
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-490-890-921

Alignment Scores:
Pred. No.:                2.36e-291       Length:          1940
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       80              Gaps:            0

US-10-757-745-2 (1-362) x US-60-490-890-921 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
```

```
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260

Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280

Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340

Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360

Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362

Db       1106 ATATTG 1111

RESULT 6
US-60-568-073-582
; Sequence 582, Application US/60568073
; GENERAL INFORMATION:
;  APPLICANT: Huang, Pei
;  APPLICANT:  Han, Xia
;  APPLICANT:  Shaw, Peter
;  APPLICANT:  Clark, Edwin
;  TITLE OF INVENTION: COLON CANCER BIOMARKERS
;  FILE REFERENCE: 10205 PSP
;  CURRENT APPLICATION NUMBER: US/60/568,073
;  CURRENT FILING DATE:  2004-05-04
;  NUMBER OF SEQ ID NOS: 1305
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 582
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-568-073-582

Alignment Scores:
Pred. No.:              2.36e-291       Length:         1940
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     81              Gaps:           0

US-10-757-745-2 (1-362) x US-60-568-073-582 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20

Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40

Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60

Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220

Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240

Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260

Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280

Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340

Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360

Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362

Db       1106 ATATTG 1111

RESULT 7
US-60-671-501-156
; Sequence 156, Application US/60671501
; GENERAL INFORMATION:
;  APPLICANT: Wim Van Criekinge
;  APPLICANT:  Josef Straub
;  TITLE OF INVENTION: METHYLATION MARKERS FOR DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF CANCERS
;  FILE REFERENCE: 006791.00001
```

```
;  CURRENT APPLICATION NUMBER: US/60/671,501
;  CURRENT FILING DATE:  2005-04-15
;  NUMBER OF SEQ ID NOS: 420
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 156
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-671-501-156

Alignment Scores:
Pred. No.:                2.36e-291      Length:          1940
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       82             Gaps:            0

US-10-757-745-2 (1-362) x US-60-671-501-156 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
              ||||||
Db       1106 ATATTG 1111

RESULT 8
US-60-685-372-523
; Sequence 523, Application US/60685372
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John V
;  APPLICANT:  Bradley, Walter E
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall D
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul V
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: GeneMap of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 059908-5005-PR
;  CURRENT APPLICATION NUMBER: US/60/685,372
;  CURRENT FILING DATE:  2005-05-31
;  NUMBER OF SEQ ID NOS: 2738
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 523
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-685-372-523

Alignment Scores:
Pred. No.:                2.36e-291      Length:          1940
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       82             Gaps:            0

US-10-757-745-2 (1-362) x US-60-685-372-523 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
                 ||||||
Db       1106 ATATTG 1111
```

```
RESULT 9
US-60-704-740-28
; Sequence 28, Application US/60704740
; GENERAL INFORMATION:
;  APPLICANT: Wang, Yixin
;  APPLICANT:  Zhang, Yi
;  APPLICANT:  Atkins, David
;  APPLICANT:  Sieuwerts, Anieta M.
;  APPLICANT:  Smid, Marcel
;  APPLICANT:  Klijn, Jan G.M.
;  APPLICANT:  Martens, John W.M.
;  APPLICANT:  Foekens, John A.
;  TITLE OF INVENTION: Predicting Bone Relapse of Breast Cancer
;  FILE REFERENCE: VDX 5026 USPSP
;  CURRENT APPLICATION NUMBER: US/60/704,740
;  CURRENT FILING DATE:  2005-08-02
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: human
US-60-704-740-28

Alignment Scores:
Pred. No.:                2.36e-291       Length:          1940
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       83              Gaps:            0

US-10-757-745-2 (1-362) x US-60-704-740-28 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
              ||||||
Db       1106 ATATTG 1111

RESULT 10
US-60-710-726-741
; Sequence 741, Application US/60710726
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John V
;  APPLICANT:  Bradley, Walter E
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Fournier, Helene
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Serre, Valerie
;  APPLICANT:  Eerdewegh, Paul V
;  APPLICANT:  Little, Randall D
;  APPLICANT:  Keith , Tim
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: GeneMap of the human genes associated with Crohn's disease
;  FILE REFERENCE: 059908-5009-PR
;  CURRENT APPLICATION NUMBER: US/60/710,726
;  CURRENT FILING DATE:  2005-08-24
;  NUMBER OF SEQ ID NOS: 20166
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 741
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homosapiens
US-60-710-726-741

Alignment Scores:
Pred. No.:              2.36e-291       Length:         1940
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     83              Gaps:           0

US-10-757-745-2 (1-362) x US-60-710-726-741 (1-1940)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 265

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 325

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 385

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1105

Qy        361 IleLeu 362
              ||||||
Db       1106 ATATTG 1111

RESULT 11
US-09-488-725A-1200
; Sequence 1200, Application US/09488725A
; GENERAL INFORMATION:
```

```
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/09/488,725A
;  CURRENT FILING DATE:  2000-12-22
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7144
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-09-488-725A-1200

Alignment Scores:
Pred. No.:               2.38e-291      Length:          1948
Score:                   1902.00        Matches:         362
Percent Similarity:      100.0%         Conservative:    0
Best Local Similarity:   100.0%         Mismatches:      0
Query Match:             100.0%         Indels:          0
DB:                      24             Gaps:            0

US-10-757-745-2 (1-362) x US-09-488-725A-1200 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124

RESULT 12
US-10-037-270-889
; Sequence 889, Application US/10037270
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/10/037,270
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
```

```
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-037-270-889

Alignment Scores:
Pred. No.:                2.38e-291      Length:          1948
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       40             Gaps:            0

US-10-757-745-2 (1-362) x US-10-037-270-889 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124

RESULT 13
US-10-105-837-889
; Sequence 889, Application US/10105837
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BDIVA
;  CURRENT APPLICATION NUMBER: US/10/105,837
;  CURRENT FILING DATE:  2002-03-21
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-105-837-889

Alignment Scores:
Pred. No.:                2.38e-291      Length:          1948
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       42             Gaps:            0

US-10-757-745-2 (1-362) x US-10-105-837-889 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgProProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124

RESULT 14
US-10-117-722-889
; Sequence 889, Application US/10117722

; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BCIP
;  CURRENT APPLICATION NUMBER: US/10/117,722
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-117-722-889

Alignment Scores:
Pred. No.:                2.38e-291       Length:          1948
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       42              Gaps:            0

US-10-757-745-2 (1-362) x US-10-117-722-889 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578
```

```
Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124
```

RESULT 15
US-10-122-851-889
; Sequence 889, Application US/10122851
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BDV3
; CURRENT APPLICATION NUMBER: US/10/122,851
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-122-851-889

Alignment Scores:

Pred. No.:	2.38e-291	Length:	1948
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	42	Gaps:	0

US-10-757-745-2 (1-362) x US-10-122-851-889 (1-1948)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998
```

```
Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124

RESULT 16
US-10-258-898A-1200
; Sequence 1200, Application US/10258898A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/258,898A
;  CURRENT FILING DATE:  2002-10-29
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-258-898A-1200

Alignment Scores:
Pred. No.:                2.38e-291       Length:          1948
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       43              Gaps:            0

US-10-757-745-2 (1-362) x US-10-258-898A-1200 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124

RESULT 17
US-10-286-897-1200
; Sequence 1200, Application US/10286897
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/286,897
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US/09/552,317
```

```
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US/09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US/09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US/09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US/09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US/09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US/09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-286-897-1200

Alignment Scores:
Pred. No.:              2.38e-291     Length:          1948
Score:                  1902.00       Matches:         362
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     43            Gaps:            0

US-10-757-745-2 (1-362) x US-10-286-897-1200 (1-1948)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 98

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 158

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 218

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 278

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 338

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 398

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1118

Qy        361 IleLeu 362
              ||||||
Db       1119 ATATTG 1124

RESULT 18
US-09-471-275-1314
; Sequence 1314, Application US/09471275
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Contigs Obtained
;  TITLE OF INVENTION: From Various Libraries
;  FILE REFERENCE: 782
;  CURRENT APPLICATION NUMBER: US/09/471,275
;  CURRENT FILING DATE: 1999-12-23
;  EARLIER APPLICATION NUMBER: US 09/235,076
;  EARLIER FILING DATE: 1999-01-20
;  EARLIER APPLICATION NUMBER: US 09/234,611
;  EARLIER FILING DATE: 1999-01-22
;  EARLIER APPLICATION NUMBER: US 09/240,371
;  EARLIER FILING DATE: 1999-01-29
;  EARLIER APPLICATION NUMBER: US 09/277,227
;  EARLIER FILING DATE: 1999-03-25
;  EARLIER APPLICATION NUMBER: US 09/271,490
;  EARLIER FILING DATE: 1999-03-18
;  EARLIER APPLICATION NUMBER: US 09/293,972
;  EARLIER FILING DATE: 1999-04-15
;  EARLIER APPLICATION NUMBER: US 09/274,861
;  EARLIER FILING DATE: 1999-03-23
;  EARLIER APPLICATION NUMBER: US 60/125,453
;  EARLIER FILING DATE: 1999-03-19
;  EARLIER APPLICATION NUMBER: US 60/126,605
;  EARLIER FILING DATE: 1999-03-26
;  EARLIER APPLICATION NUMBER: US 09/306,350
;  EARLIER FILING DATE: 1999-05-07
;  EARLIER APPLICATION NUMBER: US 09/399,720
;  EARLIER FILING DATE: 1999-09-21
;  EARLIER APPLICATION NUMBER: US 09/404,284
;  EARLIER FILING DATE: 1999-09-21
;  EARLIER APPLICATION NUMBER:
;  EARLIER FILING DATE: 1999-12-16
```

```
;  NUMBER OF SEQ ID NOS: 10451
;  SOFTWARE: pt_CT_genes Version 1.0
; SEQ ID NO 1314
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-471-275-1314
```

Alignment Scores:

Pred. No.:	2.41e-291	Length:	1958
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	24	Gaps:	0

US-10-757-745-2 (1-362) x US-09-471-275-1314 (1-1958)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 110

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 170

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 230

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 290

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 350

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 410

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1130

Qy        361 IleLeu 362
              ||||||
Db       1131 ATATTG 1136
```

```
RESULT 19
US-09-488-725A-4772
; Sequence 4772, Application US/09488725A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/09/488,725A
;  CURRENT FILING DATE:  2000-12-22
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7144
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-488-725A-4772
```

Alignment Scores:

Pred. No.:	2.41e-291	Length:	1958
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	24	Gaps:	0

US-10-757-745-2 (1-362) x US-09-488-725A-4772 (1-1958)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 110
```

```
Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 170

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 230

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 290

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 350

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 410

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1130

Qy        361 IleLeu 362
               ||||||
Db       1131 ATATTG 1136

RESULT 20
```

```
US-09-488-725B-8295
; Sequence 8295, Application US/09488725B
; GENERAL INFORMATION:
;  APPLICANT: Yuanhua T. Tang
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Ankura Sinku
;  APPLICANT:  Chenghua Liu
;  APPLICANT:  Radoje T. Drmanac
;  TITLE OF INVENTION: Novel Contigs Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 784
;  CURRENT APPLICATION NUMBER: US/09/488,725B
;  CURRENT FILING DATE:  2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/004,182
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: US 09/034,341
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 09/045,400
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: US 09/321,214
;  PRIOR FILING DATE: 1999-05-26
;  PRIOR APPLICATION NUMBER: US 09/131,598
;  PRIOR FILING DATE: 1998-08-10
;  PRIOR APPLICATION NUMBER: US 09/170,294
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: US 09/179,473
;  PRIOR FILING DATE: 1998-10-27
;  PRIOR APPLICATION NUMBER: US 09/181,430
;  PRIOR FILING DATE: 1998-10-28
;  PRIOR APPLICATION NUMBER: US 09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  PRIOR APPLICATION NUMBER: US 09/234,611
;  PRIOR FILING DATE: 1999-01-22
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10289
;  SOFTWARE: pt_CT_genes Version 1.01
; SEQ ID NO 8295
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-488-725B-8295

Alignment Scores:
Pred. No.:              2.41e-291     Length:          1958
Score:                  1902.00       Matches:         362
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     24            Gaps:            0

US-10-757-745-2 (1-362) x US-09-488-725B-8295 (1-1958)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 110

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 170

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 230

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 290
```

```
Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 350

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 410

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1130

Qy        361 IleLeu 362
              ||||||
Db       1131 ATATTG 1136

RESULT 21
US-09-552-317-8295
; Sequence 8295, Application US/09552317
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Sinku, Ankura
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids
;  TITLE OF INVENTION:  and Polypeptides
;  FILE REFERENCE: 784CIP
;  CURRENT APPLICATION NUMBER: US/09/552,317
;  CURRENT FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 10289
;  SOFTWARE: pt_CT_genes Version 1.01
; SEQ ID NO 8295
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-552-317-8295

Alignment Scores:
Pred. No.:              2.41e-291       Length:         1958
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     27              Gaps:           0

US-10-757-745-2 (1-362) x US-09-552-317-8295 (1-1958)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 110

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 170

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 230

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 290

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 350

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 410

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770
```

```
Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1130

Qy        361 IleLeu 362
              ||||||
Db       1131 ATATTG 1136
```

RESULT 22
US-10-258-898A-4772
; Sequence 4772, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/258,898A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
; LENGTH: 1958
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-258-898A-4772

Alignment Scores:

Pred. No.:	2.41e-291	Length:	1958
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	43	Gaps:	0

US-10-757-745-2 (1-362) x US-10-258-898A-4772 (1-1958)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 110

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 170

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 230

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 290

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 350

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 410

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1130

Qy        361 IleLeu 362
              ||||||
```

```
Db       1131 ATATTG 1136

RESULT 23
US-10-286-897-4772
; Sequence 4772, Application US/10286897
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/286,897
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US/09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US/09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US/09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US/09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US/09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US/09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US/09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-286-897-4772

Alignment Scores:
Pred. No.:                2.41e-291      Length:          1958
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       43             Gaps:            0

US-10-757-745-2 (1-362) x US-10-286-897-4772 (1-1958)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 110

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 170

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 230

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 290

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 350

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 410

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1130

Qy        361 IleLeu 362
              ||||||
Db       1131 ATATTG 1136

RESULT 24
US-10-302-689A-59957
; Sequence 59957, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ballinger, Dennis
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Liu, Jin
;  APPLICANT:  Loeb, Deborah
;  APPLICANT:  Montgomery, Julia, R.
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
```

```
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 59957
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-302-689A-59957

Alignment Scores:
Pred. No.:                2.41e-291      Length:          1958
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       48             Gaps:            0

US-10-757-745-2 (1-362) x US-10-302-689A-59957 (1-1958)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 110

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 170

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 230

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 290

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 350

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 410

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1130

Qy        361 IleLeu 362
              ||||||
Db       1131 ATATTG 1136

RESULT 25
US-10-302-689A-77389
; Sequence 77389, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ballinger, Dennis
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Liu, Jin
;  APPLICANT:  Loeb, Deborah
;  APPLICANT:  Montgomery, Julia, R.
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
```

```
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 77389
;    LENGTH: 1958
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-302-689A-77389
```

Alignment Scores:

Pred. No.:	2.41e-291	Length:	1958
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	48	Gaps:	0

US-10-757-745-2 (1-362) x US-10-302-689A-77389 (1-1958)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 110

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 170

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 230

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 290

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 350

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 410

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1130

Qy        361 IleLeu 362
              ||||||
Db       1131 ATATTG 1136
```

```
RESULT 26
US-09-606-680-2868
; Sequence 2868, Application US/09606680
; GENERAL INFORMATION:
;  APPLICANT: Lloyd, Clare M.
;  APPLICANT:  Williamson, Mark
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 1600.1131-001
;  CURRENT APPLICATION NUMBER: US/09/606,680
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER: 60/141,227
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 60/141,226
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 4394
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2868
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-606-680-2868
```

Alignment Scores:

Pred. No.:	3.23e-291	Length:	2262
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	28	Gaps:	0

US-10-757-745-2 (1-362) x US-09-606-680-2868 (1-2262)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494
```

```
Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
Db        495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
Db        555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
Db        615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
Db       1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy        361 IleLeu 362
Db       1395 ATATTG 1400

RESULT 27
US-09-606-776-4085
; Sequence 4085, Application US/09606776
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  APPLICANT:  Myers, Paul
;  APPLICANT:  Gearing, David P.
;  APPLICANT:  Pan, Yang
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 1600.1129-001
;  CURRENT APPLICATION NUMBER: US/09/606,776
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER: 60/141,578
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 60/141,379
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: 60/141,138
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: 60/141,581
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 5415
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4085
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-606-776-4085

Alignment Scores:
Pred. No.:                3.23e-291      Length:          2262
Score:                    1902.00        Matches:         362
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       28             Gaps:            0

US-10-757-745-2 (1-362) x US-09-606-776-4085 (1-2262)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
Db        315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
Db        375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
Db        435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
Db        495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
Db        555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
Db        615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
```

```
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy        361 IleLeu 362
              ||||||
Db       1395 ATATTG 1400

RESULT 28
US-09-644-869-9498
; Sequence 9498, Application US/09644869
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  APPLICANT:  McCarthy, Sean A.
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Monahan, John
;  APPLICANT:  Richardson, Jennifer
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1182-001
;  CURRENT APPLICATION NUMBER: US/09/644,869
;  CURRENT FILING DATE:  2000-08-28
;  PRIOR APPLICATION NUMBER:  60/151,062
;  PRIOR FILING DATE:  1999-08-27
;  NUMBER OF SEQ ID NOS:  9708
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9498
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-644-869-9498

Alignment Scores:
Pred. No.:              3.23e-291       Length:         2262
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     29              Gaps:           0
```

US-10-757-745-2 (1-362) x US-09-644-869-9498 (1-2262)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy        361 IleLeu 362
              ||||||
Db       1395 ATATTG 1400

RESULT 29
US-09-652-121-6794
; Sequence 6794, Application US/09652121
; GENERAL INFORMATION:
;  APPLICANT:  DiStefano, Peter
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1188-001
;  CURRENT APPLICATION NUMBER: US/09/652,121
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,129
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  7615
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6794
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-121-6794

Alignment Scores:
Pred. No.:              3.23e-291      Length:          2262
Score:                  1902.00        Matches:         362
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            100.0%         Indels:          0
DB:                     29             Gaps:            0

US-10-757-745-2 (1-362) x US-09-652-121-6794 (1-2262)

Qy           1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy          21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy          41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494

Qy          61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy          81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy         101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy         121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy         221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy         241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy         261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy         281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy         301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy         321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy         341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy         361 IleLeu 362
               ||||||
Db        1395 ATATTG 1400

RESULT 30
US-09-652-122-3756
; Sequence 3756, Application US/09652122
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1194-001
;  CURRENT APPLICATION NUMBER: US/09/652,122
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,421
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  5020
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3756
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-122-3756

Alignment Scores:
Pred. No.:              3.23e-291      Length:          2262
Score:                  1902.00        Matches:         362
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            100.0%         Indels:          0
```

```
DB:                    29            Gaps:          0

US-10-757-745-2 (1-362) x US-09-652-122-3756 (1-2262)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy        361 IleLeu 362
              ||||||
Db       1395 ATATTG 1400

RESULT 31
US-09-652-128-9329
; Sequence 9329, Application US/09652128
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  APPLICANT:  Holtzman, Douglas A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1171-001
;  CURRENT APPLICATION NUMBER: US/09/652,128
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,133
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  10265
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9329
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-128-9329

Alignment Scores:
Pred. No.:              3.23e-291       Length:         2262
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     29              Gaps:           0

US-10-757-745-2 (1-362) x US-09-652-128-9329 (1-2262)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794
```

```
Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy        361 IleLeu 362
              ||||||
Db       1395 ATATTG 1400
```

RESULT 32
US-09-652-816-9572
; Sequence 9572, Application US/09652816
; GENERAL INFORMATION:
; APPLICANT: Gutierrez-Ramos, Jose-Carlos
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1177-001
; CURRENT APPLICATION NUMBER: US/09/652,816
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,111
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 9647
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9572
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2262)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-816-9572

Alignment Scores:

Pred. No.:	3.23e-291	Length:	2262
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	29	Gaps:	0

US-10-757-745-2 (1-362) x US-09-652-816-9572 (1-2262)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334
```

```
Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy        361 IleLeu 362
              ||||||
Db       1395 ATATTG 1400

RESULT 33
US-09-700-000-6419
; Sequence 6419, Application US/09700000
; GENERAL INFORMATION:
;  APPLICANT:  Richardson, Jennifer
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2022-001
;  CURRENT APPLICATION NUMBER: US/09/700,000
;  CURRENT FILING DATE:  2000-10-30
;  PRIOR APPLICATION NUMBER:  60/162,619
;  PRIOR FILING DATE:  1999-10-29
;  NUMBER OF SEQ ID NOS:  7171
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6419
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-700-000-6419

Alignment Scores:
Pred. No.:                3.23e-291       Length:          2262
Score:                    1902.00         Matches:         362
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       31              Gaps:            0

US-10-757-745-2 (1-362) x US-09-700-000-6419 (1-2262)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy        361 IleLeu 362
              ||||||
Db       1395 ATATTG 1400

RESULT 34
US-09-710-280-2045
; Sequence 2045, Application US/09710280
; GENERAL INFORMATION:
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Gearing, David P.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2007-001
;  CURRENT APPLICATION NUMBER: US/09/710,280
;  CURRENT FILING DATE:  2000-11-10
;  PRIOR APPLICATION NUMBER:  60/164,537
;  PRIOR FILING DATE:  1999-11-10
;  NUMBER OF SEQ ID NOS:  2763
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2045
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-710-280-2045

Alignment Scores:
```

Pred. No.:	3.23e-291	Length:	2262
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	31	Gaps:	0

US-10-757-745-2 (1-362) x US-09-710-280-2045 (1-2262)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy        361 IleLeu 362
              ||||||
Db       1395 ATATTG 1400
```

RESULT 35
US-09-710-282-3926
; Sequence 3926, Application US/09710282
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Kingsbury, Gillian A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2004-001
; CURRENT APPLICATION NUMBER: US/09/710,282
; CURRENT FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 60/164,256
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4662
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3926
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2262)
; OTHER INFORMATION: n = A,T,C or G
US-09-710-282-3926

Alignment Scores:

Pred. No.:	3.23e-291	Length:	2262
Score:	1902.00	Matches:	362
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	31	Gaps:	0

US-10-757-745-2 (1-362) x US-09-710-282-3926 (1-2262)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy        361 IleLeu 362
                 ||||||
Db       1395 ATATTG 1400

RESULT 36
US-09-726-802-2141
; Sequence 2141, Application US/09726802
; GENERAL INFORMATION:
;  APPLICANT:  Acton, Susan L.
;  APPLICANT:  Falb, Dean A.
;  APPLICANT:  .
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2025-001
;  CURRENT APPLICATION NUMBER: US/09/726,802
;  CURRENT FILING DATE:  2000-11-30
;  PRIOR APPLICATION NUMBER:  60/168,012
;  PRIOR FILING DATE:  1999-11-30
;  NUMBER OF SEQ ID NOS:  2872
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2141
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-726-802-2141

Alignment Scores:
Pred. No.:              3.23e-291       Length:         2262
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2 (1-362) x US-09-726-802-2141 (1-2262)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 374

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 434

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 494

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 554

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        555 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 614

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 674

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214
```

```
Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1394

Qy        361 IleLeu 362
              ||||||
Db       1395 ATATTG 1400

RESULT 37
US-09-981-353-88
; Sequence 88, Application US/09981353
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE: PA-0038 US
;  CURRENT APPLICATION NUMBER: US/09/981,353
;  CURRENT FILING DATE:  2001-10-11
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 232992.1
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

Alignment Scores:
Pred. No.:              6.34e-291       Length:         3152
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     38              Gaps:           0

US-10-757-745-2 (1-362) x US-09-981-353-88 (1-3152)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 1311

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 1371

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 1431

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 1491

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1492 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 1551

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 1611

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 1671

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1672 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 1731

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1732 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1791

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1792 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1851

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1852 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1911

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1912 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1971

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1972 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 2031

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2032 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 2091

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2092 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 2151

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2152 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 2211

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2212 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 2271

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2272 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 2331

Qy        361 IleLeu 362
              ||||||
Db       2332 ATATTG 2337

RESULT 38
US-60-239-841-88
; Sequence 88, Application US/60239841
; GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION:  GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE:  PA-0038 P
;  CURRENT APPLICATION NUMBER: US/60/239,841
;  CURRENT FILING DATE:  2000-10-11
;  NUMBER OF SEQ ID NOS:  194
;  SOFTWARE:  PERL Program
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 232992.1
```

```
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-60-239-841-88

Alignment Scores:
Pred. No.:              6.34e-291       Length:         3152
Score:                  1902.00         Matches:        362
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     77              Gaps:           0

US-10-757-745-2 (1-362) x US-60-239-841-88 (1-3152)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 1311

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 1371

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 1431

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 1491

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1492 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 1551

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 1611

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 1671

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1672 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 1731

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1732 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1791

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1792 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1851

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1852 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1911

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1912 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1971

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1972 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 2031

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2032 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 2091

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       2092 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 2151

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2152 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 2211

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2212 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 2271

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2272 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 2331

Qy        361 IleLeu 362
              ||||||
Db       2332 ATATTG 2337

RESULT 39
PCT-US00-05883-170
; Sequence 170, Application PC/TUS0005883
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT: Steve Ruben
;  TITLE OF INVENTION: Human Colon Cancer Associated Gene Sequences and Polypeptides
;  FILE REFERENCE: PA102PCT
;  CURRENT APPLICATION NUMBER: PCT/US00/05883
;  CURRENT FILING DATE: 2000-03-08
;  EARLIER APPLICATION NUMBER: 60/124,270
;  EARLIER FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
PCT-US00-05883-170

Alignment Scores:
Pred. No.:              2e-290          Length:         1296
Score:                  1894.00         Matches:        361
Percent Similarity:     99.7%           Conservative:   0
Best Local Similarity:  99.7%           Mismatches:     1
Query Match:            99.6%           Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2 (1-362) x PCT-US00-05883-170 (1-1296)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 83

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 143

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 203

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 263

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
```

```
Db        264 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 323

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        324 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 383

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        384 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 443

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        444 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 503

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db        504 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACA 563

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db        564 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 623

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220

Db        624 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTT 683

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240

Db        684 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 743

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260

Db        744 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 803

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280

Db        804 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 863

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        864 GTGGATGTCTGGGAGTTTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAA 923

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        924 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 983

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340

Db        984 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1043

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360

Db       1044 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1103

Qy        361 IleLeu 362

Db       1104 ATATTG 1109

RESULT 40
US-09-925-299-170
; Sequence 170, Application US/09925299
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:              2e-290          Length:         1296
Score:                  1894.00         Matches:        361
Percent Similarity:     99.7%           Conservative:   0
Best Local Similarity:  99.7%           Mismatches:     1
Query Match:            99.6%           Indels:         0
DB:                     36              Gaps:           0

US-10-757-745-2 (1-362) x US-09-925-299-170 (1-1296)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20

Db         24 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 83

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40

Db         84 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 143

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60

Db        144 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 203

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db        204 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 263

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        264 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 323

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        324 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 383

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        384 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 443

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        444 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 503

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db        504 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACA 563

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db        564 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 623

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220

Db        624 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTT 683

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240

Db        684 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 743

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 803

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 863

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 GTGGATGTCTGGGAGTTTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAA 923

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 983

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        984 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1043

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1103

Qy        361 IleLeu 362
                ||||||
Db       1104 ATATTG 1109
```

RESULT 41
US-09-629-469A-13209
; Sequence 13209, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1086)
US-09-629-469A-13209

Alignment Scores:
Pred. No.: 6.28e-290 Length: 1898

Score:	1893.00	Matches:	361
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	99.5%	Indels:	0
DB:	28	Gaps:	0

US-10-757-745-2 (1-362) x US-09-629-469A-13209 (1-1898)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1080

Qy        361 IleLeu 362
                 ||||||
Db       1081 ATATTG 1086

RESULT 42
US-10-917-503-13209
; Sequence 13209, Application US/10917503
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
;   LENGTH: 1898
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1086)
US-10-917-503-13209

Alignment Scores:
Pred. No.:              6.28e-290       Length:         1898
Score:                  1893.00         Matches:        361
Percent Similarity:     99.7%           Conservative:   0
Best Local Similarity:  99.7%           Mismatches:     1
Query Match:            99.5%           Indels:         0
DB:                     63              Gaps:           0

US-10-757-745-2 (1-362) x US-10-917-503-13209 (1-1898)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1080

Qy        361 IleLeu 362
              ||||||
Db       1081 ATATTG 1086

RESULT 43
US-09-705-256A-7708
; Sequence 7708, Application US/09705256A
```

```
; GENERAL INFORMATION:
;  APPLICANT: Ma, Xiao-Jun; Dotson, Stanton B.:  Monsanto Company
;  TITLE OF INVENTION: Tumor Associated Molecules (TAMs): Targets for diagnosis, treatme
;  TITLE OF INVENTION:  and prevention of cancer
;  FILE REFERENCE: 3214
;  CURRENT APPLICATION NUMBER: US/09/705,256A
;  CURRENT FILING DATE:  1999-05-11
;  PRIOR APPLICATION NUMBER: US 60/164,285
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 8259
; SEQ ID NO 7708
;   LENGTH: 1930
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-705-256A-7708

Alignment Scores:
Pred. No.:                2.85e-289       Length:          1930
Score:                    1889.00         Matches:         362
Percent Similarity:       99.7%           Conservative:    0
Best Local Similarity:    99.7%           Mismatches:      0
Query Match:              99.3%           Indels:          1
DB:                       31              Gaps:            0

US-10-757-745-2 (1-362) x US-09-705-256A-7708 (1-1930)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArg-ProGluThrIleSerGl 80
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGAGCCTGAAACCATCTCTGA 265

Qy         80 uProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIl 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 GCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAAT 325

Qy        100 eSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAs 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 CAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAA 385

Qy        120 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 TATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTT 445

Qy        140 uAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTy 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 AGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTA 505

Qy        160 rLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheTh 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCAC 565

Qy        180 rAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 AGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCC 625

Qy        200 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 685

Qy        220 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCA 745

Qy        240 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 GTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGC 805

Qy        260 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 AGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACAT 865

Qy        280 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGl 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 TGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACA 925

Qy        300 nMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePh 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 AATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTT 985

Qy        320 eArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGl 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 CAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGA 1045

Qy        340 uLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIl 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATAT 1105

Qy        360 eIleLeu 362
              |||||||
Db       1106 AATATTG 1112

RESULT 44
US-60-164-285-7708
; Sequence 7708, Application US/60164285
; GENERAL INFORMATION:
;  APPLICANT: Ma, Xiao-Jun
;  TITLE OF INVENTION: Tumor Associated Molecules (TAMs): Targets for diagnosis, treatme
;  FILE REFERENCE: 3214
;  CURRENT APPLICATION NUMBER: US/60/164,285
;  CURRENT FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 8259
; SEQ ID NO 7708
;   LENGTH: 1930
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-164-285-7708

Alignment Scores:
Pred. No.:                2.85e-289       Length:          1930
Score:                    1889.00         Matches:         362
Percent Similarity:       99.7%           Conservative:    0
Best Local Similarity:    99.7%           Mismatches:      0
Query Match:              99.3%           Indels:          1
DB:                       75              Gaps:            0

US-10-757-745-2 (1-362) x US-60-164-285-7708 (1-1930)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 85

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 145

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 205

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArg-ProGluThrIleSerGl 80
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        206 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGAGCCTGAAACCATCTCTGA 265

Qy         80 uProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIl 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        266 GCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAAT 325

Qy        100 eSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAs 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 CAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAA 385

Qy        120 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 TATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTT 445

Qy        140 uAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTy 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 AGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTA 505

Qy        160 rLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheTh 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCAC 565

Qy        180 rAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePr 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 AGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCC 625

Qy        200 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 685

Qy        220 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCA 745

Qy        240 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 GTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGC 805

Qy        260 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 AGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACAT 865

Qy        280 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGl 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 TGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACA 925

Qy        300 nMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePh 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 AATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTT 985

Qy        320 eArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGl 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 CAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGA 1045

Qy        340 uLysLeuAspCysGlyArgProProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIl 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATAT 1105

Qy        360 eIleLeu 362
              |||||||
Db       1106 AATATTG 1112

RESULT 45
US-09-585-944-2184
; Sequence 2184, Application US/09585944
; GENERAL INFORMATION:
;  APPLICANT: Hodgson, David M.
;  APPLICANT: Lincoln, Stephen E.
;  APPLICANT: Russo, Frank D.
;  APPLICANT: Spiro, Peter A.
;  APPLICANT: Banville, Steve C.
;  APPLICANT: Bratcher, Shawn R.
;  APPLICANT: Dufour, Gerard E.
;  APPLICANT: Cohen, Howard J.
;  APPLICANT: Rosen, Bruce
;  APPLICANT: Chalup, Michael S.
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Jones, Anissa L.
;  APPLICANT: Yu, Jimmy Y.
;  APPLICANT: Greenawalt, Lila B.
;  APPLICANT: Panzer, Scott R.
;  APPLICANT: Roseberry, Ann M.
;  APPLICANT: Wright, Rachel J.
;  APPLICANT: Daniels, Susan E.
;  TITLE OF INVENTION: SECRETORY MOLECULES
;  FILE REFERENCE: PT-0064 US
;  CURRENT APPLICATION NUMBER: US/09/585,944
;  CURRENT FILING DATE: 2000-06-02
;  EARLIER APPLICATION NUMBER: US 60/137,412
;  EARLIER FILING DATE: 1999-06-03
;  EARLIER APPLICATION NUMBER: US 60/147,542
;  EARLIER FILING DATE: 1999-08-05
;  EARLIER APPLICATION NUMBER: US 60/147,501
;  EARLIER FILING DATE: 1999-08-05
;  EARLIER APPLICATION NUMBER: US 60/147,500
;  EARLIER FILING DATE: 1999-08-05
;  NUMBER OF SEQ ID NOS: 2515
;  SOFTWARE: PERL Program
; SEQ ID NO 2184
;  LENGTH: 3080
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  OTHER INFORMATION: Incyte ID No: 232992.7
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 125, 1098
;  OTHER INFORMATION: a, t, c, g, or other
US-09-585-944-2184

Alignment Scores:
Pred. No.:                7.38e-289      Length:          3080
Score:                    1889.00        Matches:         362
Percent Similarity:       99.7%          Conservative:    0
Best Local Similarity:    99.7%          Mismatches:      0
Query Match:              99.3%          Indels:          1
DB:                       27             Gaps:            0

US-10-757-745-2 (1-362) x US-09-585-944-2184 (1-3080)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1179 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 1238

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1239 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 1298

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1299 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 1358

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1359 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 1418

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1419 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 1478

Qy        101 SerProSerGluAspThr-GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAs 120
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1479 AGCCCATCTGAAGATACTCCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAA 1538

Qy        120 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLe 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1539 TATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTT 1598

Qy        140 uAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTy 160
```

```
Db       1599 AGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTA 1658

Qy        160 rLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheTh 180

Db       1659 CCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCAC 1718

Qy        180 rAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePr 200

Db       1719 AGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCC 1778

Qy        200 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220

Db       1779 AAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCT 1838

Qy        220 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 240

Db       1839 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCA 1898

Qy        240 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 260

Db       1899 GTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGC 1958

Qy        260 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 280

Db       1959 AGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACAT 2018

Qy        280 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGl 300

Db       2019 TGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACA 2078

Qy        300 nMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePh 320

Db       2079 AATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTT 2138

Qy        320 eArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGl 340

Db       2139 CAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGA 2198

Qy        340 uLysLeuAspCysGlyArgProProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIl 360

Db       2199 AAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATAT 2258

Qy        360 eIleLeu 362

Db       2259 AATATTG 2265

RESULT 46
US-60-172-360-956/c
; Sequence 956, Application US/60172360
; GENERAL INFORMATION:
;  APPLICANT: Morris, MacDonald
;  APPLICANT: Lal, Preeti
;  APPLICANT: Diep, Dinh
;  TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
;  TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymorp
;  FILE REFERENCE: GX-0007 P
;  CURRENT APPLICATION NUMBER: US/60/172,360
;  CURRENT FILING DATE: 1999-12-16
;  NUMBER OF SEQ ID NOS: 29838
;  SOFTWARE: PERL Program
; SEQ ID NO 956
;  LENGTH: 3082
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  OTHER INFORMATION: Incyte ID No: 232992.7c
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 1983, 2958
;  OTHER INFORMATION: a, t, c, g, or other
US-60-172-360-956

Alignment Scores:
Pred. No.:              7.39e-289       Length:         3082
Score:                  1889.00         Matches:        362
Percent Similarity:     99.7%           Conservative:   0
Best Local Similarity:  99.7%           Mismatches:     0
Query Match:            99.3%           Indels:         1
DB:                     75              Gaps:           0

US-10-757-745-2 (1-362) x US-60-172-360-956 (1-3082)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20

Db       1902 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 1843

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40

Db       1842 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 1783

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60

Db       1782 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 1723

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db       1722 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 1663

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db       1662 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 1603

Qy        101 SerProSerGluAspThr-GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAs 120

Db       1602 AGCCCATCTGAAGATACTCCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAA 1543

Qy        120 nIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLe 140

Db       1542 TATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTT 1483

Qy        140 uAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTy 160

Db       1482 AGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTA 1423

Qy        160 rLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheTh 180

Db       1422 CCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCAC 1363

Qy        180 rAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePr 200

Db       1362 AGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCC 1303

Qy        200 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220

Db       1302 AAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCT 1243

Qy        220 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 240

Db       1242 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCA 1183

Qy        240 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 260

Db       1182 GTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGC 1123

Qy        260 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 280

Db       1122 AGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACAT 1063

Qy        280 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGl 300

Db       1062 TGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACA 1003

Qy        300 nMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePh 320
```

```
Db       1002 AATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTT 943

Qy        320 eArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGl 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 CAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGA 883

Qy        340 uLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIl 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 AAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATAT 823

Qy        360 eIleLeu 362
              |||||||
Db        822 AATATTG 816

RESULT 47
US-60-443-566-1795
; Sequence 1795, Application US/60443566
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001447
;  CURRENT APPLICATION NUMBER: US/60/443,566
;  CURRENT FILING DATE:  2003-01-30
;  NUMBER OF SEQ ID NOS: 25102
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1795
;   LENGTH: 1919
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-443-566-1795

Alignment Scores:
Pred. No.:              2.9e-285      Length:          1919
Score:                  1864.00       Matches:         356
Percent Similarity:     98.3%         Conservative:    0
Best Local Similarity:  98.3%         Mismatches:      6
Query Match:            98.0%         Indels:          0
DB:                     79            Gaps:            0

US-10-757-745-2 (1-362) x US-60-443-566-1795 (1-1919)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 77

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 137

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||  ||||||||||||  |||
Db        138 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCC 197

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 257

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 317

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 377

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 617

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        618 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 677

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 737

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        738 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 797

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        798 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 857

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 917

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 977

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1037

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1038 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1097

Qy        361 IleLeu 362
              ||||||
Db       1098 ATATTG 1103

RESULT 48
US-60-452-680-10426
; Sequence 10426, Application US/60452680
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  GRUPE, Andrew
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001450
;  CURRENT APPLICATION NUMBER: US/60/452,680
;  CURRENT FILING DATE:  2003-03-07
;  NUMBER OF SEQ ID NOS: 116213
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10426
;   LENGTH: 1919
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-452-680-10426

Alignment Scores:
Pred. No.:              2.9e-285      Length:          1919
Score:                  1864.00       Matches:         356
Percent Similarity:     98.3%         Conservative:    0
Best Local Similarity:  98.3%         Mismatches:      6
Query Match:            98.0%         Indels:          0
DB:                     79            Gaps:            0
```

```
US-10-757-745-2 (1-362) x US-60-452-680-10426 (1-1919)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 77

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 137

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||  ||||||||||||   |||
Db        138 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCC 197

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 257

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 317

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 377

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 617

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        618 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 677

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 737

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        738 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 797

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        798 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 857

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 917

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 977

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1037

Qy        341 LysLeuAspCysGlyArgProProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1038 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1097

Qy        361 IleLeu 362
              ||||||
Db       1098 ATATTG 1103

RESULT 49
US-60-455-444-3324
; Sequence 3324, Application US/60455444
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001455
;  CURRENT APPLICATION NUMBER: US/60/455,444
;  CURRENT FILING DATE:  2003-03-18
;  NUMBER OF SEQ ID NOS: 50986
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3324
;   LENGTH: 1919
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-455-444-3324

Alignment Scores:
Pred. No.:              2.9e-285      Length:          1919
Score:                  1864.00       Matches:         356
Percent Similarity:     98.3%         Conservative:    0
Best Local Similarity:  98.3%         Mismatches:      6
Query Match:            98.0%         Indels:          0
DB:                     79            Gaps:            0

US-10-757-745-2 (1-362) x US-60-455-444-3324 (1-1919)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 77

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 137

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||  ||||||||||||   |||
Db        138 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCC 197

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 257

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 317

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 377

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 617
```

```
Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        618 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 677

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 737

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db        738 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 797

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        798 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 857

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 917

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 977

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1037

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1038 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1097

Qy        361 IleLeu 362
              ||||||
Db       1098 ATATTG 1103

RESULT 50
US-60-465-241-3324
; Sequence 3324, Application US/60465241
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001468
;  CURRENT APPLICATION NUMBER: US/60/465,241
;  CURRENT FILING DATE:  2003-04-23
;  NUMBER OF SEQ ID NOS: 258418
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3324
;   LENGTH: 1919
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-465-241-3324

Alignment Scores:
Pred. No.:              2.9e-285      Length:          1919
Score:                  1864.00       Matches:         356
Percent Similarity:     98.3%         Conservative:    0
Best Local Similarity:  98.3%         Mismatches:      6
Query Match:            98.0%         Indels:          0
DB:                     79            Gaps:            0

US-10-757-745-2 (1-362) x US-60-465-241-3324 (1-1919)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 77

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 137

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||   |||
Db        138 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCC 197

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 257

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 317

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 377

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 617

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        618 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 677

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 737

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db        738 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 797

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        798 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 857

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 917

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 977

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1037

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1038 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1097

Qy        361 IleLeu 362
              ||||||
Db       1098 ATATTG 1103

RESULT 51
US-10-170-235-26834/c
; Sequence 26834, Application US/10170235
; GENERAL INFORMATION:
```

```
;  APPLICANT: VENTER, J. Craig
;  TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
;  TITLE OF INVENTION:  TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
;  FILE REFERENCE: CL001380
;  CURRENT APPLICATION NUMBER: US/10/170,235
;  CURRENT FILING DATE:  2003-03-17
;  NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 26834
;   LENGTH: 1849
;   TYPE: DNA
;   ORGANISM: HUMAN
US-10-170-235-26834

Alignment Scores:
Pred. No.:              3.37e-253       Length:         1849
Score:                  1664.00         Matches:        318
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            87.5%           Indels:         0
DB:                     42              Gaps:           0

US-10-757-745-2 (1-362) x US-10-170-235-26834 (1-1849)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 895

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 835

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 775

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 715

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 655

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 595

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 535

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 475

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 415

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 355

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 295

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 235

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        234 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 175

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 115

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 55

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 1

RESULT 52
US-60-443-566-1789/c
; Sequence 1789, Application US/60443566
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001447
;  CURRENT APPLICATION NUMBER: US/60/443,566
;  CURRENT FILING DATE:  2003-01-30
;  NUMBER OF SEQ ID NOS: 25102
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1789
;   LENGTH: 1849
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-443-566-1789

Alignment Scores:
Pred. No.:              4.23e-247       Length:         1849
Score:                  1626.00         Matches:        312
Percent Similarity:     98.1%           Conservative:   0
Best Local Similarity:  98.1%           Mismatches:     6
Query Match:            85.5%           Indels:         0
DB:                     79              Gaps:           0

US-10-757-745-2 (1-362) x US-60-443-566-1789 (1-1849)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 895

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 835

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCC 775

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 715

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 655

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 595

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 535

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        534 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 475

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        474 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 415

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 355

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db        354 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 295

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 235

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db        234 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 175

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        174 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 115

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 55

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 1

RESULT 53
US-60-452-680-10422/c
; Sequence 10422, Application US/60452680
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  GRUPE, Andrew
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001450
;  CURRENT APPLICATION NUMBER: US/60/452,680
;  CURRENT FILING DATE:  2003-03-07
;  NUMBER OF SEQ ID NOS: 116213
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10422
;   LENGTH: 1849
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-452-680-10422

Alignment Scores:
Pred. No.:              4.23e-247       Length:         1849
Score:                  1626.00         Matches:        312
Percent Similarity:     98.1%           Conservative:   0
Best Local Similarity:  98.1%           Mismatches:     6
Query Match:            85.5%           Indels:         0
DB:                     79              Gaps:           0

US-10-757-745-2 (1-362) x US-60-452-680-10422 (1-1849)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 895

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 835

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              |||||||||||||||||||||||||||||||||||||   ||||||||||||   |||

Db        834 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCC 775

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 715

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 655

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 595

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 535

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 475

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        474 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 415

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 355

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db        354 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 295

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 235

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db        234 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 175

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        174 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 115

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 55

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 1

RESULT 54
US-60-455-444-3323/c
; Sequence 3323, Application US/60455444
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001455
;  CURRENT APPLICATION NUMBER: US/60/455,444
;  CURRENT FILING DATE:  2003-03-18
;  NUMBER OF SEQ ID NOS: 50986
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3323
;   LENGTH: 1849
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-455-444-3323
```

Alignment Scores:

Pred. No.:	4.23e-247	Length:	1849
Score:	1626.00	Matches:	312
Percent Similarity:	98.1%	Conservative:	0
Best Local Similarity:	98.1%	Mismatches:	6
Query Match:	85.5%	Indels:	0
DB:	79	Gaps:	0

US-10-757-745-2 (1-362) x US-60-455-444-3323 (1-1849)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 895

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 835

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCC 775

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 715

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 655

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 595

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 535

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 475

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 415

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 355

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db        354 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 295

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 235

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        234 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 175

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        174 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 115

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 55

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 1
```

RESULT 55
US-60-465-241-3323/c
; Sequence 3323, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3323
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-465-241-3323

Alignment Scores:

Pred. No.:	4.23e-247	Length:	1849
Score:	1626.00	Matches:	312
Percent Similarity:	98.1%	Conservative:	0
Best Local Similarity:	98.1%	Mismatches:	6
Query Match:	85.5%	Indels:	0
DB:	79	Gaps:	0

US-10-757-745-2 (1-362) x US-60-465-241-3323 (1-1849)

```
Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 895

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 835

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCC 775

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 715

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 655

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 595

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 535

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 475

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 415

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 355

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db        354 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 295
```

```
Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 235

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              |||||||||||||||||||||||    |||||||||||||||||||||||||||||||||||
Db        234 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 175

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              |||||||||||||||||||||    |||||||||||||||||||||||||||||||||||
Db        174 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 115

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 55

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 1

RESULT 56
US-09-644-868-7772
; Sequence 7772, Application US/09644868
; GENERAL INFORMATION:
;  APPLICANT:  White, David
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Pan, Yang
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1168-001
;  CURRENT APPLICATION NUMBER: US/09/644,868
;  CURRENT FILING DATE:  2000-08-28
;  PRIOR APPLICATION NUMBER:  60/151,063
;  PRIOR FILING DATE:  1999-08-27
;  NUMBER OF SEQ ID NOS:  10075
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 7772
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-644-868-7772

Alignment Scores:
Pred. No.:              9.68e-235       Length:         1852
Score:                  1549.00         Matches:        309
Percent Similarity:     85.4%           Conservative:   0
Best Local Similarity:  85.4%           Mismatches:     7
Query Match:            81.4%           Indels:         47
DB:                     29              Gaps:           1

US-10-757-745-2 (1-362) x US-09-644-868-7772 (1-1852)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 92

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 152

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 212

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 272

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 332

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 392

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 452

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 512

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 572

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 632

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 692

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 752

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 812

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||
Db        813 GGAGATACAAATCTAAGGGATCGAGAGGG-AGN---------------------------- 844

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        844 ------------------------------------------------------------ 844

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        845 ------------------------------------------------NNNNNNNNN 853

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 NNAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 913

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        914 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTANATATA 973

Qy        361 IleLeu 362
              ||||||
Db        974 ATATTG 979

RESULT 57
US-09-649-161-8751
; Sequence 8751, Application US/09649161
; GENERAL INFORMATION:
;  APPLICANT:  Gutierrez-Ramos, Jose-Carlos
;  APPLICANT:  Villeval, Jean-Luc M.G.
;  APPLICANT:  Fraser, Christopher C.
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Hodge, Martin R.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1178-001
;  CURRENT APPLICATION NUMBER: US/09/649,161
```

```
;  CURRENT FILING DATE:  2000-08-28
;  PRIOR APPLICATION NUMBER:  60/151,060
;  PRIOR FILING DATE:  1999-08-27
;  NUMBER OF SEQ ID NOS:  10788
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8751
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-649-161-8751

Alignment Scores:
Pred. No.:                9.68e-235       Length:          1852
Score:                    1549.00         Matches:         309
Percent Similarity:       85.4%           Conservative:    0
Best Local Similarity:    85.4%           Mismatches:      7
Query Match:              81.4%           Indels:          47
DB:                       29              Gaps:            1

US-10-757-745-2 (1-362) x US-09-649-161-8751 (1-1852)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 92

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 152

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 212

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 272

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 332

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 392

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 452

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 512

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 572

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 632

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 692

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 752

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 812

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||||
Db        813 GGAGATACAAATCTAAGGGATCGAGAGGG-AGN---------------------------- 844

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        844 ---------------------------------------------------------- 844

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        845 ----------------------------------------------------NNNNNNNNN 853

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 NNAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 913

Qy        341 LysLeuAspCysGlyArgProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        914 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTANATATA 973

Qy        361 IleLeu 362
              ||||||
Db        974 ATATTG 979

RESULT 58
US-09-652-109-8165
; Sequence 8165, Application US/09652109
; GENERAL INFORMATION:
;  APPLICANT:  McCarthy, Sean A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1180-001
;  CURRENT APPLICATION NUMBER: US/09/652,109
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,128
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  10105
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8165
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-109-8165

Alignment Scores:
Pred. No.:                9.68e-235       Length:          1852
Score:                    1549.00         Matches:         309
Percent Similarity:       85.4%           Conservative:    0
Best Local Similarity:    85.4%           Mismatches:      7
Query Match:              81.4%           Indels:          47
DB:                       29              Gaps:            1

US-10-757-745-2 (1-362) x US-09-652-109-8165 (1-1852)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 92

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 152

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 212
```

```
Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 272

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 332

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 392

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 452

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 512

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 572

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 632

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 692

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 752

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 812

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
               ||||||||||||||||||||||||||
Db        813 GGAGATACAAATCTAAGGGATCGAGAGGG-AGN--------------------------- 844

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        844 ------------------------------------------------------------ 844

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        845 -----------------------------------------------------NNNNNNNNN 853

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 NNAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 913

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        914 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTANATATA 973

Qy        361 IleLeu 362
               ||||||
Db        974 ATATTG 979

RESULT 59
US-09-652-128-7441
; Sequence 7441, Application US/09652128
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  APPLICANT:  Holtzman, Douglas A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1171-001
;  CURRENT APPLICATION NUMBER: US/09/652,128
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,133
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  10265
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 7441
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-128-7441

Alignment Scores:
Pred. No.:              9.68e-235       Length:         1852
Score:                  1549.00         Matches:        309
Percent Similarity:     85.4%           Conservative:   0
Best Local Similarity:  85.4%           Mismatches:     7
Query Match:            81.4%           Indels:         47
DB:                     29              Gaps:           1

US-10-757-745-2 (1-362) x US-09-652-128-7441 (1-1852)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 92

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 152

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 212

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 272

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 332

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 392

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 452

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 512

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 572

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 632

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 692

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        693 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 752

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 812

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||
Db        813 GGAGATACAAATCTAAGGGATCGAGAGGG-AGN---------------------------- 844

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        844 ------------------------------------------------------------ 844

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        845 -----------------------------------------------------NNNNNNNNN 853

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 NNAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 913

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db        914 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTANATATA 973

Qy        361 IleLeu 362
              ||||||
Db        974 ATATTG 979

RESULT 60
US-09-652-816-8491
; Sequence 8491, Application US/09652816
; GENERAL INFORMATION:
;  APPLICANT:  Gutierrez-Ramos, Jose-Carlos
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1177-001
;  CURRENT APPLICATION NUMBER: US/09/652,816
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER:  60/152,111
;  PRIOR FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  9647
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8491
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-816-8491

Alignment Scores:
Pred. No.:              9.68e-235     Length:          1852
Score:                  1549.00       Matches:         309
Percent Similarity:     85.4%         Conservative:    0
Best Local Similarity:  85.4%         Mismatches:      7
Query Match:            81.4%         Indels:          47
DB:                     29            Gaps:            1

US-10-757-745-2 (1-362) x US-09-652-816-8491 (1-1852)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 92

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 152

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 212

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 272

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 332

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 392

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 452

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 512

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 572

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 632

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 692

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 752

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 812

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              ||||||||||||||||||||||||
Db        813 GGAGATACAAATCTAAGGGATCGAGAGGG-AGN---------------------------- 844

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        844 ------------------------------------------------------------ 844

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        845 -----------------------------------------------------NNNNNNNNN 853

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 NNAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 913

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db        914 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTANATATA 973

Qy        361 IleLeu 362
              ||||||
Db        974 ATATTG 979

RESULT 61
US-09-652-914-8450
; Sequence 8450, Application US/09652914
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
```

```
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1193-001
;  CURRENT APPLICATION NUMBER: US/09/652,914
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER:  60/152,112
;  PRIOR FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  9677
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8450
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-914-8450

Alignment Scores:
Pred. No.:              9.68e-235      Length:          1852
Score:                  1549.00        Matches:         309
Percent Similarity:     85.4%          Conservative:    0
Best Local Similarity:  85.4%          Mismatches:      7
Query Match:            81.4%          Indels:          47
DB:                     29             Gaps:            1

US-10-757-745-2 (1-362) x US-09-652-914-8450 (1-1852)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 92

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 152

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 GCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 212

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 272

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 332

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 392

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 452

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 512

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 572

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 632

Qy        201 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 692

Qy        221 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 752

Qy        241 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 812

Qy        261 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 280
              |||||||||||||||||||||||||
Db        813 GGAGATACAAATCTAAGGGATCGAGAGGG-AGN---------------------------- 844

Qy        281 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 300

Db        844 ------------------------------------------------------------ 844

Qy        301 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 320

Db        845 ----------------------------------------------------NNNNNNNNN 853

Qy        321 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 NNAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 913

Qy        341 LysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeuCysAsnLeuAspIle 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db        914 AAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTANATATA 973

Qy        361 IleLeu 362
              ||||||
Db        974 ATATTG 979

RESULT 62
US-10-757-745-3
; Sequence 3, Application US/10757745
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (122)..(1234)
;   OTHER INFORMATION:
US-10-757-745-3

Alignment Scores:
Pred. No.:              2.42e-188      Length:          1312
Score:                  1258.00        Matches:         240
Percent Similarity:     80.6%          Conservative:    43
Best Local Similarity:  68.4%          Mismatches:      64
Query Match:            66.1%          Indels:          4
DB:                     61             Gaps:            3

US-10-757-745-2 (1-362) x US-10-757-745-3 (1-1312)

Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31
              |||||||||:::   |||  |||      ||||||:::|||||||||   |||:::
Db        191 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 247

Qy         32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51
```

```
Db        248 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 307

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71

Db        308 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 367

Qy         72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91

Db        368 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 427

Qy         92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111

Db        428 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 484

Qy        112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131

Db        485 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 544

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151

Db        545 AGGGCTCGAGGGGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 604

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171

Db        605 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 664

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191

Db        665 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTT 724

Qy        192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211

Db        725 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 784

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231

Db        785 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 844

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251

Db        845 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 904

Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271

Db        905 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 964

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291

Db        965 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 1024

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311

Db       1025 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1084

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331

Db       1085 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1138

Qy        332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHis 351

Db       1139 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1198

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362

Db       1199 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1231

RESULT 63
US-10-144-771-826
; Sequence 826, Application US/10144771
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig
;  TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
;  FILE REFERENCE: CL001321
;  CURRENT APPLICATION NUMBER: US/10/144,771
;  CURRENT FILING DATE:  2002-05-15
;  NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 826
;   LENGTH: 1944
;   TYPE: DNA
;   ORGANISM: HUMAN
US-10-144-771-826

Alignment Scores:
Pred. No.:                5.4e-188      Length:           1944
Score:                    1258.00       Matches:          240
Percent Similarity:       80.6%         Conservative:     43
Best Local Similarity:    68.4%         Mismatches:       64
Query Match:              66.1%         Indels:           4
DB:                       42            Gaps:             3

US-10-757-745-2 (1-362) x US-10-144-771-826 (1-1944)

Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31

Db        117 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 173

Qy         32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51

Db        174 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 233

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71

Db        234 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 293

Qy         72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91

Db        294 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 353

Qy         92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111

Db        354 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 410

Qy        112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131

Db        411 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 470

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151

Db        471 AGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 530

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171

Db        531 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 590

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191

Db        591 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTT 650

Qy        192 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 211

Db        651 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 710

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231

Db        711 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 770

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251

Db        771 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 830

Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271

Db        831 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 890

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
```

```
                :::|||||||||||||||:::|||:::   |||   |||||||||||||||||||||||
Db        891 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 950

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
              ||||||||||||||||||||||||:::   |||:::||||||   |||   ||||||
Db        951 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1010

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              |||   |||||||||||||||||||||||   |||||||||||||||:::||||||
Db       1011 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1064

Qy        332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
              :::|||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db       1065 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1124

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||||||   |||:::::::::|||
Db       1125 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1157

RESULT 64
US-60-360-207-826
; Sequence 826, Application US/60360207
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig
;  TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
;  FILE REFERENCE: CL001321
;  CURRENT APPLICATION NUMBER: US/60/360,207
;  CURRENT FILING DATE:  2002-03-01
;  NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 826
;   LENGTH: 1944
;   TYPE: DNA
;   ORGANISM: HUMAN
US-60-360-207-826

Alignment Scores:
Pred. No.:                5.4e-188      Length:          1944
Score:                    1258.00       Matches:         240
Percent Similarity:       80.6%         Conservative:    43
Best Local Similarity:    68.4%         Mismatches:      64
Query Match:              66.1%         Indels:          4
DB:                       78            Gaps:            3

US-10-757-745-2 (1-362) x US-60-360-207-826 (1-1944)

Qy         12 GluAlaAlaGluGluGluGlyGluProGluValLysLysArgArgLeuLeuCysValGlu 31
              |||||||||:::   |||   |||      ||||||:::|||||||||   |||:::
Db        117 GAAGCAGCACAGGCGGAG---GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 173

Qy         32 PheAlaSerValAlaSerCysAspAlaAlaValAlaGlnCysPheLeuAlaGluAsnAsp 51
              ||||||   |||      ||||||      :::          |||   |||||||||
Db        174 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 233

Qy         52 TrpGluMetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeu 71
              |||:::   ::::::||||||::::::|||||||||   |||      :::
Db        234 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 293

Qy         72 GluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu 91
              |||:::|||   |||      :::   :::   ||||||||||||||||||||:::
Db        294 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 353

Qy         92 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 111
                 |||:::|||      :::   |||||||||      |||      |||:::
Db        354 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCT---GGAACTCCTCTAGAAGATAGC 410

Qy        112 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 131
              |||      |||   ||||||||||||||||||||||||||||      ||||||   |||
Db        411 AGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAG 470

Qy        132 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 151
              ||||||||||||||||||   |||||||||||||||||||||||:::|||||||||
Db        471 AGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 530

Qy        152 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 171
              ||||||||||||||||||   :::||||||||||||||:::::::::|||   ||||||
Db        531 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 590

Qy        172 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 191
              ||||||:::|||||||||||||||||||||||||:::|||||||||   |||||||||
Db        591 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTT 650

Qy        192 LysSerGlnGluIleIleProProSerProSerThrLysMetMetArgAsnLeuLeuCysVal 211
              |||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db        651 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 710

Qy        212 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 231
              :::|||::::::   |||||||||   ||||||||||||||||||||||||||||||||
Db        711 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 770

Qy        232 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 251
                 |||:::|||||||||:::   |||||||||   ||||||   ||||||||||||||||
Db        771 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 830

Qy        252 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 271
              |||:::|||   ||||||||||||||||||||||||||||||||||||:::||||||
Db        831 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 890

Qy        272 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 291
              :::|||||||||||||||:::|||:::   |||   |||||||||||||||||||||||
Db        891 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 950

Qy        292 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 311
              ||||||||||||||||||||||||:::   |||:::||||||   |||   ||||||
Db        951 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1010

Qy        312 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 331
              |||   |||||||||||||||||||||||   |||||||||||||||:::||||||
Db       1011 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1064

Qy        332 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHis 351
              :::|||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db       1065 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1124

Qy        352 TrpGlyLeuLeuCysAsnLeuAspIleIleLeu 362
              ||||||||||||||||   |||:::::::::|||
Db       1125 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTG 1157

RESULT 65
PCT-US01-08631-22058
; Sequence 22058, Application PC/TUS0108631
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-049
;  CURRENT APPLICATION NUMBER: PCT/US01/08631
;  CURRENT FILING DATE:  2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22058
;   LENGTH: 1088
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1)..(687)
;   OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=1199.
PCT-US01-08631-22058

Alignment Scores:
Pred. No.:                4.88e-179     Length:          1088
```

Score:	1199.00	Matches:	226
Percent Similarity:	99.1%	Conservative:	1
Best Local Similarity:	98.7%	Mismatches:	2
Query Match:	63.0%	Indels:	0
DB:	1	Gaps:	0

US-10-757-745-2 (1-362) x PCT-US01-08631-22058 (1-1088)

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy        333 SerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHisTrp 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 660

Qy        353 GlyLeuLeuCysAsnLeuAspIleIle 361
              |||||||||||||||||||||||||||
Db        661 GGTCTTCTGTGCAACTTAGATATAATA 687
```

RESULT 66
US-10-302-689A-136240
; Sequence 136240, Application US/10302689A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Asundi, Vinod
; APPLICANT: Ballinger, Dennis
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; APPLICANT: Loeb, Deborah
; APPLICANT: Montgomery, Julia, R.
; APPLICANT: Pace, Ann M.
; APPLICANT: Sheridan, James P.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 502CIP
; CURRENT APPLICATION NUMBER: US/10/302,689A
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 10/084,643
; PRIOR FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: 09/989,660
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 10/014,487
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 09/952,981
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/905,059
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 09/898,888
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 09/919,002
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 158931
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 136240
; LENGTH: 1088
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-302-689A-136240

Alignment Scores:

Pred. No.:	4.88e-179	Length:	1088
Score:	1199.00	Matches:	226
Percent Similarity:	99.1%	Conservative:	1
Best Local Similarity:	98.7%	Mismatches:	2
Query Match:	63.0%	Indels:	0
DB:	48	Gaps:	0

US-10-757-745-2 (1-362) x US-10-302-689A-136240 (1-1088)

```
Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480
```

```
Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy        333 SerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHisTrp 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 660

Qy        353 GlyLeuLeuCysAsnLeuAspIleIle 361
              |||||||||||||||||||||||||||
Db        661 GGTCTTCTGTGCAACTTAGATATAATA 687

RESULT 67
US-10-450-763-22058
; Sequence 22058, Application US/10450763
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22058
;   LENGTH: 1088
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1)..(687)
;   OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=1199.
US-10-450-763-22058

Alignment Scores:
Pred. No.:              4.88e-179       Length:         1088
Score:                  1199.00         Matches:        226
Percent Similarity:     99.1%           Conservative:   1
Best Local Similarity:  98.7%           Mismatches:     2
Query Match:            63.0%           Indels:         0
DB:                     51              Gaps:           0

US-10-757-745-2 (1-362) x US-10-450-763-22058 (1-1088)

Qy        133 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy        153 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy        173 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy        193 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy        213 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        233 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        253 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        273 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy        293 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy        313 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy        333 SerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgProSerAspHisTrp 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 660

Qy        353 GlyLeuLeuCysAsnLeuAspIleIle 361
              |||||||||||||||||||||||||||
Db        661 GGTCTTCTGTGCAACTTAGATATAATA 687

RESULT 68
US-09-629-469A-4908
; Sequence 4908, Application US/09629469A
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/09/629,469A
;  CURRENT FILING DATE:  2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4908
;   LENGTH: 752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (685)..(685)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (694)..(694)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (740)..(740)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-09-629-469A-4908

Alignment Scores:
Pred. No.:              1.23e-176       Length:         752
Score:                  1182.00         Matches:        238
Percent Similarity:     94.8%           Conservative:   1
Best Local Similarity:  94.4%           Mismatches:     10
Query Match:            62.1%           Indels:         3
DB:                     28              Gaps:           0

US-10-757-745-2 (1-362) x US-09-629-469A-4908 (1-752)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy        201 SerThrLysMetMetArg-AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 660

Qy        220 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 240
              |||||||   ||||||   |||    :::|||   ||||||       ||||||||||||||
Db        661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy        240 lnLeuLysMetValLeu-LysLysMetGln 249
              ||    ||||||   |||    |||||||||
Db        721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750

RESULT 69

US-10-917-503-4908
; Sequence 4908, Application US/10917503
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4908
;   LENGTH: 752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (685)..(685)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (694)..(694)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (740)..(740)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-10-917-503-4908

Alignment Scores:
Pred. No.:              1.23e-176       Length:         752
Score:                  1182.00         Matches:        238
Percent Similarity:     94.8%           Conservative:   1
Best Local Similarity:  94.4%           Mismatches:     10
Query Match:            62.1%           Indels:         3
DB:                     63              Gaps:           0

US-10-757-745-2 (1-362) x US-10-917-503-4908 (1-752)

Qy          1 MetGluLeuGlySerCysLeuGluGlyGlyArgGluAlaAlaGluGluGluGlyGluPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCT 60

Qy         21 GluValLysLysArgArgLeuLeuCysValGluPheAlaSerValAlaSerCysAspAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCC 120

Qy         41 AlaValAlaGlnCysPheLeuAlaGluAsnAspTrpGluMetGluArgAlaLeuAsnSer 60
```

```
Db        121 GCGGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCC 180

Qy         61 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGlu 80

Db        181 TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAG 240

Qy         81 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIle 100

Db        241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 300

Qy        101 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 120

Db        301 AGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360

Qy        121 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 140

Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy        141 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 160

Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy        161 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 180

Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy        181 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 200

Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTCCA 600

Qy        201 SerThrLysMetMetArg-AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 220

Db        601 AGTACCAAAATGATGAGAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 660

Qy        220 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 240

Db        661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy        240 lnLeuLysMetValLeu-LysLysMetGln 249

Db        721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750
```

RESULT 70
US-09-248-178-63/c
; Sequence 63, Application US/09248178
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF BREAST CANCER AND
; TITLE OF INVENTION: METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C2
; CURRENT APPLICATION NUMBER: US/09/248,178
; CURRENT FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-248-178-63

Alignment Scores:

Pred. No.:	1.63e-175	Length:	1079
Score:	1177.00	Matches:	226
Percent Similarity:	99.6%	Conservative:	0
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	61.9%	Indels:	1
DB:	22	Gaps:	0

US-10-757-745-2 (1-362) x US-09-248-178-63 (1-1079)

```
Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155

Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175

Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195

Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215

Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235

Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255

Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275

Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295

Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315

Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335

Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355

Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362

Db        418 TGCAACTTAGATATAATATTG 398
```

RESULT 71
US-09-288-950-63/c
; Sequence 63, Application US/09288950
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C3
; CURRENT APPLICATION NUMBER: US/09/288,950
; CURRENT FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-288-950-63

Alignment Scores:

Pred. No.:	1.63e-175	Length:	1079
Score:	1177.00	Matches:	226
Percent Similarity:	99.6%	Conservative:	0
Best Local Similarity:	99.6%	Mismatches:	1

```
Query Match:            61.9%       Indels:        1
DB:                     22          Gaps:          0

US-10-757-745-2 (1-362) x US-09-288-950-63 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398

RESULT 72
US-09-346-327-63/c
; Sequence 63, Application US/09346327
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Xu, Jiangchun
;  APPLICANT: Dillon, Davin C.
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
;  TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.446C4
;  CURRENT APPLICATION NUMBER: US/09/346,327
;  CURRENT FILING DATE: 1999-07-02
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-346-327-63

Alignment Scores:
Pred. No.:              1.63e-175   Length:        1079
Score:                  1177.00     Matches:       226
Percent Similarity:     99.6%       Conservative:  0
Best Local Similarity:  99.6%       Mismatches:    1
Query Match:            61.9%       Indels:        1
DB:                     23          Gaps:          0

US-10-757-745-2 (1-362) x US-09-346-327-63 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398

RESULT 73
US-09-687-507-63/c
; Sequence 63, Application US/09687507
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Retter, Marc W.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE:  210121.446C6
;  CURRENT APPLICATION NUMBER: US/09/687,507
;  CURRENT FILING DATE:  2000-10-12
;  NUMBER OF SEQ ID NOS:  116
;  SOFTWARE:  FastSEQ for Windows Version 3.0
```

; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-687-507-63

Alignment Scores:

Pred. No.:	1.63e-175	Length:	1079
Score:	1177.00	Matches:	226
Percent Similarity:	99.6%	Conservative:	0
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	61.9%	Indels:	1
DB:	29	Gaps:	0

US-10-757-745-2 (1-362) x US-09-687-507-63 (1-1079)

```
Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398
```

RESULT 74
US-09-687-507A-63/c
; Sequence 63, Application US/09687507A
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C6
; CURRENT APPLICATION NUMBER: US/09/687,507A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-687-507A-63

Alignment Scores:

Pred. No.:	1.63e-175	Length:	1079
Score:	1177.00	Matches:	226
Percent Similarity:	99.6%	Conservative:	0
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	61.9%	Indels:	1
DB:	29	Gaps:	0

US-10-757-745-2 (1-362) x US-09-687-507A-63 (1-1079)

```
Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
Db        418 TGCAACTTAGATATAATATTG 398
```

RESULT 75
US-09-745-288-63/c
; Sequence 63, Application US/09745288

```
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  TITLE OF INVENTION:  COMPOUNDS FOR IMMUNOTHERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; . FILE REFERENCE:  210121.446D1
;  CURRENT APPLICATION NUMBER: US/09/745,288
;  CURRENT FILING DATE:  2000-12-19
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-745-288-63

Alignment Scores:
Pred. No.:                1.63e-175       Length:          1079
Score:                    1177.00         Matches:         226
Percent Similarity:       99.6%           Conservative:    0
Best Local Similarity:    99.6%           Mismatches:      1
Query Match:              61.9%           Indels:          1
DB:                       31              Gaps:            0

US-10-757-745-2 (1-362) x US-09-745-288-63 (1-1079)

Qy        136 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        156 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        176 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        196 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        216 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        236 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        256 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        276 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        296 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        316 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        336 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPheProSerAspHisTrpGlyLeuLeu 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTG 419

Qy        356 CysAsnLeuAspIleIleLeu 362
              |||||||||||||||||||||
```

```
Db        418 TGCAACTTAGATATAATATTG 398

Search completed: March 11, 2006, 12:08:30
Job time : 14992.7 secs
```

```
Db      34768 AAGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 34709

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81

Db      34708 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTGTGATGAAATAGACTTTAAACCA 34649

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98

Db      34648 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 34589

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117

Db      34588 GAAGAGCTAACTATCCTAAATGTATATGCACCCAATACA---------GGAGCATCCAGA 34538

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137

Db      34537 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 34484

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145

Db      34483 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 34427

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163

Db      34426 GTTAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGTGGACCTAATAGACATC 34367

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177

Db      34366 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 34328

RESULT 59
US-10-868-184C-10277/c
; Sequence 10277, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10277
;   LENGTH: 49375
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10277

Alignment Scores:
Pred. No.:              234          Length:          49375
Score:                  86.50        Matches:         51
Percent Similarity:     43.4%        Conservative:    45
Best Local Similarity:  23.1%        Mismatches:      80

Query Match:            7.3%         Indels:          45
DB:                     6            Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10277 (1-49375)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21

Db      47846 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAACACACAGACTGGCAAATTGGATAAAG 47787

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41

Db      47786 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 47745

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61

Db      47744 ACATGCAGAGACAAACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 47685

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81

Db      47684 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCGGATAAAACAGACTTTAAACCA 47625

Qy         82 ThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnValSerGly 97

Db      47624 ACAAAGATCAAAAGAGACAAAGAAGGCCATTATATAATGGTAAAGGCATCAGTTCAACAA 47565

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117

Db      47564 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 47514

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137

Db      47513 TTC------ATAAAGCAAGTCCTTAGAGACCTACAAAAAGACTTAGACTCCCACACAATA 47460

Qy        138 IlePheAlaGlyAsp------------------------ThrAsnLeuArgAspArgGl 149

Db      47459 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACAAGACAGAAA 47403

Qy        149 uValThrArgCysGly-GlyLeu---------------ProAsn-AsnIleValAspVal 163

Db      47402 GTTAACAAGGATATCCTGGAATGGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATA 47343

Qy        164 ----------------------------TrpGluPheLeuGlyLysProLysHisCys 173

Db      47342 CACAGAACTCTCCACCCCAAATCAACAGAATATACATTCTTCTCAGCACCA---CATTGC 47286

Qy        174 GlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192

Db      47285 ACTTATTCCAAAATTGACCACATAGTTGAAAATAAAGCACTTCTCAGCAAATGTAAA 47229

RESULT 60
US-10-868-184D-10277/c
; Sequence 10277, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
```

```
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10277
;   LENGTH: 49375
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10277

Alignment Scores:
Pred. No.:                234             Length:          49375
Score:                    86.50           Matches:         51
Percent Similarity:       43.4%           Conservative:    45
Best Local Similarity:    23.1%           Mismatches:      80
Query Match:              7.3%            Indels:          45
DB:                       6               Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10277 (1-49375)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::     :::     :::  ::::::::::
Db      47846 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAACACACAGACTGGCAAATTGGATAAAG 47787

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                 |||  |||      :::||||||                  ::::::|||
Db      47786 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 47745

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                ||      :::         |||  |||      :::   |||
Db      47744 ACATGCAGAGACAAACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 47685

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                 ||||||:::   |||  :::     :::              |||
Db      47684 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCGGATAAAACAGACTTTAAACCA 47625

Qy         82 ThrLysMetMetArg----------AsnLeuLeuCysValHisValAsnValSerGly 97
              ||||||:::   |||          :::  :::   |||      :::|||
Db      47624 ACAAAGATCAAAAGAGACAAAGAAGGCCATTATATAATGGTAAAGGCATCAGTTCAACAA 47565

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                ||||||   ::::::    :::     :::|||           |||   |||
Db      47564 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 47514

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                :::|||   ||||||:::   :::|||:::      :::|||   |||:::
Db      47513 TTC------ATAAAGCAAGTCCTTAGAGACCTACAAAAAGACTTAGACTCCCACACAATA 47460

Qy        138 IlePheAlaGlyAsp------------------------ThrAsnLeuArgAspArgGl 149
              |||      ||||||                      |||:::   :::||||||::
Db      47459 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACAAGACAGAAA 47403

Qy        149 uValThrArgCysGly-GlyLeu--------------ProAsn-AsnIleValAspVal 163
              ::::||||||     |||:::              |||::: :::::::::|||:::
Db      47402 GTTAACAAGGATATCCTGGAATGGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATA 47343

Qy        164 ----------------------------TrpGluPheLeuGlyLysProLysHisCys 173
                                         :::   |||         |||   ||||||
Db      47342 CACAGAACTCTCCACCCCAAATCAACAGAATATACATTCTTCTCAGCACCA---CATTGC 47286

Qy        174 GlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
                |||:::      :::         |||     :::  :::     ||||||
Db      47285 ACTTATTCCAAAATTGACCACATAGTTGAAAATAAAGCACTTCTCAGCAAATGTAAA 47229

RESULT 61
US-11-045-004-1/c
; Sequence 1, Application US/11045004
; GENERAL INFORMATION:
;  APPLICANT: BUCHRIESER, CARMEN
;  APPLICANT:  FRANGEUL, LIONEL
;  APPLICANT:  COUVE, ELISABETH
;  APPLICANT:  RUSNIOK, CHRISTOPHE
;  APPLICANT:  FSIHI, HAFIDA
;  APPLICANT:  DEHOUX, PIERRE
;  APPLICANT:  DUSSURGET, OLIVIER
;  APPLICANT:  CHETOUANI, FARID
;  APPLICANT:  NEDJARI, HAFED
;  APPLICANT:  GLASER, PHILIPPE
;  APPLICANT:  KUNST, FRANCK
;  APPLICANT:  COSSART, PASCALE
;  APPLICANT:  DANIELS, JUSTIN
;  APPLICANT:  GOEBBEI, WERNER
;  APPLICANT:  KREFT, JURGEN
;  APPLICANT:  KUHN, MICHAEL
;  APPLICANT:  NG, EVA
;  APPLICANT:  VAZQUEZ-BOLAND, ANTONIO
;  APPLICANT:  DOMINGUEZ-BERNAL, GUSTAVO
;  APPLICANT:  GARRIDO-GARCIA, PATRICIA
;  APPLICANT:  TIERREZ-MARTINEZ, ALBERTO
;  APPLICANT:  AMEND, ALEXANDRA
;  APPLICANT:  CHAKRABORTY, TRINAD
;  APPLICANT:  DOMANN, EUGEN
;  APPLICANT:  HAIN, THORSTEN
;  APPLICANT:  BERCHE, PATRICK
;  APPLICANT:  CHARBIT, ALAIN
;  APPLICANT:  DURANT, LIONEL
;  APPLICANT:  PEREZ-DIAZ, JOSE-CLAUDIO
;  APPLICANT:  BAQUERO, FERNANDO
;  APPLICANT:  GARCIA DEL PORTILLO, FRANCISCO
;  APPLICANT:  GOMEZ-LOPEZ, NURIA
;  APPLICANT:  MADUENIO, ENCARNA
;  APPLICANT:  PABLOS, BETRIZ DE
;  APPLICANT:  WEHLAND, JURGEN
;  APPLICANT:  KARST, UWE
;  APPLICANT:  ENTIAN, KARL-DIETER
;  APPLICANT:  HAUF, JORG
;  APPLICANT:  ROSE, MATTHIAS
;  APPLICANT:  VOSS, HAMUT
;  TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
;  FILE REFERENCE: 05394.0018-02
;  CURRENT APPLICATION NUMBER: US/11/045,004
;  CURRENT FILING DATE:  2005-01-28
;  PRIOR APPLICATION NUMBER: 10/637,657
;  PRIOR FILING DATE: 2003-08-11
;  PRIOR APPLICATION NUMBER: 10/257,023
;  PRIOR FILING DATE: 2002-10-08
;  PRIOR APPLICATION NUMBER: PCT/FR01/01118
;  PRIOR FILING DATE: 2001-04-11
;  PRIOR APPLICATION NUMBER: FR 00/04,629
;  PRIOR FILING DATE: 2000-04-11
;  NUMBER OF SEQ ID NOS: 2854
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
;   LENGTH: 2944528
;   TYPE: DNA
;   ORGANISM: Listeria monocytogenes
US-11-045-004-1

Alignment Scores:
Pred. No.:                1.62e+05        Length:          2944528
Score:                    86.50           Matches:         36
Percent Similarity:       44.1%           Conservative:    28
Best Local Similarity:    24.8%           Mismatches:      58
Query Match:              7.3%            Indels:          23
DB:                       9               Gaps:            7

US-10-757-745-2_COPY_121_347 (1-227) x US-11-045-004-1 (1-2944528)

Qy         18 SerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyr 37
                |||   |||       |||      |||  :::   ::::::|||   :::   |||
Db     348341 TCATTACTCGATAAATATCAATGGGATTTCATGGGTGTGGAGGAACCTCTTCTGCCACAA 348282
```

```
Qy         38 TyrSerTyrLeuLysLysArgSerSer---AsnTyrGluIleIle---------ThrGly 53
              :::|||:::          :::|||     :::          |||
Db     348281 ATGCGCGACATGAAAGAAATGCAGGATTGGGATTATTTTGGGGTAGGACGAGATGACGGT 348222

Qy         54 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu------ 71
             |||:::|||   |||||||||:::            :::|||     |||
Db     348221 TTTGAAAAGGGAGAATTTACTGCAGTTTTTTACAATTCTACTCGTTTTACCTTGTTGCAA 348162

Qy         72 --------------------LysSerGlnGluIleIleProPheProSerThrLysMet 84
                                  ::::::  ::::::          ||||||   |||
Db     348161 GAAGGTCATTTTTGGCTTTCCGAAACGCCTGATGTTCCTTCCATTCACTCAACAGCCATG 348102

Qy         85 MetArgAsnLeuLeuCysVal---------HisValAsnValSerGlyAsnGluLeuCys 101
                  :::||||||              :::      |||   :::
Db     348101 TTTCCGCGA---ATTTGTGTTTGGGGGAAATTCGCCGATTCGGACGGAAAACAGTTCTAT 348045

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu------ArgMetAsn 119
              :::     :::||||||:::           |||   :::|||      |||:::
Db     348044 ATTTTTAACACTCATTTGGAT-----------CACATCTCTGAGGAAGCTCGTCTATTT 347997

Qy        120 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 139
                 :::::::::|||||||||               |||::::::   ||||||
Db     347996 GCAAGCCAACTTTTACTAAAAAAAGCTGCCACCATTGCAGAAAATTCACCCGTGATTATA 347937

Qy        140 AlaGlyAspThrAsn 144
                |||||   |||
Db     347936 TTAGGCGATTTTAAC 347922
```

RESULT 62
US-60-762-056-9276
; Sequence 9276, Application US/60762056
; GENERAL INFORMATION:
; APPLICANT: Trinklein, Nathan D
; APPLICANT: Aldred, Shelley Force
; APPLICANT: Cooper, Sara J
; APPLICANT: Myers, Richard M
; TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
; TITLE OF INVENTION: EXPRESSION REGULATORY ELEMENTS
; FILE REFERENCE: 33102-701.102
; CURRENT APPLICATION NUMBER: US/60/762,056
; CURRENT FILING DATE: 2006-01-24
; NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 9276
; LENGTH: 2300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-762-056-9276

Alignment Scores:

Pred. No.:	2.02	Length:	2300
Score:	86.00	Matches:	51
Percent Similarity:	45.5%	Conservative:	36
Best Local Similarity:	26.7%	Mismatches:	71
Query Match:	7.2%	Indels:	34
DB:	11	Gaps:	9

US-10-757-745-2_COPY_121_347 (1-227) x US-60-762-056-9276 (1-2300)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||        ||||||   |||   :::   :::::::::             |||
Db       1141 CTAAATGCTCCAATTGAAAGACACAGA---CTGGCAAATTGGATAAAGACTCAAGATCCA 1197

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::||||||                :::||||||     |||
Db       1198 TCAGTGTGCTGTATTCAGGAA-----------------ACCTATCTCACATGCAGAGAC 1239

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::      |||  |||     :::   |||                     |||
Db       1240 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 1299

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||::::::          :::          |||      ||||||:::
Db       1300 AAGGCAGGGGTTGAAATCCCAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 1359

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::            ::::::   |||   |||:::            ||||||
Db       1360 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAATAAGAAGAGCTAACT 1419

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::     :::     :::|||          |||   |||          :::
Db       1420 GTCCTAAATATATATTCACCCAATACA---------GGAGCACCCAGATTC------ATA 1464

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db       1465 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACAATAATAATG---GGA 1521

Qy        142 Asp------------------------ThrAsnLeuArgAspArgGluValThrArgCy 153
              |||                       |||:::   |||||||||:::::::|||
Db       1522 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTTAACAAGCAT 1581

Qy        153 sGlyGlyLeuProAsn-----------------AsnIleValAspValTrpGluPheLe 167
              |||      |||                   :::::::::|||::::::      ||
Db       1582 ACCCGGGAA-TTGAACTCATCTCTGCACCAAGGGGACCTAATAGACATCTACAGAACTCT 1640

Qy        167 uGlyLysProLysHisCysGlnTyrThrTrp 177
              |     ||||||     :::|||||||:::
Db       1641 C---CACCCTAAATCAACAGAATATACATTT 1668
```

RESULT 63
US-10-868-184C-10381/c
; Sequence 10381, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10381
; LENGTH: 6044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-868-184C-10381

Alignment Scores:

Pred. No.:	9.5	Length:	6044
Score:	86.00	Matches:	45
Percent Similarity:	42.6%	Conservative:	36
Best Local Similarity:	23.7%	Mismatches:	77
Query Match:	7.2%	Indels:	32
DB:	6	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10381 (1-6044)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::             |||
Db       4003 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 3947

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
               |||      :::||||||                 ::::::|||      |||
Db       3946 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 3905

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       :::   |||                |||
Db       3904 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 3845

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::  |||  :::       :::         |||      ||||||:::
Db       3844 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 3785

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::            ::::::  |||  |||:::            ||||||
Db       3784 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 3725

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||            |||   |||         :::
Db       3724 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 3680

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||       :::|||       :::|||   |||:::|||      |||
Db       3679 AAGCAAGTCCTCAGTGACCTACAAAGAGTCTTAGACTCCCACACATTAATAATG---GGA 3623

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db       3622 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 3563

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
               :::                    :::::::::|||::::::      |||
Db       3562 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 3503

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
              ||||||      :::||||||:::
Db       3502 ---CACCCCAAATCAACAGAATATACATTT 3476

RESULT 64
US-10-868-184D-10381/c
; Sequence 10381, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10381
;   LENGTH: 6044
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10381

Alignment Scores:
Pred. No.:                9.5          Length:          6044
Score:                    86.00        Matches:         45
Percent Similarity:       42.6%        Conservative:    36
Best Local Similarity:    23.7%        Mismatches:      77
Query Match:              7.2%         Indels:          32
DB:                       6            Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10381 (1-6044)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::             |||
Db       4003 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 3947

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
               |||      :::||||||                 ::::::|||      |||
Db       3946 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 3905

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       :::   |||                |||
Db       3904 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 3845

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::  |||  :::       :::         |||      ||||||:::
Db       3844 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 3785

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::            ::::::  |||  |||:::            ||||||
Db       3784 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 3725

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||            |||   |||         :::
Db       3724 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 3680

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||       :::|||       :::|||   |||:::|||      |||
Db       3679 AAGCAAGTCCTCAGTGACCTACAAAGAGTCTTAGACTCCCACACATTAATAATG---GGA 3623

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db       3622 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 3563

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
               :::                    :::::::::|||::::::      |||
Db       3562 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 3503

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
              ||||||      :::||||||:::
Db       3502 ---CACCCCAAATCAACAGAATATACATTT 3476

RESULT 65
US-10-868-184C-10280/c
; Sequence 10280, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
```

```
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10280
;   LENGTH: 6065
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10280

Alignment Scores:
Pred. No.:                9.55          Length:          6065
Score:                    86.00         Matches:         45
Percent Similarity:       42.6%         Conservative:    36
Best Local Similarity:    23.7%         Mismatches:      77
Query Match:              7.2%          Indels:          32
DB:                       6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10280 (1-6065)

Qy           6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
Db        4012 CTAAATTCTCCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 3956

Qy          26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
Db        3955 TCAGTGTGCTGTATTCAGGAA----------------ACCCATCTCACGTGCAGAGAC 3914

Qy          46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db        3913 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 3854

Qy          66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db        3853 AAGGCAGGGGTTACAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 3794

Qy          86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
Db        3793 AGAGACAAAGAAAGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 3734

Qy         102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db        3733 ATCCTAAATATATATGCACCCAATACA---------GGAGCATCCAGATTC------ATA 3689

Qy         122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db        3688 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 3632

Qy         142 AspThrAsnLeu-----------------------------------------ArgAsp 147
Db        3631 GACTGTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 3572

Qy         148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
Db        3571 ACCCAGGAATTGAACTCATCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 3512

Qy         168 GlyLysProLysHisCysGlnTyrThrTrp 177
Db        3511 ---CACCCCAAATCAACAGAATATACGTTT 3485

RESULT 66
US-10-868-184D-10280/c
; Sequence 10280, Application US/10868184D

; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10280
;   LENGTH: 6065
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10280

Alignment Scores:
Pred. No.:                9.55          Length:          6065
Score:                    86.00         Matches:         45
Percent Similarity:       42.6%         Conservative:    36
Best Local Similarity:    23.7%         Mismatches:      77
Query Match:              7.2%          Indels:          32
DB:                       6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10280 (1-6065)

Qy           6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
Db        4012 CTAAATTCTCCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 3956

Qy          26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
Db        3955 TCAGTGTGCTGTATTCAGGAA----------------ACCCATCTCACGTGCAGAGAC 3914

Qy          46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db        3913 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 3854

Qy          66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db        3853 AAGGCAGGGGTTACAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 3794

Qy          86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
Db        3793 AGAGACAAAGAAAGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 3734

Qy         102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db        3733 ATCCTAAATATATATGCACCCAATACA---------GGAGCATCCAGATTC------ATA 3689

Qy         122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db        3688 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 3632
```

```
Qy        142 AspThrAsnLeu-------------------------------------------ArgAsp 147
              |||   |||                                              :::|||
Db       3631 GACTGTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 3572

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::                  :::::::::|||::::::      |||
Db       3571 ACCCAGGAATTGAACTCATCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 3512

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                  ||||||      :::|||||||:::
Db       3511 ---CACCCCAAATCAACAGAATATACGTTT 3485

RESULT 67
US-10-868-184C-10386/c
; Sequence 10386, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10386
;   LENGTH: 6495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10386

Alignment Scores:
Pred. No.:              10.7          Length:         6495
Score:                  86.00         Matches:        45
Percent Similarity:     43.2%         Conservative:   37
Best Local Similarity:  23.7%         Mismatches:     76
Query Match:            7.2%          Indels:         32
DB:                     6             Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10386 (1-6495)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::           |||
Db       4438 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 4382

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::||||||                   :::::::|||      |||
Db       4381 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGTAGAGAC 4340

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::      |||   |||      :::   |||                  |||
Db       4339 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 4280

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
```

```
              |||:::   |||   :::         :::         |||       ||||||:::
Db       4279 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 4220

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::          ::::::   |||   |||:::            ||||||
Db       4219 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 4160

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::     :::|||          |||   |||         :::
Db       4159 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 4115

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||          :::|||   |||:::|||      |||
Db       4114 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 4058

Qy        142 AspThrAsnLeu-------------------------------------------ArgAsp 147
              |||   |||                                              :::|||
Db       4057 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 3998

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::      :::          :::::::::|||::::::      |||
Db       3997 ACCCAGGAATTGAAATCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 3938

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                  ||||||      :::|||||||:::
Db       3937 ---CACCCCAAATCAACAGAATATACATTT 3911

RESULT 68
US-10-868-184D-10386/c
; Sequence 10386, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10386
;   LENGTH: 6495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10386

Alignment Scores:
Pred. No.:              10.7          Length:         6495
Score:                  86.00         Matches:        45
Percent Similarity:     43.2%         Conservative:   37
Best Local Similarity:  23.7%         Mismatches:     76
Query Match:            7.2%          Indels:         32
DB:                     6             Gaps:           8
```

```
US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10386 (1-6495)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::              |||
Db       4438 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 4382

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
              |||      :::||||||                 ::::::|||      |||
Db       4381 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGTAGAGAC 4340

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       :::   |||                  |||
Db       4339 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 4280

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db       4279 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 4220

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::            ::::::   |||   |||:::            ||||||
Db       4219 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 4160

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||   |||         :::
Db       4159 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 4115

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db       4114 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 4058

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db       4057 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 3998

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::      :::                 :::::::::|||::::::      |||
Db       3997 ACCCAGGAATTGAAATCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 3938

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                    ||||||      :::|||||||:::
Db       3937 ---CACCCCAAATCAACAGAATATACATTT 3911

RESULT 69
US-10-868-184C-10688/c
; Sequence 10688, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046

;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10688
;   LENGTH: 14411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184C-10688

Alignment Scores:
Pred. No.:                38.2          Length:          14411
Score:                    86.00         Matches:         45
Percent Similarity:       42.6%         Conservative:    36
Best Local Similarity:    23.7%         Mismatches:      77
Query Match:              7.2%          Indels:          32
DB:                       6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10688 (1-14411)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::              |||
Db       4016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 3960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
              |||      :::||||||                 ::::::|||      |||
Db       3959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCATGTGCAGAGAC 3918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       :::   |||                  |||
Db       3917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 3858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db       3857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 3798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::            ::::::   |||   |||:::            ||||||
Db       3797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 3738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||   |||         :::
Db       3737 ATCCTAAATATGTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 3693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db       3692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 3636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db       3635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 3576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::                         :::::::::|||::::::      |||
Db       3575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 3516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                    ||||||      :::|||||||:::
Db       3515 ---CACCCCAAATCAGCAGAATATACATTT 3489

RESULT 70
US-10-868-184D-10688/c
; Sequence 10688, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
```

```
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10688
;   LENGTH: 14411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10688

Alignment Scores:
Pred. No.:                38.2          Length:          14411
Score:                    86.00         Matches:         45
Percent Similarity:       42.6%         Conservative:    36
Best Local Similarity:    23.7%         Mismatches:      77
Query Match:              7.2%          Indels:          32
DB:                       6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10688 (1-14411)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::     :::|||  |||   :::  |||::::::           |||
Db       4016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 3960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::||||||                 ::::::|||      |||
Db       3959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCATGTGCAGAGAC 3918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::     |||  |||       :::   |||           |||
Db       3917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 3858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||     ||||||:::
Db       3857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 3798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::          ::::::   |||   |||:::             ||||||
Db       3797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 3738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::   :::    :::|||          |||   |||          :::
Db       3737 ATCCTAAATATGTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 3693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||        :::|||  |||:::|||      |||
Db       3692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 3636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db       3635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 3576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
              :::                        ::::::::::|||::::::      |||
Db       3575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 3516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                |||||      :::|||||||:::
Db       3515 ---CACCCCAAATCAGCAGAATATACATTT 3489

RESULT 71
```

```
US-11-266-748A-60910/c
; Sequence 60910, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 60910
;   LENGTH: 256164
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-60910

Alignment Scores:
Pred. No.:                3.83e+03      Length:          256164
Score:                    86.00         Matches:         45
Percent Similarity:       41.5%         Conservative:    36
Best Local Similarity:    23.1%         Mismatches:      80
Query Match:              7.2%          Indels:          34
DB:                       9             Gaps:            7

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-60910 (1-256164)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ::::::||||||          |||:::      :::         :::   :::::::::
Db     218493 GTAAATGGGCTA---------AATGTGCCAATTAAAAGACACAGACTAGTAAATTGGATA 218443

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                  |||   |||       :::|||||||                      ::::::
Db     218442 AAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAG-----------------ACCCAT 218401

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||     |||   |||:::          |||   |||          :::   |||
Db     218400 CTCACATGCAGAGACTCACATAGGCTCAACATAAAGGGATAGAGGAAGATCTACCAAGCA 218341

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                    ||||||:::   |||   :::         :::          |||
Db     218340 AATGGAGAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAA 218281

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly--------- 97
                  ||||||:::   |||:::           ::::::   |||   |||
Db     218280 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATGAATTCAA 218221

Qy         98 ---AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                 ||||||   ::::::      :::         :::|||                |||
Db     218220 CAAGAAGAGCTAACTATCCTAAATATATATGCACCCAATACA--------------GAA 218176

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
                    :::|||   ||||||:::   :::|||          :::|||   |||
Db     218175 GCACCCAGATTCATAAAACAAGTCCTTAGAGACCTACAAAGAAATTTAGACTCCCACACA 218116
```

```
Qy        137 ValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeu 156
              :::|||      ||||||   |||           :::   |||      |||
Db     218115 ATAATAATG---GGAGACTTTAACACCCCACTATCAACATTAGACAGATCAACGGGACAG 218059

Qy        157 ProAsnAsn-----------------------------------------IleValAsp 162
                 |||                                            ::::::|||
Db     218058 AAAGTTAACAAGGATATCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGAC 217999

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::::::      |||      ||||||      :::||||||::
Db     217998 ATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACTTTC 217957

RESULT 72
US-11-266-748A-184392
; Sequence 184392, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 184392
;   LENGTH: 1349
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-184392

Alignment Scores:
Pred. No.:              1.01          Length:          1349
Score:                  85.50         Matches:         49
Percent Similarity:     39.8%         Conservative:    29
Best Local Similarity:  25.0%         Mismatches:      70
Query Match:            7.2%          Indels:          49
DB:                     9             Gaps:            6

US-10-757-745-2_COPY_121_347 (1-227) x US-11-266-748A-184392 (1-1349)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||      |||   :::|||              :::||||||
Db        718 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 777

Qy         33 ValIleProProTyr-------------------------------------TyrSer 39
              :::      |||                                        |||
Db        778 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 837

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::     |||             :::||||||      ||||||       ||
Db        838 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 882

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
              ||||      :::   :::||||||||||         :::|||   |||
Db        883 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 942

Qy         76 ----------------------------------------------IleIleProPhePr 80
                                                            |||:::      ||
Db        943 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 1002

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::     ||||||          :::::::|||:::   :::   ::
Db       1003 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1056

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::||||||                |||       ||
Db       1057 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1115

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |       :::|||   :::|||   :::|||                :::::   ||
Db       1116 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1169

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||              ||||||:::||||||      |||             ||
Db       1170 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1211

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              |   ::::::   :::   |||           |||   :::|||
Db       1212 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATAT 1257

RESULT 73
US-10-461-673-2142/c
; Sequence 2142, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
```

```
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 2142
;    LENGTH: 1575
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (1)..(1152)
US-10-461-673-2142

Alignment Scores:
Pred. No.:                1.29          Length:          1575
Score:                    85.50         Matches:         45
Percent Similarity:       39.3%         Conservative:    39
Best Local Similarity:    21.0%         Mismatches:      84
Query Match:              7.2%          Indels:          46
DB:                       6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-461-673-2142 (1-1575)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
              :::  :::|||:::::::   ||||||     :::     :::|||   :::
Db       1053 GAAGCGGTTGACATCAGTAAACTGTCTCCTTCTTCTAGCTTCATCTGCTAATTCCAGTGC 994

Qy         22 LeuTyrSerProAspVal--------------------------------IlePheLeu 30
              ||||||     ::::::                               |||:::::
Db        993 TTGTACAAGAACAATATCATCATTAGAGGAGAAAATGGTCAGAGGAGGTGTATCTGGATC 934

Qy         31 GlnGluValIle-----------------------------------------ProPro 36
              :::||||||                                           |||
Db        933 AGGGAAGTTACGCTGAAGTGGATCATAGTGGATGCCATCATAAATAAGCAGAACCCTTTT 874

Qy         37 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 56
                |||   :::     ::::::  |||         :::      |||
Db        873 GGTATATCCTGCATCTTCCCCAAAACGATCAATTCTTACTGTGTGTGTATCCACTACACA 814

Qy         57 GlyTyrPheThrAlaIleMetLeuLysLysSerArg---ValLysLeuLysSerGlnGlu 75
              |||||||||  :::  |||  ::::::|||  :::      |||
Db        813 ---TATTTCACATTGGTAAAACTTGGACAAAATCGATATCTCTATTGCTCCTCCCCAAGT 757

Qy         76 IleIleProPhe---ProSerThrLysMetMet-----------ArgAsnLeuLeuCys 90
              :::||||||||   |||  |||  ::::::          :::      |||
Db        756 GTCATCCCTTTTGATCCAGTCACAGTACTCTTGATTTGTTTTTCCCAGTATTGCCTCACT 697

Qy         91 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 110
              :::  |||  ::::::       ||||||       |||||||||      |||
Db        696 ATAGAAGTCTGGATCGCTTGCTACAATTTGTGCTATGAGGCGTCTCATCTCAGGGGCACA 637

Qy        111 -ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGl 130
               |||                          ::::::|||:::   :::
Db        636 AGCTGGATTCAAGACTCCTCCTTCGACGACATAGTACACACTAGTAAAGAGGCAAGAGTT 577

Qy        130 uAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluVa 150
              :::      :::      ::::::      |||      :::|||      |||
Db        576 GTCTGCTGGGACCACGGTTCTGGTAAGCACAGGCAAAGTTTCCCTGACGTAACTAGAAGC 517

Qy        150 lThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPr 170
              |||                        :::|||:::      |||   ||||||
Db        516 ACCACGT--------------------TTAGTAAATGCAGGTGAACTTCTGGGC----- 483

Qy        170 oLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
                        |||   :::   |||   |||
Db        482 ------------CTGGTTTGGTCTTCTTCAATGATCAGC 456

RESULT 74
US-10-868-184C-10629/c
; Sequence 10629, Application US/10868184C
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184C
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10629
;    LENGTH: 6060
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-868-184C-10629

Alignment Scores:
Pred. No.:                11.2          Length:          6060
Score:                    85.50         Matches:         43
Percent Similarity:       42.3%         Conservative:    39
Best Local Similarity:    22.2%         Mismatches:      79
Query Match:              7.2%          Indels:          33
DB:                       6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184C-10629 (1-6060)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::      :::   :::   :::::::::
Db       4038 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 3979

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||      :::|||||||                    :::::|||
Db       3978 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 3937

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::          |||  |||      :::   |||
Db       3936 ATGAGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 3877

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                   ||||||:::   |||   :::      :::      |||
Db       3876 GGAAAACAAAAAAAGGCAGAGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3817

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::            ::::::   |||   |||:::
Db       3816 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAACGGTAAAGGGATCAATTCAACAA 3757

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||     ::::::      :::      :::|||          |||   |||
Db       3756 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3706

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                    :::|||   ||||||       :::|||          :::|||   |||:::
Db       3705 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 3652

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
                 |||      ||||||   |||
```

```
Db       3651 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 3595

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                  :::|||   :::                          :::::::::|||:::
Db       3594 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 3535

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::||||||:::
Db       3534 TACAGAACTCTC---CACCCTAAATCAACAGAATATACATTT 3496

RESULT 75
US-10-868-184D-10629/c
; Sequence 10629, Application US/10868184D
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al
;  TITLE OF INVENTION: Human Secreted Proteins
;  FILE REFERENCE: PS805
;  CURRENT APPLICATION NUMBER: US/10/868,184D
;  CURRENT FILING DATE:  2004-06-16
;  PRIOR APPLICATION NUMBER: 60/278,650
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 09/833,245
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US01/11988
;  PRIOR FILING DATE: 2001-04-12
;  PRIOR APPLICATION NUMBER: PCT/US00/06043
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06012
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06058
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06044
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06059
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06042
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: PCT/US00/06014
;  PRIOR FILING DATE: 2000-03-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 13046
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10629
;   LENGTH: 6060
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-868-184D-10629

Alignment Scores:
Pred. No.:              11.2          Length:         6060
Score:                  85.50         Matches:        43
Percent Similarity:     42.3%         Conservative:   39
Best Local Similarity:  22.2%         Mismatches:     79
Query Match:            7.2%          Indels:         33
DB:                     6             Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-868-184D-10629 (1-6060)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::         :::         :::   :::::::::
Db       4038 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 3979

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||      :::||||||                 :::::::|||
Db       3978 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 3937

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::         |||   |||       :::   |||
Db       3936 ATGAGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 3877

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                |||||:::   |||   :::         :::         |||
Db       3876 GGAAAACAAAAAAAGGCAGAGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3817

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::             ::::::   |||   |||:::
Db       3816 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAACGGTAAAGGGATCAATTCAACAA 3757

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||                |||   |||
Db       3756 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3706

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   ||||||      :::|||          :::|||   |||:::
Db       3705 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 3652

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db       3651 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 3595

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                  :::|||   :::                          :::::::::|||:::
Db       3594 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 3535

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::||||||:::
Db       3534 TACAGAACTCTC---CACCCTAAATCAACAGAATATACATTT 3496

Search completed: March 11, 2006, 00:04:49
Job time : 831.106 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:19:50 ; Search time 369.917 Seconds
(without alignments)
1345.232 Million cell updates/sec

Title: US-10-757-745-2_COPY_145_153
Perfect score: 45
Sequence: 1 PDVIFLQEV 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170333_4165/app_query.fasta_1
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=rnpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss02h -USER=US10757745_@CGN_1_1_20983_@runat_09032006_170333_4165
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUSB_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUSC_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US075_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US076_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US077_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US078_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US079_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US101_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US104_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US105_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US112_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US117_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US600_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US601_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US603_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US605_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US606_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US607_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	45	100.0	176	1	PCT-US02-02781-1367	Sequence 1367, Ap
c 2	45	100.0	176	34	US-09-878-178-50	Sequence 50, Appl
3	45	100.0	176	34	US-09-878-722-121	Sequence 121, App
4	45	100.0	176	35	US-09-904-456-121	Sequence 121, App
c 5	45	100.0	176	40	US-10-046-935-50	Sequence 50, Appl
c 6	45	100.0	176	40	US-10-060-036-1367	Sequence 1367, Ap
c 7	45	100.0	176	42	US-10-146-502-50	Sequence 50, Appl
c 8	45	100.0	201	79	US-60-443-566-20813	Sequence 20813, A

	9	45	100.0	201	79	US-60-443-566-20854 Sequence 20854, A
c	10	45	100.0	201	79	US-60-452-680-100212 Sequence 100212,
	11	45	100.0	201	79	US-60-452-680-100240 Sequence 100240,
c	12	45	100.0	201	79	US-60-455-444-41303 Sequence 41303, A
	13	45	100.0	201	79	US-60-455-444-41320 Sequence 41320, A
c	14	45	100.0	201	79	US-60-465-241-41303 Sequence 41303, A
	15	45	100.0	201	79	US-60-465-241-41320 Sequence 41320, A
c	16	45	100.0	201	79	US-60-465-241-204808 Sequence 204808,
	17	45	100.0	201	79	US-60-465-241-204907 Sequence 204907,
	18	45	100.0	221	27	US-09-540-766-44489 Sequence 44489, A
	19	45	100.0	221	74	US-60-017-556-150 Sequence 150, App
	20	45	100.0	229	27	US-09-540-766-28296 Sequence 28296, A
	21	45	100.0	229	74	US-60-017-556-751 Sequence 751, App
	22	45	100.0	237	20	US-09-041-720-965 Sequence 965, App
	23	45	100.0	237	26	US-09-539-334-23928 Sequence 23928, A
	24	45	100.0	237	74	US-60-040-197-965 Sequence 965, App
	25	45	100.0	248	20	US-09-015-080-208 Sequence 208, App
	26	45	100.0	248	26	US-09-540-212A-31516 Sequence 31516, A
	27	45	100.0	250	19	US-08-961-526-1051 Sequence 1051, Ap
	28	45	100.0	250	26	US-09-539-334-31143 Sequence 31143, A
	29	45	100.0	252	12	US-08-276-163A-3853 Sequence 3853, Ap
	30	45	100.0	252	12	US-08-276-163B-3853 Sequence 3853, Ap
	31	45	100.0	252	12	US-08-276-163D-3853 Sequence 3853, Ap
	32	45	100.0	252	33	US-09-840-145-3853 Sequence 3853, Ap
	33	45	100.0	252	35	US-09-912-293-29561 Sequence 29561, A
	34	45	100.0	266	18	US-08-834-645-54 Sequence 54, Appl
	35	45	100.0	266	27	US-09-540-766-4618 Sequence 4618, Ap
	36	45	100.0	292	19	US-08-917-046-231 Sequence 231, App
	37	45	100.0	292	27	US-09-540-766-3985 Sequence 3985, Ap
	38	45	100.0	303	20	US-09-041-720-1771 Sequence 1771, Ap
	39	45	100.0	303	26	US-09-539-334-1416 Sequence 1416, Ap
	40	45	100.0	303	74	US-60-040-197-1771 Sequence 1771, Ap
	41	45	100.0	396	25	US-09-528-409-24161 Sequence 24161, A
	42	45	100.0	396	36	US-09-933-524-24161 Sequence 24161, A
	43	45	100.0	396	36	US-09-933-524A-24161 Sequence 24161, A
	44	45	100.0	397	25	US-09-528-409-26306 Sequence 26306, A
	45	45	100.0	397	36	US-09-933-524-26306 Sequence 26306, A
	46	45	100.0	397	36	US-09-933-524A-26306 Sequence 26306, A
	47	45	100.0	401	1	PCT-US02-02870-1363 Sequence 1363, Ap
	48	45	100.0	401	25	US-09-528-409-74640 Sequence 74640, A
	49	45	100.0	401	36	US-09-933-524-74640 Sequence 74640, A
	50	45	100.0	401	36	US-09-933-524A-74640 Sequence 74640, A
	51	45	100.0	401	40	US-10-066-543-1363 Sequence 1363, Ap
	52	45	100.0	401	43	US-10-214-403-1363 Sequence 1363, Ap
	53	45	100.0	414	24	US-09-404-549-2518 Sequence 2518, Ap
	54	45	100.0	414	24	US-09-404-549A-2518 Sequence 2518, Ap
	55	45	100.0	417	33	US-09-824-559-4370 Sequence 4370, Ap
	56	45	100.0	418	23	US-09-359-067-7230 Sequence 7230, Ap
	57	45	100.0	420	25	US-09-528-409-24160 Sequence 24160, A
	58	45	100.0	420	36	US-09-933-524-24160 Sequence 24160, A
	59	45	100.0	420	36	US-09-933-524A-24160 Sequence 24160, A
	60	45	100.0	433	25	US-09-528-409-74096 Sequence 74096, A
	61	45	100.0	433	36	US-09-933-524-74096 Sequence 74096, A
	62	45	100.0	433	36	US-09-933-524A-74096 Sequence 74096, A
	63	45	100.0	444	33	US-09-864-761-16762 Sequence 16762, A
	64	45	100.0	444	42	US-10-182-997-5244 Sequence 5244, Ap
	65	45	100.0	444	77	US-60-236-359-6954 Sequence 6954, Ap
	66	45	100.0	449	23	US-09-304-649-421 Sequence 421, App
	67	45	100.0	449	33	US-09-840-424-421 Sequence 421, App
	68	45	100.0	483	19	US-08-905-144-190 Sequence 190, App
	69	45	100.0	483	27	US-09-547-599C-825 Sequence 825, App
	70	45	100.0	483	54	US-10-631-402-825 Sequence 825, App
	71	45	100.0	483	54	US-10-631-441-825 Sequence 825, App
	72	45	100.0	483	54	US-10-631-441A-825 Sequence 825, App
	73	45	100.0	485	28	US-09-606-776-1036 Sequence 1036, Ap
	74	45	100.0	485	29	US-09-644-868-6183 Sequence 6183, Ap
	75	45	100.0	502	76	US-60-217-080-8309 Sequence 8309, Ap

ALIGNMENTS

RESULT 1
PCT-US02-02781-1367/c

```
; Sequence 1367, Application PC/TUS0202781
; GENERAL INFORMATION:
;  APPLICANT: Corixa Corporation
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Persing, David H.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF PANCREATIC CANCER
;  FILE REFERENCE: 210121.566PC
;  CURRENT APPLICATION NUMBER: PCT/US02/02781
;  CURRENT FILING DATE:  2002-01-30
;  NUMBER OF SEQ ID NOS: 4560
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1367
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US02-02781-1367

Alignment Scores:
Pred. No.:                12.5          Length:          176
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       1             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x PCT-US02-02781-1367 (1-176)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145

RESULT 2
US-09-878-178-50/c
; Sequence 50, Application US/09878178
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527
;  CURRENT APPLICATION NUMBER: US/09/878,178
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 2237
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-878-178-50

Alignment Scores:
Pred. No.:                12.5          Length:          176
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       34            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-878-178-50 (1-176)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145

RESULT 3
US-09-878-722-121
; Sequence 121, Application US/09878722
```

; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Hepler, William T.
; APPLICANT: Clapper, Jonathan
; APPLICANT: Wang, Aijun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.524
; CURRENT APPLICATION NUMBER: US/09/878,722
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 245
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-722-121

Alignment Scores:

Pred. No.:	12.5	Length:	176
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	34	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-878-722-121 (1-176)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          6 CCAGATGTGATATTTCTACAGGAAGTT 32
```

RESULT 4
US-09-904-456-121
; Sequence 121, Application US/09904456
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.524C1
; CURRENT APPLICATION NUMBER: US/09/904,456
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 247
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-456-121

Alignment Scores:

Pred. No.:	12.5	Length:	176
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	35	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-904-456-121 (1-176)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          6 CCAGATGTGATATTTCTACAGGAAGTT 32
```

RESULT 5
US-10-046-935-50/c
; Sequence 50, Application US/10046935
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-50

Alignment Scores:

Pred. No.:	12.5	Length:	176
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	40	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-046-935-50 (1-176)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145
```

RESULT 6
US-10-060-036-1367/c
; Sequence 1367, Application US/10060036
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1367
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-1367

Alignment Scores:

Pred. No.:	12.5	Length:	176
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	40	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-060-036-1367 (1-176)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145
```

RESULT 7
US-10-146-502-50/c
; Sequence 50, Application US/10146502
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun

```
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C2
;  CURRENT APPLICATION NUMBER: US/10/146,502
;  CURRENT FILING DATE:  2002-05-14
;  NUMBER OF SEQ ID NOS: 2241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-146-502-50

Alignment Scores:
Pred. No.:                12.5          Length:         176
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       42            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-146-502-50 (1-176)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145

RESULT 8
US-60-443-566-20813/c
; Sequence 20813, Application US/60443566
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001447
;  CURRENT APPLICATION NUMBER: US/60/443,566
;  CURRENT FILING DATE:  2003-01-30
;  NUMBER OF SEQ ID NOS: 25102
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20813
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-443-566-20813

Alignment Scores:
Pred. No.:                14.6          Length:         201
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       79            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-443-566-20813 (1-201)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        164 CCAGATGTGATATTTCTACAGGAAGTT 138

RESULT 9
US-60-443-566-20854
; Sequence 20854, Application US/60443566
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001447
;  CURRENT APPLICATION NUMBER: US/60/443,566
;  CURRENT FILING DATE:  2003-01-30
;  NUMBER OF SEQ ID NOS: 25102
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20854
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-443-566-20854

Alignment Scores:
Pred. No.:                14.6          Length:         201
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       79            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-443-566-20854 (1-201)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         38 CCAGATGTGATATTTCTACAGGAAGTT 64

RESULT 10
US-60-452-680-100212/c
; Sequence 100212, Application US/60452680
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  GRUPE, Andrew
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001450
;  CURRENT APPLICATION NUMBER: US/60/452,680
;  CURRENT FILING DATE:  2003-03-07
;  NUMBER OF SEQ ID NOS: 116213
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100212
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-452-680-100212

Alignment Scores:
Pred. No.:                14.6          Length:         201
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       79            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-452-680-100212 (1-201)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        164 CCAGATGTGATATTTCTACAGGAAGTT 138

RESULT 11
US-60-452-680-100240
; Sequence 100240, Application US/60452680
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  GRUPE, Andrew
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001450
;  CURRENT APPLICATION NUMBER: US/60/452,680
;  CURRENT FILING DATE:  2003-03-07
;  NUMBER OF SEQ ID NOS: 116213
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100240
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-452-680-100240
```

Alignment Scores:

Pred. No.:	14.6	Length:	201
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	79	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-452-680-100240 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         38 CCAGATGTGATATTTCTACAGGAAGTT 64
```

RESULT 12
US-60-455-444-41303/c
; Sequence 41303, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41303
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-455-444-41303

Alignment Scores:

Pred. No.:	14.6	Length:	201
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	79	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-455-444-41303 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        164 CCAGATGTGATATTTCTACAGGAAGTT 138
```

RESULT 13
US-60-455-444-41320
; Sequence 41320, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41320
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-455-444-41320

Alignment Scores:

Pred. No.:	14.6	Length:	201
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	79	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-455-444-41320 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         38 CCAGATGTGATATTTCTACAGGAAGTT 64
```

RESULT 14
US-60-465-241-41303/c
; Sequence 41303, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41303
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-465-241-41303

Alignment Scores:

Pred. No.:	14.6	Length:	201
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	79	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-465-241-41303 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        164 CCAGATGTGATATTTCTACAGGAAGTT 138
```

RESULT 15
US-60-465-241-41320
; Sequence 41320, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41320
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-465-241-41320

Alignment Scores:

Pred. No.:	14.6	Length:	201
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	79	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-465-241-41320 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         38 CCAGATGTGATATTTCTACAGGAAGTT 64
```

```
RESULT 16
US-60-465-241-204808/c
; Sequence 204808, Application US/60465241
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001468
;  CURRENT APPLICATION NUMBER: US/60/465,241
;  CURRENT FILING DATE:  2003-04-23
;  NUMBER OF SEQ ID NOS: 258418
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204808
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-465-241-204808

Alignment Scores:
Pred. No.:                14.6          Length:         201
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       79            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-465-241-204808 (1-201)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        164 CCAGATGTGATATTTCTACAGGAAGTT 138

RESULT 17
US-60-465-241-204907
; Sequence 204907, Application US/60465241
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001468
;  CURRENT APPLICATION NUMBER: US/60/465,241
;  CURRENT FILING DATE:  2003-04-23
;  NUMBER OF SEQ ID NOS: 258418
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204907
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-465-241-204907

Alignment Scores:
Pred. No.:                14.6          Length:         201
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       79            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-465-241-204907 (1-201)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         38 CCAGATGTGATATTTCTACAGGAAGTT 64

RESULT 18
US-09-540-766-44489
; Sequence 44489, Application US/09540766
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE
;  FILE REFERENCE:  PD-1024 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,766
;  CURRENT FILING DATE:  2000-03-30
;  PRIOR APPLICATION DATA removed - refer tp PALM or File Wrapper
;  NUMBER OF SEQ ID NOS:  77960
;  SOFTWARE:  PERL Program
; SEQ ID NO 44489
;   LENGTH: 221
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00178725
US-09-540-766-44489

Alignment Scores:
Pred. No.:                16.3          Length:         221
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       27            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-540-766-44489 (1-221)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         56 CCAGATGTGATATTTCTACAGGAAGTT 82

RESULT 19
US-60-017-556-150
; Sequence 150, Application US/60017556
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuve, Laura L.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Naughton, Rebecca E.
;    APPLICANT:  Klingler, Tod M.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  HUMAN ASCENDING COLON
;    NUMBER OF SEQUENCES:  1719
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/60/017,556
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LUTHER, BARBARA J.
;      REGISTRATION NUMBER:  33954
;      REFERENCE/DOCKET NUMBER:  PD-0173P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 852-0195
;  INFORMATION FOR SEQ ID NO:  150:
;    SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  221 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     IMMEDIATE SOURCE:
;       CLONE:  1333829
US-60-017-556-150

Alignment Scores:
Pred. No.:                16.3          Length:         221
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       74            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-017-556-150 (1-221)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         56 CCAGATGTGATATTTCTACAGGAAGTT 82

RESULT 20
US-09-540-766-28296
; Sequence 28296, Application US/09540766
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE
;  FILE REFERENCE:  PD-1024 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,766
;  CURRENT FILING DATE:  2000-03-30
;  PRIOR APPLICATION DATA removed - refer tp PALM or File Wrapper
;  NUMBER OF SEQ ID NOS:  77960
;  SOFTWARE:  PERL Program
; SEQ ID NO 28296
;   LENGTH: 229
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00536627
;   NAME/KEY: unsure
;   LOCATION: 104, 174, 178, 190, 193, 212, 223-224
;   OTHER INFORMATION: a, t, c, g, or other
US-09-540-766-28296

Alignment Scores:
Pred. No.:                17            Length:         229
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       27            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-540-766-28296 (1-229)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         37 CCAGATGTGATATTTCTACAGGAAGTT 63

RESULT 21
US-60-017-556-751
; Sequence 751, Application US/60017556
;  GENERAL INFORMATION:
;     APPLICANT:  Gooding, Douglas H.
;     APPLICANT:  Stuve, Laura L.
;     APPLICANT:  Stuart, Susan G.
;     APPLICANT:  Ito, Laura Y.
;     APPLICANT:  Akerblom, Ingrid E.
;     APPLICANT:  Delegeane, Angelo M.
;     APPLICANT:  Naughton, Rebecca E.
;     APPLICANT:  Klingler, Tod M.
;     TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;     TITLE OF INVENTION:  HUMAN ASCENDING COLON
;     NUMBER OF SEQUENCES:  1719
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;       STREET:  3174 PORTER DRIVE
;       CITY:  PALO ALTO
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/60/017,556
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  LUTHER, BARBARA J.
;       REGISTRATION NUMBER:  33954
;       REFERENCE/DOCKET NUMBER:  PD-0173P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 855-0555
;       TELEFAX:  (415) 852-0195
;   INFORMATION FOR SEQ ID NO:  751:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  229 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     IMMEDIATE SOURCE:
;       CLONE:  1335662
US-60-017-556-751

Alignment Scores:
Pred. No.:                17            Length:         229
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       74            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-017-556-751 (1-229)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         37 CCAGATGTGATATTTCTACAGGAAGTT 63

RESULT 22
US-09-041-720-965
; Sequence 965, Application US/09041720
;  GENERAL INFORMATION:
;     APPLICANT:  Gooding, Douglas H.
;     APPLICANT:  Stuve, Laura L.
;     APPLICANT:  Stuart, Susan G.
;     APPLICANT:  Ito, Laura Y.
;     APPLICANT:  Akerblom, Ingrid E.
;     APPLICANT:  Delegeane, Angelo M.
;     APPLICANT:  Naughton, Rebecca E.
;     APPLICANT:  Klingler, Tod M.
;     TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;     TITLE OF INVENTION:  ADRENAL GLAND TUMOR
;     NUMBER OF SEQUENCES:  4388
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
```

```
;       STREET:  3174 PORTER DRIVE
;       CITY:  PALO ALTO
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/041,720
;       FILING DATE:  HEREWITH
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/040,197
;       FILING DATE:  MARCH 14, 1997
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CERRONE, MICHAEL C.
;       REGISTRATION NUMBER:  39,132
;       REFERENCE/DOCKET NUMBER:  PD-0333 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (650) 855-0555
;       TELEFAX:  (650) 845-4166
;   INFORMATION FOR SEQ ID NO:  965:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  237 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     IMMEDIATE SOURCE:
;       CLONE:  2822981H1
US-09-041-720-965

Alignment Scores:
Pred. No.:                17.6          Length:          237
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       20            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-041-720-965 (1-237)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         66 CCAGATGTGATATTTCTACAGGAAGTT 92

RESULT 23
US-09-539-334-23928
; Sequence 23928, Application US/09539334
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF ENDOCRINE SYSTEM    TISSUE
;  FILE REFERENCE:  PD-1026 CIP
;  CURRENT APPLICATION NUMBER: US/09/539,334
;  CURRENT FILING DATE:  2000-03-30
;  "Prior application data removed - refer to PALM or file wrapper"
;  NUMBER OF SEQ ID NOS:  38381
;  SOFTWARE:  PERL Program
; SEQ ID NO 23928
;   LENGTH: 237
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00710528
;   NAME/KEY: unsure
;   LOCATION: 3, 190, 208
;   OTHER INFORMATION: a, t, c, g, or other
US-09-539-334-23928

Alignment Scores:
Pred. No.:                17.6          Length:          237
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       26            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-539-334-23928 (1-237)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         66 CCAGATGTGATATTTCTACAGGAAGTT 92

RESULT 24
US-60-040-197-965
; Sequence 965, Application US/60040197
;  GENERAL INFORMATION:
;     APPLICANT:  Gooding, Douglas H.
;     APPLICANT:  Stuve, Laura L.
;     APPLICANT:  Stuart, Susan G.
;     APPLICANT:  Ito, Laura Y.
;     APPLICANT:  Akerblom, Ingrid E.
;     APPLICANT:  Delegeane, Angelo M.
;     APPLICANT:  Naughton, Rebecca E.
;     APPLICANT:  Klingler, Tod M.
;     TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;     TITLE OF INVENTION:  ADRENAL GLAND TUMOR
;     NUMBER OF SEQUENCES:  2437
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;       STREET:  3174 PORTER DRIVE
;       CITY:  PALO ALTO
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/60/040,197
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CERRONE, MICHAEL C.
;       REGISTRATION NUMBER:  39,132
;       REFERENCE/DOCKET NUMBER:  PD-0333P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 855-0555
;       TELEFAX:  (415) 845-4166
;   INFORMATION FOR SEQ ID NO:  965:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  237 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     IMMEDIATE SOURCE:
;       CLONE:  2822981H1
US-60-040-197-965

Alignment Scores:
Pred. No.:                17.6          Length:          237
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
```

```
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     74              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-040-197-965 (1-237)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         66 CCAGATGTGATATTTCTACAGGAAGTT 92

RESULT 25
US-09-015-080-208
; Sequence 208, Application US/09015080
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuve, Laura L.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Naughton, Rebecca E.
;    APPLICANT:  Klingler, Tod M.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  LUNG WITH SQUAMOUS CELL CARCINOMA
;    NUMBER OF SEQUENCES:  4133
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/015,080
;      FILING DATE:  Herewith
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,039
;      FILING DATE:  January 28, 1997
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PD-0310 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  208:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  248 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  2603934H1
US-09-015-080-208

Alignment Scores:
Pred. No.:              18.6            Length:         248
Score:                  45.00           Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     20              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-015-080-208 (1-248)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        101 CCAGATGTGATATTTCTACAGGAAGTT 127

RESULT 26
US-09-540-212A-31516
; Sequence 31516, Application US/09540212A
; GENERAL INFORMATION:
;  APPLICANT: Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION: POLYNUCLEOTIDES OF AIRWAY AND LUNG SYSTEM TISSUE
;  FILE REFERENCE: PD-1034 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,212A
;  CURRENT FILING DATE:  2000-03-31
;  NUMBER OF SEQ ID NOS: 67551
;  SOFTWARE: PERL Program
; SEQ ID NO 31516
;   LENGTH: 248
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00625939
;   NAME/KEY: unsure
;   LOCATION: 61
;   OTHER INFORMATION: a, t, c, g, or other
US-09-540-212A-31516

Alignment Scores:
Pred. No.:              18.6            Length:         248
Score:                  45.00           Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     26              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-540-212A-31516 (1-248)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        101 CCAGATGTGATATTTCTACAGGAAGTT 127

RESULT 27
US-08-961-526-1051
; Sequence 1051, Application US/08961526
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuve, Laura L.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Naughton, Rebecca E.
;    APPLICANT:  Klingler, Tod M.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  THYROID WITH ADENOMA
;    NUMBER OF SEQUENCES:  2070
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,526
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PD-0271 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 845-4166
;  INFORMATION FOR SEQ ID NO:  1051:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  250 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  2192904H1
US-08-961-526-1051

Alignment Scores:
Pred. No.:                18.8          Length:          250
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       19            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-961-526-1051 (1-250)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         73 CCAGATGTGATATTTCTACAGGAAGTT 99

RESULT 28
US-09-539-334-31143
; Sequence 31143, Application US/09539334
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF ENDOCRINE SYSTEM    TISSUE
;  FILE REFERENCE:  PD-1026 CIP
;  CURRENT APPLICATION NUMBER: US/09/539,334
;  CURRENT FILING DATE:  2000-03-30
;  "Prior application data removed - refer to PALM or file wrapper"
;  NUMBER OF SEQ ID NOS:  38381
;  SOFTWARE:  PERL Program
; SEQ ID NO 31143
;   LENGTH: 250
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00939287
;   NAME/KEY: unsure
;   LOCATION: 33
;   OTHER INFORMATION: a, t, c, g, or other
US-09-539-334-31143

Alignment Scores:
Pred. No.:                18.8          Length:          250
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       26            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-539-334-31143 (1-250)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         73 CCAGATGTGATATTTCTACAGGAAGTT 99

RESULT 29
US-08-276-163A-3853
; Sequence 3853, Application US/08276163A
;  GENERAL INFORMATION:
;    APPLICANT:  Haseltine, William
;    APPLICANT:  Fitzgerald, Lisa
;    APPLICANT:  Adams, Mark
;    APPLICANT:  Lee, Normal
;    APPLICANT:  Fuldner, Rebecca
;    APPLICANT:  Fleischmann, Robert
;    APPLICANT:  Bult, Carol
;    APPLICANT:  Blake, Judy
;    APPLICANT:  White, Owen
;    APPLICANT:  Clayton, Rebecca
;    APPLICANT:  Pellegrino, Susan
;    TITLE OF INVENTION:  Human Genes, Sequences, and Expression
;    NUMBER OF SEQUENCES:  15312
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart, &
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 5.0
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/276,163A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Olstein, Elliot
;      REGISTRATION NUMBER:  24,025
;      REFERENCE/DOCKET NUMBER:  325800-175
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 994-1700
;      TELEFAX:  (201) 994-1744
;  INFORMATION FOR SEQ ID NO:  3853:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  252 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-276-163A-3853

Alignment Scores:
Pred. No.:                19            Length:          252
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       12            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-276-163A-3853 (1-252)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          2 CCAGATGTGATATTTCTACAGGAAGTT 28

RESULT 30
US-08-276-163B-3853
```

```
; Sequence 3853, Application US/08276163B
; GENERAL INFORMATION:
;   APPLICANT:  Adams, Mark D.
;   APPLICANT:  Bult, Carol J.
;   APPLICANT:  Blake, Judith A.
;   APPLICANT:  Fitzgerald, Lisa
;   APPLICANT:  Fleischmann, Robert D.
;   APPLICANT:  Fraser, Claire M.
;   APPLICANT:  Fuldner, Rebecca A.
;   APPLICANT:  Gocayne, Jeannine D.
;   APPLICANT:  Kelley, Jenny
;   APPLICANT:  Kirkness, Ewen F.
;   APPLICANT:  Sutton, Granger G., III
;   APPLICANT:  Haseltine, William A.
;   APPLICANT:  Rosen, Craig A.
;   APPLICANT:  Ruben, Steve
;   APPLICANT:  Dillon, Patrick J.
;   APPLICANT:  Li, Haodong
;   TITLE OF INVENTION:  Human Genes, Sequences and Expression Products
;   NUMBER OF SEQUENCES:  15314
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Human Genome Sciences, Inc.
;     STREET:  9410 Key West Avenue
;     CITY:  Rockville
;     STATE:  Maryland
;     COUNTRY:  USA
;     ZIP:  20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;     COMPUTER:  HP Vectra 486/33
;     OPERATING SYSTEM:  MSDOS version 6.2
;     SOFTWARE:  ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/276,163B
;     FILING DATE:  July 15, 1994
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Michele M. Wales
;     REGISTRATION NUMBER:  P-43,975
;     REFERENCE/DOCKET NUMBER:  PO-14
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (301) 309-8504
;     TELEFAX:  (301) 309-8439
;  INFORMATION FOR SEQ ID NO:  3853:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  252 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
US-08-276-163B-3853

Alignment Scores:
Pred. No.:                19              Length:         252
Score:                    45.00           Matches:        9
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       12              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-276-163B-3853 (1-252)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          2 CCAGATGTGATATTTCTACAGGAAGTT 28

RESULT 31
US-08-276-163D-3853
; Sequence 3853, Application US/08276163D
; GENERAL INFORMATION:
;  APPLICANT: Adams, et. al.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
;  FILE REFERENCE: PO14
;  CURRENT APPLICATION NUMBER: US/08/276,163D
;  CURRENT FILING DATE:  1994-07-15
;  NUMBER OF SEQ ID NOS: 15314
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3853
;   LENGTH: 252
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-08-276-163D-3853

Alignment Scores:
Pred. No.:                19              Length:         252
Score:                    45.00           Matches:        9
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       12              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-276-163D-3853 (1-252)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          2 CCAGATGTGATATTTCTACAGGAAGTT 28

RESULT 32
US-09-840-145-3853
; Sequence 3853, Application US/09840145
; GENERAL INFORMATION:
;  APPLICANT: Adams, et. al.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
;  FILE REFERENCE: PO14C1
;  CURRENT APPLICATION NUMBER: US/09/840,145
;  CURRENT FILING DATE:  2001-04-24
;  PRIOR APPLICATION NUMBER: 08/276,163
;  PRIOR FILING DATE: 1994-07-15
;  NUMBER OF SEQ ID NOS: 15314
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3853
;   LENGTH: 252
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-840-145-3853

Alignment Scores:
Pred. No.:                19              Length:         252
Score:                    45.00           Matches:        9
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       33              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-840-145-3853 (1-252)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          2 CCAGATGTGATATTTCTACAGGAAGTT 28

RESULT 33
US-09-912-293-29561
; Sequence 29561, Application US/09912293
; GENERAL INFORMATION:
;  APPLICANT: Rosen, et. al.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 100
;  FILE REFERENCE: PO-100
;  CURRENT APPLICATION NUMBER: US/09/912,293
;  CURRENT FILING DATE:  2001-07-26
;  PRIOR APPLICATION NUMBER: 08/103,744
;  PRIOR FILING DATE: 1993-08-09
;  PRIOR APPLICATION NUMBER: 09/249,651
;  PRIOR FILING DATE: 1999-02-12
;  PRIOR APPLICATION NUMBER: 08/104,507
```

```
;  PRIOR FILING DATE: 1993-08-09
;  PRIOR APPLICATION NUMBER: 08/196,363
;  PRIOR FILING DATE: 1994-02-15
;  PRIOR APPLICATION NUMBER: 09/859,490
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: 08/196,362
;  PRIOR FILING DATE: 1994-02-15
;  PRIOR APPLICATION NUMBER: 08/221,623
;  PRIOR FILING DATE: 1994-03-31
;  PRIOR APPLICATION NUMBER: 08/220,691
;  PRIOR FILING DATE: 1994-03-31
;  PRIOR APPLICATION NUMBER: 09/741,830
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 09/813,155
;  PRIOR FILING DATE: 2001-03-21
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 244538
; SEQ ID NO 29561
;   LENGTH: 252
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-912-293-29561

Alignment Scores:
Pred. No.:                19              Length:         252
Score:                    45.00           Matches:        9
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       35              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-912-293-29561 (1-252)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          2 CCAGATGTGATATTTCTACAGGAAGTT 28

RESULT 34
US-08-834-645-54
; Sequence 54, Application US/08834645
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuve, Laura L.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Naughton, Rebecca E.
;    APPLICANT:  Klingler, Tod M.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  HUMAN COLON
;    NUMBER OF SEQUENCES:  1149
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,645
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/017,170
;      FILING DATE:  APRIL 10, 1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PD-0139 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 845-4166
;  INFORMATION FOR SEQ ID NO:  54:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  266 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  903518
US-08-834-645-54

Alignment Scores:
Pred. No.:                20.2            Length:         266
Score:                    45.00           Matches:        9
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       18              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-834-645-54 (1-266)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        157 CCAGATGTGATATTTCTACAGGAAGTT 183

RESULT 35
US-09-540-766-4618
; Sequence 4618, Application US/09540766
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE
;  FILE REFERENCE:  PD-1024 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,766
;  CURRENT FILING DATE:  2000-03-30
;  PRIOR APPLICATION DATA removed - refer tp PALM or File Wrapper
;  NUMBER OF SEQ ID NOS:  77960
;  SOFTWARE:  PERL Program
; SEQ ID NO 4618
;   LENGTH: 266
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00655028
;   NAME/KEY: unsure
;   LOCATION: 82, 146, 255
;   OTHER INFORMATION: a, t, c, g, or other
US-09-540-766-4618

Alignment Scores:
Pred. No.:                20.2            Length:         266
Score:                    45.00           Matches:        9
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       27              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-540-766-4618 (1-266)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        157 CCAGATGTGATATTTCTACAGGAAGTT 183
```

```
RESULT 36
US-08-917-046-231
; Sequence 231, Application US/08917046
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuve, Laura L.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Naughton, Rebecca E.
;    APPLICANT:  Klingler, Tod M.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  HUMAN COLON
;    NUMBER OF SEQUENCES:  2876
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/917,046
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/024,538
;      FILING DATE:  AUGUST 22, 1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PD-0226 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  231:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  292 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  1843738
US-08-917-046-231

Alignment Scores:
Pred. No.:                22.5          Length:          292
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       19            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-917-046-231 (1-292)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         23 CCAGATGTGATATTTCTACAGGAAGTT 49

RESULT 37
US-09-540-766-3985
; Sequence 3985, Application US/09540766
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE
;  FILE REFERENCE:  PD-1024 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,766
;  CURRENT FILING DATE:  2000-03-30
;  PRIOR APPLICATION DATA removed - refer tp PALM or File Wrapper
;  NUMBER OF SEQ ID NOS:  77960
;  SOFTWARE:  PERL Program
; SEQ ID NO 3985
;   LENGTH: 292
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00498694
US-09-540-766-3985

Alignment Scores:
Pred. No.:                22.5          Length:          292
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       27            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-540-766-3985 (1-292)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         23 CCAGATGTGATATTTCTACAGGAAGTT 49

RESULT 38
US-09-041-720-1771
; Sequence 1771, Application US/09041720
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuve, Laura L.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Naughton, Rebecca E.
;    APPLICANT:  Klingler, Tod M.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  ADRENAL GLAND TUMOR
;    NUMBER OF SEQUENCES:  4388
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/041,720
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/040,197
;      FILING DATE:  MARCH 14, 1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
```

```
;      REFERENCE/DOCKET NUMBER:  PD-0333 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1771:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  303 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  2824678H1
US-09-041-720-1771

Alignment Scores:
Pred. No.:                23.5          Length:         303
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       20            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-041-720-1771 (1-303)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         64 CCAGATGTGATATTTCTACAGGAAGTT 90

RESULT 39
US-09-539-334-1416
; Sequence 1416, Application US/09539334
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF ENDOCRINE SYSTEM    TISSUE
;  FILE REFERENCE:  PD-1026 CIP
;  CURRENT APPLICATION NUMBER: US/09/539,334
;  CURRENT FILING DATE:  2000-03-30
;  "Prior application data removed - refer to PALM or file wrapper"
;  NUMBER OF SEQ ID NOS:  38381
;  SOFTWARE:  PERL Program
; SEQ ID NO 1416
;   LENGTH: 303
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00810048
US-09-539-334-1416

Alignment Scores:
Pred. No.:                23.5          Length:         303
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       26            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-539-334-1416 (1-303)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         64 CCAGATGTGATATTTCTACAGGAAGTT 90

RESULT 40
US-60-040-197-1771
; Sequence 1771, Application US/60040197
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuve, Laura L.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Naughton, Rebecca E.
;    APPLICANT:  Klingler, Tod M.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  ADRENAL GLAND TUMOR
;    NUMBER OF SEQUENCES:  2437
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/60/040,197
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PD-0333P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 845-4166
;  INFORMATION FOR SEQ ID NO:  1771:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  303 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  2824678H1
US-60-040-197-1771

Alignment Scores:
Pred. No.:                23.5          Length:         303
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       74            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-040-197-1771 (1-303)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         64 CCAGATGTGATATTTCTACAGGAAGTT 90

RESULT 41
US-09-528-409-24161
; Sequence 24161, Application US/09528409
; GENERAL INFORMATION:
;  APPLICANT:  Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION:  Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE:  774
;  CURRENT APPLICATION NUMBER: US/09/528,409
;  CURRENT FILING DATE:  2000-03-17
```

```
;  PRIOR APPLICATION NUMBER:  60/125,453
;  PRIOR FILING DATE:  1999-03-19
;  NUMBER OF SEQ ID NOS:  116231
;  SOFTWARE:  Hy-patent.pl Version 3.1
; SEQ ID NO 24161
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-528-409-24161

Alignment Scores:
Pred. No.:                32.2          Length:         396
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       25            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-528-409-24161 (1-396)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        249 CCAGATGTGATATTTCTACAGGAAGTT 275

RESULT 42
US-09-933-524-24161
; Sequence 24161, Application US/09933524
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 24161
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524-24161

Alignment Scores:
Pred. No.:                32.2          Length:         396
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       36            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-933-524-24161 (1-396)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        249 CCAGATGTGATATTTCTACAGGAAGTT 275

RESULT 43
US-09-933-524A-24161
; Sequence 24161, Application US/09933524A
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524A
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 24161
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524A-24161

Alignment Scores:
Pred. No.:                32.2          Length:         396
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       36            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-933-524A-24161 (1-396)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        249 CCAGATGTGATATTTCTACAGGAAGTT 275

RESULT 44
US-09-528-409-26306
; Sequence 26306, Application US/09528409
; GENERAL INFORMATION:
;  APPLICANT:  Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION:  Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE:  774
;  CURRENT APPLICATION NUMBER: US/09/528,409
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER:  60/125,453
;  PRIOR FILING DATE:  1999-03-19
;  NUMBER OF SEQ ID NOS:  116231
;  SOFTWARE:  Hy-patent.pl Version 3.1
; SEQ ID NO 26306
;   LENGTH: 397
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-528-409-26306

Alignment Scores:
Pred. No.:                32.3          Length:         397
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       25            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-528-409-26306 (1-397)
```

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        282 CCAGATGTGATATTTCTACAGGAAGTT 308

RESULT 45
US-09-933-524-26306
; Sequence 26306, Application US/09933524
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 26306
;   LENGTH: 397
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-933-524-26306

Alignment Scores:
Pred. No.:                32.3          Length:          397
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       36            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-933-524-26306 (1-397)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        282 CCAGATGTGATATTTCTACAGGAAGTT 308

RESULT 46
US-09-933-524A-26306
; Sequence 26306, Application US/09933524A
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524A
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 26306
;   LENGTH: 397
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-933-524A-26306

Alignment Scores:
Pred. No.:                32.3          Length:          397
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       36            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-933-524A-26306 (1-397)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        282 CCAGATGTGATATTTCTACAGGAAGTT 308

RESULT 47
PCT-US02-02870-1363
; Sequence 1363, Application PC/TUS0202870
; GENERAL INFORMATION:
;  APPLICANT: Corixa Coporation
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563PC
;  CURRENT APPLICATION NUMBER: PCT/US02/02870
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1363
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US02-02870-1363

Alignment Scores:
Pred. No.:                32.7          Length:          401
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       1             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x PCT-US02-02870-1363 (1-401)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         55 CCAGATGTGATATTTCTACAGGAAGTT 81

RESULT 48
US-09-528-409-74640
; Sequence 74640, Application US/09528409
; GENERAL INFORMATION:
;  APPLICANT:  Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION:  Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE:  774
;  CURRENT APPLICATION NUMBER: US/09/528,409
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER:  60/125,453
;  PRIOR FILING DATE:  1999-03-19
;  NUMBER OF SEQ ID NOS:  116231
;  SOFTWARE:  Hy-patent.pl Version 3.1
; SEQ ID NO 74640
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-528-409-74640
```

```
Alignment Scores:
Pred. No.:                32.7          Length:          401
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       25            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-528-409-74640 (1-401)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         37 CCAGATGTGATATTTCTACAGGAAGTT 63

RESULT 49
US-09-933-524-74640
; Sequence 74640, Application US/09933524
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 74640
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-933-524-74640

Alignment Scores:
Pred. No.:                32.7          Length:          401
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       36            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-933-524-74640 (1-401)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         37 CCAGATGTGATATTTCTACAGGAAGTT 63

RESULT 50
US-09-933-524A-74640
; Sequence 74640, Application US/09933524A
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524A
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 74640
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-933-524A-74640

Alignment Scores:
Pred. No.:                32.7          Length:          401
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       36            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-933-524A-74640 (1-401)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         37 CCAGATGTGATATTTCTACAGGAAGTT 63

RESULT 51
US-10-066-543-1363
; Sequence 1363, Application US/10066543
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1363
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-066-543-1363

Alignment Scores:
Pred. No.:                32.7          Length:          401
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       40            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-066-543-1363 (1-401)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         55 CCAGATGTGATATTTCTACAGGAAGTT 81

RESULT 52
US-10-214-403-1363
; Sequence 1363, Application US/10214403
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Chenault, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
```

; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563C1
; CURRENT APPLICATION NUMBER: US/10/214,403
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 3420
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1363
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-214-403-1363

Alignment Scores:

Pred. No.:	32.7	Length:	401
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	43	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-214-403-1363 (1-401)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         55 CCAGATGTGATATTTCTACAGGAAGTT 81
```

RESULT 53
US-09-404-549-2518
; Sequence 2518, Application US/09404549
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN FETAL THYMUS LIBRARY
; FILE REFERENCE: MLN98-49pM
; CURRENT APPLICATION NUMBER: US/09/404,549
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60/101,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106,434
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107.258
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 6007
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2518
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-549-2518

Alignment Scores:

Pred. No.:	33.9	Length:	414
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	24	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-404-549-2518 (1-414)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        103 CCAGATGTGATATTTCTACAGGAAGTT 129
```

RESULT 54
US-09-404-549A-2518
; Sequence 2518, Application US/09404549A
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN FETAL THYMUS LIBRARY
; FILE REFERENCE: MLN98-49pM
; CURRENT APPLICATION NUMBER: US/09/404,549A
; CURRENT FILING DATE: 1999-09-24
; PRIOR APPLICATION NUMBER: 60/101,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106,434
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107.258
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 6007
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2518
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-549A-2518

Alignment Scores:

Pred. No.:	33.9	Length:	414
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	24	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-404-549A-2518 (1-414)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        103 CCAGATGTGATATTTCTACAGGAAGTT 129
```

RESULT 55
US-09-824-559-4370
; Sequence 4370, Application US/09824559
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Jakubowski, Joseph A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2077-001
; CURRENT APPLICATION NUMBER: US/09/824,559
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/193,437
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9929
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4370
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-824-559-4370

Alignment Scores:

Pred. No.:	34.2	Length:	417
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	33	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-824-559-4370 (1-417)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        387 CCAGATGTGATATTTCTACAGGAAGTT 413
```

RESULT 56
US-09-359-067-7230
; Sequence 7230, Application US/09359067

```
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-747CON1
;  CURRENT APPLICATION NUMBER: US/09/359,067
;  CURRENT FILING DATE: 1999-07-22
;  EARLIER APPLICATION NUMBER: US 09/131,598
;  EARLIER FILING DATE: 1998-08-10
;  NUMBER OF SEQ ID NOS: 49786
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7230
;   LENGTH: 418
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(418)
;   OTHER INFORMATION: n = A,T,C or G
US-09-359-067-7230

Alignment Scores:
Pred. No.:                34.3          Length:          418
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       23            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-359-067-7230 (1-418)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         75 CCAGATGTGATATTTCTACAGGAAGTT 101

RESULT 57
US-09-528-409-24160
; Sequence 24160, Application US/09528409
; GENERAL INFORMATION:
;  APPLICANT:  Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION:  Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE:  774
;  CURRENT APPLICATION NUMBER: US/09/528,409
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER:  60/125,453
;  PRIOR FILING DATE:  1999-03-19
;  NUMBER OF SEQ ID NOS:  116231
;  SOFTWARE:  Hy-patent.pl Version 3.1
; SEQ ID NO 24160
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(420)
;   OTHER INFORMATION: n = A,T,C or G
US-09-528-409-24160

Alignment Scores:
Pred. No.:                34.5          Length:          420
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       25            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-528-409-24160 (1-420)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        248 CCAGATGTGATATTTCTACAGGAAGTT 274

RESULT 58
US-09-933-524-24160
; Sequence 24160, Application US/09933524
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 24160
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(420)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524-24160

Alignment Scores:
Pred. No.:                34.5          Length:          420
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       36            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-933-524-24160 (1-420)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        248 CCAGATGTGATATTTCTACAGGAAGTT 274

RESULT 59
US-09-933-524A-24160
; Sequence 24160, Application US/09933524A
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524A
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 24160
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(420)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524A-24160
```

Alignment Scores:
Pred. No.: 34.5 Length: 420
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 36 Gaps: 0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-933-524A-24160 (1-420)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        248 CCAGATGTGATATTTCTACAGGAAGTT 274
```

RESULT 60
US-09-528-409-74096
; Sequence 74096, Application US/09528409
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/528,409
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,453
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 74096
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-528-409-74096

Alignment Scores:
Pred. No.: 35.7 Length: 433
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 25 Gaps: 0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-528-409-74096 (1-433)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        353 CCAGATGTGATATTTCTACAGGAAGTT 379
```

RESULT 61
US-09-933-524-74096
; Sequence 74096, Application US/09933524
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 74096
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-933-524-74096

Alignment Scores:
Pred. No.: 35.7 Length: 433
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 36 Gaps: 0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-933-524-74096 (1-433)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        353 CCAGATGTGATATTTCTACAGGAAGTT 379
```

RESULT 62
US-09-933-524A-74096
; Sequence 74096, Application US/09933524A
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524A
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 74096
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-933-524A-74096

Alignment Scores:
Pred. No.: 35.7 Length: 433
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 36 Gaps: 0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-933-524A-74096 (1-433)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        353 CCAGATGTGATATTTCTACAGGAAGTT 379
```

RESULT 63
US-09-864-761-16762
; Sequence 16762, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26

```
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16762
;   LENGTH: 444
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
US-09-864-761-16762

Alignment Scores:
Pred. No.:                36.8          Length:          444
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       33            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-864-761-16762 (1-444)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        260 CCAGATGTGATATTTCTACAGGAAGTT 286

RESULT 64
US-10-182-997-5244
; Sequence 5244, Application US/10182997
; GENERAL INFORMATION:
;  APPLICANT: Molecular Dynamics, Inc.
;  APPLICANT:  Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  ANALYSIS OF GENE EXPRESSION IN HELA
;  FILE REFERENCE: PB 0004 WO 10
;  CURRENT APPLICATION NUMBER: US/10/182,997
;  CURRENT FILING DATE:  2002-08-02
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 04 February 2000 (04.02.00)
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 26 May 2000 (26.05.00)
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 03 August 2000 (03.08.00)
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 03 October 2000 (03.10.00)
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 27 September 2000 (27.09.00)
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 21 September 2000 (21.09.00)
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 30 June 2000 (30.06.00)
;  NUMBER OF SEQ ID NOS: 26941
;  SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 5244
;   LENGTH: 444
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   FEATURE:
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
US-10-182-997-5244

Alignment Scores:
Pred. No.:                36.8          Length:          444
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       42            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-182-997-5244 (1-444)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        260 CCAGATGTGATATTTCTACAGGAAGTT 286

RESULT 65
US-60-236-359-6954
; Sequence 6954, Application US/60236359
; GENERAL INFORMATION:
;  APPLICANT:  Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION:  HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE:  MDhMORF-4P
;  CURRENT APPLICATION NUMBER: US/60/236,359
;  CURRENT FILING DATE:  2000-09-27
;  PRIOR APPLICATION NUMBER:  US 60/207,456
;  PRIOR FILING DATE:  2000-05-26
;  NUMBER OF SEQ ID NOS:  21709
;  SOFTWARE:  Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 6954
;   LENGTH: 444
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
;   OTHER INFORMATION: EST_HUMAN HIT: AA325145.1, EVALUE 7.00e-42
;   OTHER INFORMATION: NT HIT: AJ269473.1, EVALUE 2.00e-45
;   OTHER INFORMATION: SWISSPROT HIT: P18004, EVALUE 3.40e-02
US-60-236-359-6954

Alignment Scores:
Pred. No.:                36.8          Length:          444
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
```

```
DB:                    77            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-236-359-6954 (1-444)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        260 CCAGATGTGATATTTCTACAGGAAGTT 286

RESULT 66
US-09-304-649-421
; Sequence 421, Application US/09304649
; GENERAL INFORMATION:
;  APPLICANT: Gearing, David P.
;  APPLICANT: McCarthy, Sean A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
;  TITLE OF INVENTION: HUMAN MICROVASCULAR ENDOTHELIAL LIBRARY
;  FILE REFERENCE: MLN98-10pM
;  CURRENT APPLICATION NUMBER: US/09/304,649
;  CURRENT FILING DATE: 1999-05-04
;  EARLIER APPLICATION NUMBER: 60/084,098
;  EARLIER FILING DATE: 1998-05-04
;  EARLIER APPLICATION NUMBER: 60/123,523
;  EARLIER FILING DATE: 1999-03-08
;  EARLIER APPLICATION NUMBER: 60/126,974
;  EARLIER FILING DATE: 1999-03-30
;  NUMBER OF SEQ ID NOS: 1311
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 421
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(449)
;   OTHER INFORMATION: n = A,T,C or G
US-09-304-649-421

Alignment Scores:
Pred. No.:              37.3          Length:        449
Score:                  45.00         Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            100.0%        Indels:        0
DB:                     23            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-304-649-421 (1-449)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        137 CCAGATGTGATATTTCTACAGGAAGTT 163

RESULT 67
US-09-840-424-421
; Sequence 421, Application US/09840424
; GENERAL INFORMATION:
;  APPLICANT: Gearing, David P.
;  APPLICANT:  McCarthy, Sean A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
;  TITLE OF INVENTION:  HUMAN MICROVASCULAR ENDOTHELIAL LIBRARY
;  FILE REFERENCE: 1600.1010-002
;  CURRENT APPLICATION NUMBER: US/09/840,424
;  CURRENT FILING DATE:  2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/084,098
;  PRIOR FILING DATE: 1998-05-04
;  PRIOR APPLICATION NUMBER: US 60/123,523
;  PRIOR FILING DATE: 1999-03-08
;  PRIOR APPLICATION NUMBER: US 60/126,974
;  PRIOR FILING DATE: 1999-03-30
;  PRIOR APPLICATION NUMBER: US 09/304,649
;  PRIOR FILING DATE: 1999-05-04
;  NUMBER OF SEQ ID NOS: 1311
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 421
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(449)
;   OTHER INFORMATION: n = A,T,C or G
US-09-840-424-421

Alignment Scores:
Pred. No.:              37.3          Length:        449
Score:                  45.00         Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            100.0%        Indels:        0
DB:                     33            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-840-424-421 (1-449)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        137 CCAGATGTGATATTTCTACAGGAAGTT 163

RESULT 68
US-08-905-144-190
; Sequence 190, Application US/08905144
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Jean-Baptiste D.
;    APPLICANT:  Duelert, Aymeric
;    APPLICANT:  Lacroix, Bruno
;    TITLE OF INVENTION:  5' ESTs FOR SECRETED PROTEINS
;    TITLE OF INVENTION:  EXPRESSED IN PROSTATE
;    NUMBER OF SEQUENCES:  593
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  501 West Broadway
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-3505
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Win95
;      SOFTWARE:  Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/905,144
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned A.
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  GENSET.010A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  190:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  483 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  DOUBLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  CDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo Sapiens
;      TISSUE TYPE:  Normal prostate
;    FEATURE:
;      NAME/KEY:  other
;      LOCATION:  105..414
;      IDENTIFICATION METHOD:  blastn
;      OTHER INFORMATION:  identity 99
;      OTHER INFORMATION:  region 1..310
;      OTHER INFORMATION:  id T26956
```

```
;      OTHER INFORMATION:  est
;    FEATURE:
;      NAME/KEY:  other
;      LOCATION:  45..359
;      IDENTIFICATION METHOD:  blastn
;      OTHER INFORMATION:  identity 98
;      OTHER INFORMATION:  region 1..315
;      OTHER INFORMATION:  id T31666
;      OTHER INFORMATION:  est
;    FEATURE:
;      NAME/KEY:  other
;      LOCATION:  202..332
;      IDENTIFICATION METHOD:  blastn
;      OTHER INFORMATION:  identity 100
;      OTHER INFORMATION:  region 137..267
;      OTHER INFORMATION:  id R14990
;      OTHER INFORMATION:  est
;    FEATURE:
;      NAME/KEY:  other
;      LOCATION:  127..201
;      IDENTIFICATION METHOD:  blastn
;      OTHER INFORMATION:  identity 100
;      OTHER INFORMATION:  region 63..137
;      OTHER INFORMATION:  id R14990
;      OTHER INFORMATION:  est
;    FEATURE:
;      NAME/KEY:  other
;      LOCATION:  65..114
;      IDENTIFICATION METHOD:  blastn
;      OTHER INFORMATION:  identity 100
;      OTHER INFORMATION:  region 1..50
;      OTHER INFORMATION:  id R14990
;      OTHER INFORMATION:  est
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  1..120
;      IDENTIFICATION METHOD:  Von Heijne matrix
;      OTHER INFORMATION:  score 5.2
;      OTHER INFORMATION:  seq LCVEFASVASCDA/AV
US-08-905-144-190

Alignment Scores:
Pred. No.:                40.6          Length:          483
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       19            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-905-144-190 (1-483)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 69
US-09-547-599C-825
; Sequence 825, Application US/09547599C
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Duclert, Aymeric
;  APPLICANT:  Lacroix, Bruno
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
;  FILE REFERENCE: GEN-T119C1
;  CURRENT APPLICATION NUMBER: US/09/547,599C
;  CURRENT FILING DATE:  2000-04-11
;  PRIOR APPLICATION NUMBER: US 08/905,223
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,135
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,051
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,144
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,279
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/904,468
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,134
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,133
;  PRIOR FILING DATE: 1997-08-01
;  NUMBER OF SEQ ID NOS: 3475
;  SOFTWARE: Patent.pm
; SEQ ID NO 825
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: Normal prostate
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 1..120
;   OTHER INFORMATION: Von Heijne matrix
;   OTHER INFORMATION: score 5.20
;   FEATURE:
;   OTHER INFORMATION: seq LCVEFASVASCDA/AV
;   NAME/KEY: misc_feature
;   LOCATION: 105..414
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 99
;   OTHER INFORMATION: region 1..310
;   OTHER INFORMATION: id T26956
;   FEATURE:
;   OTHER INFORMATION: est
;   NAME/KEY: misc_feature
;   LOCATION: 45..359
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 98
;   OTHER INFORMATION: region 1..315
;   OTHER INFORMATION: id T31666
;   FEATURE:
;   OTHER INFORMATION: est
;   NAME/KEY: misc_feature
;   LOCATION: 202..332
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 137..267
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   OTHER INFORMATION: est
;   NAME/KEY: misc_feature
;   LOCATION: 127..201
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 63..137
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   OTHER INFORMATION: est
;   NAME/KEY: misc_feature
;   LOCATION: 65..114
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 1..50
;   OTHER INFORMATION: id R14990
;   OTHER INFORMATION: est
US-09-547-599C-825

Alignment Scores:
Pred. No.:                40.6          Length:          483
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       27            Gaps:            0
```

```
US-10-757-745-2_COPY_145_153 (1-9) x US-09-547-599C-825 (1-483)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 70
US-10-631-402-825
; Sequence 825, Application US/10631402
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Duclert, Aymeric
;  APPLICANT:  Lacroix, Bruno
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
;  FILE REFERENCE: GEN-T119C1
;  CURRENT APPLICATION NUMBER: US/10/631,402
;  CURRENT FILING DATE:  2003-07-30
;  PRIOR APPLICATION NUMBER: US/09/547,599C
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 08/905,223
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,135
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,051
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,144
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,279
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/904,468
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,134
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,133
;  PRIOR FILING DATE: 1997-08-01
;  NUMBER OF SEQ ID NOS: 3475
;  SOFTWARE: Patent.pm
; SEQ ID NO 825
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: Normal prostate
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 1..120
;   OTHER INFORMATION: Von Heijne matrix
;   FEATURE:
;   OTHER INFORMATION: score 5.20
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 105..414
;   OTHER INFORMATION: seq LCVEFASVASCDA/AV
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 99
;   OTHER INFORMATION: region 1..310
;   FEATURE:
;   OTHER INFORMATION: id T26956
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 45..359
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 98
;   OTHER INFORMATION: region 1..315
;   FEATURE:
;   OTHER INFORMATION: id T31666
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 202..332
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 137..267
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 127..201
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 63..137
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 65..114
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 1..50
;   OTHER INFORMATION: id R14990
;   OTHER INFORMATION: est
US-10-631-402-825

Alignment Scores:
Pred. No.:                40.6          Length:          483
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       54            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-631-402-825 (1-483)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 71
US-10-631-441-825
; Sequence 825, Application US/10631441
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Duclert, Aymeric
;  APPLICANT:  Lacroix, Bruno
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
;  FILE REFERENCE: GEN-T119C1
;  CURRENT APPLICATION NUMBER: US/10/631,441
;  CURRENT FILING DATE:  2003-07-30
;  PRIOR APPLICATION NUMBER: US/09/547,599C
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 08/905,223
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,135
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,051
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,144
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,279
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/904,468
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,134
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,133
;  PRIOR FILING DATE: 1997-08-01
;  NUMBER OF SEQ ID NOS: 3475
;  SOFTWARE: Patent.pm
; SEQ ID NO 825
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
```

```
;   OTHER INFORMATION: Normal prostate
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 1..120
;   OTHER INFORMATION: Von Heijne matrix
;   FEATURE:
;   OTHER INFORMATION: score 5.20
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 105..414
;   OTHER INFORMATION: seq LCVEFASVASCDA/AV
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 99
;   OTHER INFORMATION: region 1..310
;   FEATURE:
;   OTHER INFORMATION: id T26956
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 45..359
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 98
;   OTHER INFORMATION: region 1..315
;   FEATURE:
;   OTHER INFORMATION: id T31666
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 202..332
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 137..267
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 127..201
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 63..137
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 65..114
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 1..50
;   OTHER INFORMATION: id R14990
;   OTHER INFORMATION: est
US-10-631-441-825

Alignment Scores:
Pred. No.:                40.6          Length:          483
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       54            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-631-441-825 (1-483)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 72
US-10-631-441A-825
; Sequence 825, Application US/10631441A
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Duclert, Aymeric
;  APPLICANT:  Lacroix, Bruno
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
;  FILE REFERENCE: GEN-T119C1
;  CURRENT APPLICATION NUMBER: US/10/631,441A
;  CURRENT FILING DATE:  2003-07-30
;  PRIOR APPLICATION NUMBER: US/09/547,599C
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 08/905,223
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,135
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,051
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,144
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,279
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/904,468
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,134
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,133
;  PRIOR FILING DATE: 1997-08-01
;  NUMBER OF SEQ ID NOS: 3475
;  SOFTWARE: Patent.pm
; SEQ ID NO 825
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: Normal prostate
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 1..120
;   OTHER INFORMATION: Von Heijne matrix
;   FEATURE:
;   OTHER INFORMATION: score 5.20
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 105..414
;   OTHER INFORMATION: seq LCVEFASVASCDA/AV
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 99
;   OTHER INFORMATION: region 1..310
;   FEATURE:
;   OTHER INFORMATION: id T26956
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 45..359
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 98
;   OTHER INFORMATION: region 1..315
;   FEATURE:
;   OTHER INFORMATION: id T31666
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 202..332
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 137..267
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 127..201
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 63..137
;   FEATURE:
```

```
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 65..114
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 1..50
;   OTHER INFORMATION: id R14990
;   OTHER INFORMATION: est
US-10-631-441A-825

Alignment Scores:
Pred. No.:                40.6          Length:          483
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       54            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-631-441A-825 (1-483)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 73
US-09-606-776-1036
; Sequence 1036, Application US/09606776
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  APPLICANT:  Myers, Paul
;  APPLICANT:  Gearing, David P.
;  APPLICANT:  Pan, Yang
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 1600.1129-001
;  CURRENT APPLICATION NUMBER: US/09/606,776
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER: 60/141,578
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 60/141,379
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: 60/141,138
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: 60/141,581
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 5415
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1036
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-606-776-1036

Alignment Scores:
Pred. No.:                40.8          Length:          485
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       28            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-606-776-1036 (1-485)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        182 CCAGATGTGATATTTCTACAGGAAGTT 208

RESULT 74
US-09-644-868-6183
; Sequence 6183, Application US/09644868
; GENERAL INFORMATION:
;  APPLICANT:  White, David
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Pan, Yang
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1168-001
;  CURRENT APPLICATION NUMBER: US/09/644,868
;  CURRENT FILING DATE:  2000-08-28
;  PRIOR APPLICATION NUMBER:  60/151,063
;  PRIOR FILING DATE:  1999-08-27
;  NUMBER OF SEQ ID NOS:  10075
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6183
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-644-868-6183

Alignment Scores:
Pred. No.:                40.8          Length:          485
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       29            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-644-868-6183 (1-485)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        182 CCAGATGTGATATTTCTACAGGAAGTT 208

RESULT 75
US-60-217-080-8309
; Sequence 8309, Application US/60217080
; GENERAL INFORMATION:
;  APPLICANT:  Glenn, Matthew
;  APPLICANT:  Grigor, Murray R.
;  TITLE OF INVENTION:  Compositions isolated from bovine tissue
;  TITLE OF INVENTION:  and methods for their use.
;  FILE REFERENCE:  1051P2
;  CURRENT APPLICATION NUMBER: US/60/217,080
;  CURRENT FILING DATE:  2000-07-10
;  NUMBER OF SEQ ID NOS:  35169
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8309
;   LENGTH: 502
;   TYPE: DNA
;   ORGANISM: Bovine
US-60-217-080-8309

Alignment Scores:
Pred. No.:                42.5          Length:          502
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       76            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-217-080-8309 (1-502)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        347 CCTGATGTGATATTTCTACAGGAAGTC 373


Search completed: March 11, 2006, 12:05:45
Job time : 373.75 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:22:00 ; Search time 21.0571 Seconds
(without alignments)
1013.401 Million cell updates/sec

Title: US-10-757-745-2_COPY_145_153
Perfect score: 45
Sequence: 1 PDVIFLQEV 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5007642 seqs, 1185514400 residues

Total number of hits satisfying chosen parameters: 10015284

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170335_4211/app_query.fasta_1
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_240_@runat_09032006_170335_4211 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq1:*
9: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	45	100.0	752	6	US-10-917-503B-4908	Sequence 4908, Ap
	2	45	100.0	1288	9	US-11-266-748A-226184	Sequence 226184,
	3	45	100.0	1898	6	US-10-917-503B-13209	Sequence 13209, A
	4	45	100.0	1940	6	US-10-567-867-921	Sequence 921, App
	5	45	100.0	1940	11	US-60-742-219-2071	Sequence 2071, Ap
c	6	37	82.2	424	9	US-11-313-816-1993	Sequence 1993, Ap
	7	37	82.2	435	9	US-11-266-748A-429406	Sequence 429406,
c	8	37	82.2	976	6	US-10-953-349-8407	Sequence 8407, Ap
	9	37	82.2	1482	6	US-10-953-349-21907	Sequence 21907, A
	10	37	82.2	208765	6	US-10-540-898-407	Sequence 407, App
	11	36	80.0	370	7	US-11-330-082-14845	Sequence 14845, A
c	12	36	80.0	1000	9	US-11-266-748A-294777	Sequence 294777,
	13	36	80.0	1000	9	US-11-266-748A-346206	Sequence 346206,
c	14	36	80.0	1000	9	US-11-266-748A-407208	Sequence 407208,
	15	36	80.0	1000	9	US-11-266-748A-478254	Sequence 478254,
	16	36	80.0	1318	6	US-10-516-032-11	Sequence 11, Appl
c	17	36	80.0	2508	11	US-60-742-219-1191	Sequence 1191, Ap
	18	36	80.0	2813	6	US-10-516-032-9	Sequence 9, Appli
	19	36	80.0	2854	7	US-11-293-697-1656	Sequence 1656, Ap
	20	36	80.0	3100	6	US-10-516-032-7	Sequence 7, Appli
	21	36	80.0	3100	7	US-11-145-307A-183	Sequence 183, App
	22	36	80.0	3100	9	US-11-266-748A-57213	Sequence 57213, A
	23	36	80.0	3101	6	US-10-531-147-2947	Sequence 2947, Ap
	24	36	80.0	2944528	9	US-11-045-004-1	Sequence 1, Appli
c	25	35	77.8	411	7	US-11-331-019-24422	Sequence 24422, A
c	26	35	77.8	444	7	US-11-331-032-4211	Sequence 4211, Ap
c	27	35	77.8	578	9	US-11-292-078-15903	Sequence 15903, A
	28	35	77.8	1000	9	US-11-266-748A-408169	Sequence 408169,
c	29	35	77.8	1000	9	US-11-266-748A-479215	Sequence 479215,
	30	35	77.8	1239	9	US-11-266-748A-25946	Sequence 25946, A
	31	35	77.8	1963	7	US-11-342-366-1917	Sequence 1917, Ap
	32	35	77.8	1963	7	US-11-342-367-1917	Sequence 1917, Ap
	33	35	77.8	1988	6	US-10-461-673-568	Sequence 568, App
	34	35	77.8	1995	9	US-11-266-748A-32741	Sequence 32741, A
	35	35	77.8	2163	11	US-60-751-420-1685	Sequence 1685, Ap
	36	34	75.6	226	9	US-11-314-006-3974	Sequence 3974, Ap
	37	34	75.6	228	9	US-11-314-006-1805	Sequence 1805, Ap
	38	34	75.6	334	9	US-11-314-006-8	Sequence 8, Appli
	39	34	75.6	371	7	US-11-353-150-18282	Sequence 18282, A
	40	34	75.6	371	9	US-11-314-006-3002	Sequence 3002, Ap
	41	34	75.6	374	9	US-11-314-006-626	Sequence 626, App
	42	34	75.6	375	9	US-11-314-006-598	Sequence 598, App
	43	34	75.6	377	9	US-11-314-006-2239	Sequence 2239, Ap
	44	34	75.6	380	9	US-11-314-006-1145	Sequence 1145, Ap
	45	34	75.6	382	9	US-11-314-006-2650	Sequence 2650, Ap
	46	34	75.6	384	9	US-11-314-006-3428	Sequence 3428, Ap
	47	34	75.6	387	9	US-11-314-006-2731	Sequence 2731, Ap
c	48	34	75.6	388	7	US-11-331-019-34748	Sequence 34748, A
	49	34	75.6	388	9	US-11-314-006-45	Sequence 45, Appl
	50	34	75.6	396	9	US-11-314-006-1103	Sequence 1103, Ap
	51	34	75.6	400	9	US-11-314-006-2849	Sequence 2849, Ap
	52	34	75.6	401	9	US-11-314-006-3724	Sequence 3724, Ap
	53	34	75.6	403	9	US-11-314-006-3916	Sequence 3916, Ap
	54	34	75.6	404	9	US-11-314-006-54	Sequence 54, Appl
	55	34	75.6	408	9	US-11-314-006-1295	Sequence 1295, Ap
	56	34	75.6	409	9	US-11-314-006-2829	Sequence 2829, Ap
	57	34	75.6	421	9	US-11-314-006-2227	Sequence 2227, Ap
	58	34	75.6	422	9	US-11-314-006-3827	Sequence 3827, Ap
	59	34	75.6	423	6	US-10-998-197A-31	Sequence 31, Appl
	60	34	75.6	423	9	US-11-314-006-1619	Sequence 1619, Ap
	61	34	75.6	431	9	US-11-314-006-4009	Sequence 4009, Ap
	62	34	75.6	434	9	US-11-314-006-2315	Sequence 2315, Ap
	63	34	75.6	437	9	US-11-314-006-3781	Sequence 3781, Ap
	64	34	75.6	437	9	US-11-266-748A-270053	Sequence 270053,
c	65	34	75.6	437	9	US-11-266-748A-330570	Sequence 330570,
	66	34	75.6	442	9	US-11-314-006-2378	Sequence 2378, Ap
	67	34	75.6	449	9	US-11-314-006-910	Sequence 910, App
	68	34	75.6	454	9	US-11-314-006-915	Sequence 915, App
	69	34	75.6	454	9	US-11-314-006-2262	Sequence 2262, Ap
	70	34	75.6	456	9	US-11-314-006-317	Sequence 317, App
	71	34	75.6	473	9	US-11-314-006-951	Sequence 951, App
	72	34	75.6	486	9	US-11-313-816-17804	Sequence 17804, A
c	73	34	75.6	536	9	US-11-266-748A-100449	Sequence 100449,
	74	34	75.6	536	9	US-11-266-748A-153260	Sequence 153260,
c	75	34	75.6	568	9	US-11-266-748A-55838	Sequence 55838, A

ALIGNMENTS

RESULT 1

US-10-917-503B-4908
; Sequence 4908, Application US/10917503B
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/10/917,503B
; CURRENT FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US/09/629,469
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4908
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (685)..(685)
; OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (694)..(694)
; OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (740)..(740)
; OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-10-917-503B-4908

Alignment Scores:

Pred. No.:	1.76	Length:	752
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-917-503B-4908 (1-752)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459
```

RESULT 2
US-11-266-748A-226184
; Sequence 226184, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 226184
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-226184

Alignment Scores:

Pred. No.:	3.26	Length:	1288
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-226184 (1-1288)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        285 CCAGATGTGATATTTCTACAGGAAGTT 311
```

RESULT 3
US-10-917-503B-13209
; Sequence 13209, Application US/10917503B
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/10/917,503B
; CURRENT FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US/09/629,469
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767

```
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
;   LENGTH: 1898
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1086)
US-10-917-503B-13209

Alignment Scores:
Pred. No.:                5.08          Length:          1898
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-917-503B-13209 (1-1898)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 4
US-10-567-867-921
; Sequence 921, Application US/10567867
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PCT
;  CURRENT APPLICATION NUMBER: US/10/567,867
;  CURRENT FILING DATE:  2006-01-27
;  PRIOR APPLICATION NUMBER: US 60/490,890
;  PRIOR FILING DATE: 2003-07-29
;  NUMBER OF SEQ ID NOS: 2786
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-567-867-921

Alignment Scores:
Pred. No.:                5.21          Length:          1940
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-567-867-921 (1-1940)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        458 CCAGATGTGATATTTCTACAGGAAGTT 484

RESULT 5
US-60-742-219-2071
; Sequence 2071, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2071
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-2071

Alignment Scores:
Pred. No.:                5.21          Length:          1940
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       11            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-742-219-2071 (1-1940)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        458 CCAGATGTGATATTTCTACAGGAAGTT 484

RESULT 6
US-11-313-816-1993/c
; Sequence 1993, Application US/11313816
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Ruan, Yijun G.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.353 - 38-21(15404)C/US
;  CURRENT APPLICATION NUMBER: US/11/313,816
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: 09/333,534
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 18718
; SEQ ID NO 1993
;   LENGTH: 424
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB22-066-Q1-E1-H2
US-11-313-816-1993

Alignment Scores:
Pred. No.:                60.6          Length:          424
Score:                    37.00         Matches:         6
Percent Similarity:       100.0%        Conservative:    2
Best Local Similarity:    75.0%         Mismatches:      0
Query Match:              82.2%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-313-816-1993 (1-424)

Qy          1 ProAspValIlePheLeuGlnGlu 8
```

```
                  |||||||||:::|||||||||:::
Db        208 CCGGATGTAGTTTTCTTGCAAGAC 185

RESULT 7
US-11-266-748A-429406
; Sequence 429406, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 429406
;   LENGTH: 435
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (402)..(402)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-429406

Alignment Scores:
Pred. No.:                62.4          Length:          435
Score:                    37.00         Matches:         7
Percent Similarity:       87.5%         Conservative:    0
Best Local Similarity:    87.5%         Mismatches:      1
Query Match:              82.2%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-429406 (1-435)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||   ||||||||||||
Db        179 CCTGATGTCTTTTTTCTGCAGGAG 202

RESULT 8
US-10-953-349-8407/c
; Sequence 8407, Application US/10953349
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCONDED THERBY
;  FILE REFERENCE: 2750-1579PUS2
;  CURRENT APPLICATION NUMBER: US/10/953,349
;  CURRENT FILING DATE:  2004-09-30
;  NUMBER OF SEQ ID NOS: 40252
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8407
;   LENGTH: 976
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-10-953-349-8407

Alignment Scores:
Pred. No.:                157           Length:          976
Score:                    37.00         Matches:         7
Percent Similarity:       88.9%         Conservative:    1
Best Local Similarity:    77.8%         Mismatches:      1
Query Match:              82.2%         Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-953-349-8407 (1-976)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||   ||||||
Db        481 CCTGACATCATCTTCCTTCTTGAGGTT 455

RESULT 9
US-10-953-349-21907
; Sequence 21907, Application US/10953349
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCONDED THERBY
;  FILE REFERENCE: 2750-1579PUS2
;  CURRENT APPLICATION NUMBER: US/10/953,349
;  CURRENT FILING DATE:  2004-09-30
;  NUMBER OF SEQ ID NOS: 40252
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21907
;   LENGTH: 1482
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-953-349-21907

Alignment Scores:
Pred. No.:                254           Length:          1482
Score:                    37.00         Matches:         7
Percent Similarity:       88.9%         Conservative:    1
Best Local Similarity:    77.8%         Mismatches:      1
Query Match:              82.2%         Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-953-349-21907 (1-1482)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||   ||||||||||||:::
Db        377 CCCGACGTTCCCTTCCTCCAAGAAATC 403

RESULT 10
US-10-540-898-407
; Sequence 407, Application US/10540898
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: CHIR0056-101 (PP023367.0003)
;  CURRENT APPLICATION NUMBER: US/10/540,898
;  CURRENT FILING DATE:  2005-06-27
;  PRIOR APPLICATION NUMBER: US 10/330,773
;  PRIOR FILING DATE: 2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 407
;   LENGTH: 208765
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(208765)
;   OTHER INFORMATION: n = A,T,C or G
US-10-540-898-407

Alignment Scores:
```

Pred. No.:	7.3e+04	Length:	208765
Score:	37.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	82.2%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-540-898-407 (1-208765)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||:::
Db     113412 CCAGACATAATCTTTCTGCAGCAG 113435
```

RESULT 11
US-11-330-082-14845
; Sequence 14845, Application US/11330082
; GENERAL INFORMATION:
; APPLICANT: Buehler, Robert E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Coombs, Brian E.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.356 - 38-21(15760)C/US
; CURRENT APPLICATION NUMBER: US/11/330,082
; CURRENT FILING DATE: 2006-01-12
; PRIOR APPLICATION NUMBER: US 09/552,086
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 27278
; SEQ ID NO 14845
; LENGTH: 370
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: jC-gmle01810082e01a1
US-11-330-082-14845

Pred. No.:	87.6	Length:	370
Score:	36.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-330-082-14845 (1-370)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||   ||||||||||||
Db        239 CCAGACATTATTTGTCTTCAGGAGGTT 265
```

RESULT 12
US-11-266-748A-294777/c
; Sequence 294777, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 294777
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-294777

Pred. No.:	273	Length:	1000
Score:	36.00	Matches:	7
Percent Similarity:	87.5%	Conservative:	0
Best Local Similarity:	87.5%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-294777 (1-1000)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||||||||||   |||
Db        532 CCTGATGTCATTTTTCTGCATGAA 509
```

RESULT 13
US-11-266-748A-346206
; Sequence 346206, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 346206
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-346206

Pred. No.:	273	Length:	1000
Score:	36.00	Matches:	7
Percent Similarity:	87.5%	Conservative:	0
Best Local Similarity:	87.5%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-346206 (1-1000)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||||||   |||
Db        469 CCTGATGTCATTTTTCTGCATGAA 492

RESULT 14
US-11-266-748A-407208/c
; Sequence 407208, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 407208
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-407208

Alignment Scores:
Pred. No.:              273          Length:         1000
Score:                  36.00        Matches:        7
Percent Similarity:     87.5%        Conservative:   0
Best Local Similarity:  87.5%        Mismatches:     1
Query Match:            80.0%        Indels:         0
DB:                     9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-407208 (1-1000)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||||||   |||
Db        532 CCTGATGTCATTTTTCTGCATGAA 509

RESULT 15
US-11-266-748A-478254
; Sequence 478254, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 478254
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-478254

Alignment Scores:
Pred. No.:              273          Length:         1000
Score:                  36.00        Matches:        7
Percent Similarity:     87.5%        Conservative:   0
Best Local Similarity:  87.5%        Mismatches:     1
Query Match:            80.0%        Indels:         0
DB:                     9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-478254 (1-1000)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||||||   |||
Db        469 CCTGATGTCATTTTTCTGCATGAA 492

RESULT 16
US-10-516-032-11
; Sequence 11, Application US/10516032
; GENERAL INFORMATION:
;  APPLICANT: Moodie, Shonna
;  APPLICANT:  Gustafson, Thomas A.
;  TITLE OF INVENTION: METHODS OF TREATING AND DIAGNOSING
;  TITLE OF INVENTION:  DIABETES WITH CX3CR1 MODULATORS
;  FILE REFERENCE: 016325-008010US
;  CURRENT APPLICATION NUMBER: US/10/516,032
;  CURRENT FILING DATE:  2004-11-24
;  PRIOR APPLICATION NUMBER: PCT/US03/18570
;  PRIOR FILING DATE: 2003-06-10
;  PRIOR APPLICATION NUMBER: US 60/387,699
;  PRIOR FILING DATE: 2002-06-10
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 1318
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: putative G-protein coupled receptor (GCR) cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (68)..(1132)
;   OTHER INFORMATION: putative G-protein coupled receptor (GCR)
US-10-516-032-11

Alignment Scores:
Pred. No.:              375          Length:         1318
Score:                  36.00        Matches:        6
Percent Similarity:     87.5%        Conservative:   1
Best Local Similarity:  75.0%        Mismatches:     1
Query Match:            80.0%        Indels:         0
DB:                     6            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-516-032-11 (1-1318)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
```

```
Db        715 CCGGACGCTGTTTTCCTGCAAGAA 738

RESULT 17
US-60-742-219-1191/c
; Sequence 1191, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1191
;   LENGTH: 2508
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-1191

Alignment Scores:
Pred. No.:              783          Length:         2508
Score:                  36.00        Matches:        7
Percent Similarity:     87.5%        Conservative:   0
Best Local Similarity:  87.5%        Mismatches:     1
Query Match:            80.0%        Indels:         0
DB:                     11           Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-742-219-1191 (1-2508)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||||||||||   |||
Db       2040 CCTGATGTCATTTTTCTGCATGAA 2017

RESULT 18
US-10-516-032-9
; Sequence 9, Application US/10516032
; GENERAL INFORMATION:
;  APPLICANT: Moodie, Shonna
;  APPLICANT:  Gustafson, Thomas A.
;  TITLE OF INVENTION: METHODS OF TREATING AND DIAGNOSING
;  TITLE OF INVENTION:  DIABETES WITH CX3CR1 MODULATORS
;  FILE REFERENCE: 016325-008010US
;  CURRENT APPLICATION NUMBER: US/10/516,032
;  CURRENT FILING DATE:  2004-11-24
;  PRIOR APPLICATION NUMBER: PCT/US03/18570
;  PRIOR FILING DATE: 2003-06-10
;  PRIOR APPLICATION NUMBER: US 60/387,699
;  PRIOR FILING DATE: 2002-06-10
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 2813
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   OTHER INFORMATION: chemokine receptor CX3CR1
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (220)..(1284)
;   OTHER INFORMATION: CX3CR1
US-10-516-032-9

Alignment Scores:
Pred. No.:              893          Length:         2813
Score:                  36.00        Matches:        6
Percent Similarity:     87.5%        Conservative:   1
Best Local Similarity:  75.0%        Mismatches:     1
Query Match:            80.0%        Indels:         0
DB:                     6            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-516-032-9 (1-2813)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        867 CCAGACGCTGTTTTCCTGCAAGAA 890

RESULT 19
US-11-293-697-1656
; Sequence 1656, Application US/11293697
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE:  2005-12-05
;  PRIOR APPLICATION NUMBER: US/10/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1656
;   LENGTH: 2854
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-293-697-1656

Alignment Scores:
Pred. No.:              907          Length:         2854
Score:                  36.00        Matches:        6
Percent Similarity:     88.9%        Conservative:   2
Best Local Similarity:  66.7%        Mismatches:     1
Query Match:            80.0%        Indels:         0
DB:                     7            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-293-697-1656 (1-2854)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||   :::
Db       1455 CCCGATGTTGTGTTCCTGCAGCGTCTG 1481

RESULT 20
US-10-516-032-7
; Sequence 7, Application US/10516032
; GENERAL INFORMATION:
;  APPLICANT: Moodie, Shonna
;  APPLICANT:  Gustafson, Thomas A.
;  TITLE OF INVENTION: METHODS OF TREATING AND DIAGNOSING
;  TITLE OF INVENTION:  DIABETES WITH CX3CR1 MODULATORS
;  FILE REFERENCE: 016325-008010US
;  CURRENT APPLICATION NUMBER: US/10/516,032
;  CURRENT FILING DATE:  2004-11-24
;  PRIOR APPLICATION NUMBER: PCT/US03/18570
;  PRIOR FILING DATE: 2003-06-10
;  PRIOR APPLICATION NUMBER: US 60/387,699
;  PRIOR FILING DATE: 2002-06-10
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 3100
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: chemokine (C-X3-C) receptor 1 (CX3CR1), G
;   OTHER INFORMATION: protein-coupled receptor V28 cDNA
;   FEATURE:
;   NAME/KEY: CDS
```

```
;   LOCATION: (88)..(1155)
;   OTHER INFORMATION: CX3CR1
US-10-516-032-7

Alignment Scores:
Pred. No.:                998             Length:         3100
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-516-032-7 (1-3100)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        732 CCAGACGCTGTTTTCCTGCAAGAA 755

RESULT 21
US-11-145-307A-183
; Sequence 183, Application US/11145307A
; GENERAL INFORMATION:
;  APPLICANT: Arcturus Bioscience, Inc.
;  APPLICANT:  Erlander, Mark G.
;  APPLICANT:  Ma, Xiao-Jun
;  TITLE OF INVENTION: Identification of Tumors
;  FILE REFERENCE: 022041-002020US
;  CURRENT APPLICATION NUMBER: US/11/145,307A
;  CURRENT FILING DATE:  2005-06-03
;  PRIOR APPLICATION NUMBER: US 60/577,084
;  PRIOR FILING DATE: 2004-06-04
;  NUMBER OF SEQ ID NOS: 268
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 183
;   LENGTH: 3100
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-145-307A-183

Alignment Scores:
Pred. No.:                998             Length:         3100
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-145-307A-183 (1-3100)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        732 CCAGACGCTGTTTTCCTGCAAGAA 755

RESULT 22
US-11-266-748A-57213
; Sequence 57213, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 57213
;   LENGTH: 3100
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-57213

Alignment Scores:
Pred. No.:                998             Length:         3100
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-57213 (1-3100)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        732 CCAGACGCTGTTTTCCTGCAAGAA 755

RESULT 23
US-10-531-147-2947
; Sequence 2947, Application US/10531147
; GENERAL INFORMATION:
;  APPLICANT: THOMAS WU
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF TUMOR
;  FILE REFERENCE: P5102R1-US
;  CURRENT APPLICATION NUMBER: US/10/531,147
;  CURRENT FILING DATE:  2005-04-12
;  PRIOR APPLICATION NUMBER: PCT/US03/029126
;  PRIOR FILING DATE: 2003-10-15
;  PRIOR APPLICATION NUMBER: US 60/418,988
;  PRIOR FILING DATE: 2002-10-18
;  NUMBER OF SEQ ID NOS: 4622
; SEQ ID NO 2947
;   LENGTH: 3101
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-531-147-2947

Alignment Scores:
Pred. No.:                998             Length:         3101
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-531-147-2947 (1-3101)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        732 CCAGACGCTGTTTTCCTGCAAGAA 755

RESULT 24
US-11-045-004-1
; Sequence 1, Application US/11045004
; GENERAL INFORMATION:
;  APPLICANT: BUCHRIESER, CARMEN
;  APPLICANT:  FRANGEUL, LIONEL
;  APPLICANT:  COUVE, ELISABETH
;  APPLICANT:  RUSNIOK, CHRISTOPHE
```

```
;  APPLICANT:  FSIHI, HAFIDA
;  APPLICANT:  DEHOUX, PIERRE
;  APPLICANT:  DUSSURGET, OLIVIER
;  APPLICANT:  CHETOUANI, FARID
;  APPLICANT:  NEDJARI, HAFED
;  APPLICANT:  GLASER, PHILIPPE
;  APPLICANT:  KUNST, FRANCK
;  APPLICANT:  COSSART, PASCALE
;  APPLICANT:  DANIELS, JUSTIN
;  APPLICANT:  GOEBEI, WERNER
;  APPLICANT:  KREFT, JURGEN
;  APPLICANT:  KUHN, MICHAEL
;  APPLICANT:  NG, EVA
;  APPLICANT:  VAZQUEZ-BOLAND, ANTONIO
;  APPLICANT:  DOMINGUEZ-BERNAL, GUSTAVO
;  APPLICANT:  GARRIDO-GARCIA, PATRICIA
;  APPLICANT:  TIERREZ-MARTINEZ, ALBERTO
;  APPLICANT:  AMEND, ALEXANDRA
;  APPLICANT:  CHAKRABORTY, TRINAD
;  APPLICANT:  DOMANN, EUGEN
;  APPLICANT:  HAIN, THORSTEN
;  APPLICANT:  BERCHE, PATRICK
;  APPLICANT:  CHARBIT, ALAIN
;  APPLICANT:  DURANT, LIONEL
;  APPLICANT:  PEREZ-DIAZ, JOSE-CLAUDIO
;  APPLICANT:  BAQUERO, FERNANDO
;  APPLICANT:  GARCIA DEL PORTILLO, FRANCISCO
;  APPLICANT:  GOMEZ-LOPEZ, NURIA
;  APPLICANT:  MADUENIO, ENCARNA
;  APPLICANT:  PABLOS, BETRIZ DE
;  APPLICANT:  WEHLAND, JURGEN
;  APPLICANT:  KARST, UWE
;  APPLICANT:  ENTIAN, KARL-DIETER
;  APPLICANT:  HAUF, JORG
;  APPLICANT:  ROSE, MATTHIAS
;  APPLICANT:  VOSS, HAMUT
;  TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
;  FILE REFERENCE: 05394.0018-02
;  CURRENT APPLICATION NUMBER: US/11/045,004
;  CURRENT FILING DATE:  2005-01-28
;  PRIOR APPLICATION NUMBER: 10/637,657
;  PRIOR FILING DATE: 2003-08-11
;  PRIOR APPLICATION NUMBER: 10/257,023
;  PRIOR FILING DATE: 2002-10-08
;  PRIOR APPLICATION NUMBER: PCT/FR01/01118
;  PRIOR FILING DATE: 2001-04-11
;  PRIOR APPLICATION NUMBER: FR 00/04,629
;  PRIOR FILING DATE: 2000-04-11
;  NUMBER OF SEQ ID NOS: 2854
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
;   LENGTH: 2944528
;   TYPE: DNA
;   ORGANISM: Listeria monocytogenes
US-11-045-004-1

Alignment Scores:
Pred. No.:              2.25e+06        Length:         2944528
Score:                  36.00           Matches:        7
Percent Similarity:     87.5%           Conservative:   0
Best Local Similarity:  87.5%           Mismatches:     1
Query Match:            80.0%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-045-004-1 (1-2944528)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||||||||||   |||
Db    1010769 CCGGATGTGATTTTTCTTGATGAA 1010792

RESULT 25
US-11-331-019-24422/c
; Sequence 24422, Application US/11331019
; GENERAL INFORMATION:
;  APPLICANT: Fincher, Karen L.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  McCarter, David W.
;  APPLICANT:  Pear, Julie R.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.359 - 38-21(51375)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,019
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/637,086
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: US 60/149,881
;  PRIOR FILING DATE: 1999-08-19
;  NUMBER OF SEQ ID NOS: 52949
; SEQ ID NO 24422
;   LENGTH: 411
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3147-020-Q1-K1-E12
US-11-331-019-24422

Alignment Scores:
Pred. No.:              167             Length:         411
Score:                  35.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   3
Best Local Similarity:  66.7%           Mismatches:     0
Query Match:            77.8%           Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-331-019-24422 (1-411)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::::|||||||||:::|||||
Db        241 CCGAACATCATCTTCCTTAGAGAGGTT 215

RESULT 26
US-11-331-032-4211/c
; Sequence 4211, Application US/11331032
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Xiao, Jinhua
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.361 - 38-21(51462)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,032
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/666,355
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US 60/155,005
;  PRIOR FILING DATE: 1999-09-21
;  NUMBER OF SEQ ID NOS: 25978
; SEQ ID NO 4211
;   LENGTH: 444
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: uC-gsronu33B146f06b1
US-11-331-032-4211

Alignment Scores:
Pred. No.:              182             Length:         444
Score:                  35.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   3
Best Local Similarity:  66.7%           Mismatches:     0
Query Match:            77.8%           Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-331-032-4211 (1-444)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
```

```
                  |||:::::::|||||||||:::||||||
Db        193 CCGAACATCATCTTCCTTAGAGAGGTT 167

RESULT 27
US-11-292-078-15903/c
; Sequence 15903, Application US/11292078
; GENERAL INFORMATION:
;  APPLICANT: Deikman, Jill
;  APPLICANT:  Feng, Paul C.C.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Ziegler, Todd E.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.346 - 38-21(52274)C
;  CURRENT APPLICATION NUMBER: US/11/292,078
;  CURRENT FILING DATE:  2005-12-02
;  PRIOR APPLICATION NUMBER: US 10/021,323
;  PRIOR FILING DATE: 2001-12-12
;  PRIOR APPLICATION NUMBER: US 60/255,619
;  PRIOR FILING DATE: 2000-12-14
;  NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 15903
;   LENGTH: 578
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3829-032-Q6-K6-E9
US-11-292-078-15903

Alignment Scores:
Pred. No.:                247             Length:         578
Score:                    35.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   3
Best Local Similarity:    66.7%           Mismatches:     0
Query Match:              77.8%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-292-078-15903 (1-578)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
                |||:::::::|||||||||:::|||||||
Db        529 CCGAACATCATCTTCCTTAGAGAGGTT 503

RESULT 28
US-11-266-748A-408169
; Sequence 408169, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 408169
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-408169

Alignment Scores:
Pred. No.:                462             Length:         1000
Score:                    35.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
Best Local Similarity:    66.7%           Mismatches:     1
Query Match:              77.8%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-408169 (1-1000)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
                |||||||||:::   ||||||:::|||
Db        184 CCTGACGTGGTGCTGCTCCAGGATGTG 210

RESULT 29
US-11-266-748A-479215/c
; Sequence 479215, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 479215
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-479215

Alignment Scores:
Pred. No.:                462             Length:         1000
Score:                    35.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
Best Local Similarity:    66.7%           Mismatches:     1
Query Match:              77.8%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-479215 (1-1000)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
                |||||||||:::   ||||||:::|||
Db        817 CCTGACGTGGTGCTGCTCCAGGATGTG 791

RESULT 30
US-11-266-748A-25946
; Sequence 25946, Application US/11266748A
```

```
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25946
;   LENGTH: 1239
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-25946

Alignment Scores:
Pred. No.:                590          Length:          1239
Score:                    35.00        Matches:         7
Percent Similarity:       88.9%        Conservative:    1
Best Local Similarity:    77.8%        Mismatches:      1
Query Match:              77.8%        Indels:          0
DB:                       9            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-25946 (1-1239)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||   |||||||||:::
Db        408 CCCGATGTCATCAAGTTGCAAGAGATT 434

RESULT 31
US-11-342-366-1917
; Sequence 1917, Application US/11342366
; GENERAL INFORMATION:
;  APPLICANT: Hood, Leroy
;  APPLICANT:  Lin, Biaoyang
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING AND USING
;  TITLE OF INVENTION:  ORGAN-SPECIFIC PROTEINS IN BLOOD
;  FILE REFERENCE: 460092.401
;  CURRENT APPLICATION NUMBER: US/11/342,366
;  CURRENT FILING DATE:  2006-01-27
;  PRIOR APPLICATION NUMBER: 60/683,071
;  PRIOR FILING DATE: 2005-05-20
;  PRIOR APPLICATION NUMBER: 60/647,685
;  PRIOR FILING DATE: 2005-01-27
;  NUMBER OF SEQ ID NOS: 1947
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1917
;   LENGTH: 1963
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-342-366-1917

Alignment Scores:
Pred. No.:                999          Length:          1963
Score:                    35.00        Matches:         7
Percent Similarity:       87.5%        Conservative:    0
Best Local Similarity:    87.5%        Mismatches:      1
Query Match:              77.8%        Indels:          0
DB:                       7            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-342-366-1917 (1-1963)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||||||||||   |||
Db        982 CCAGATGTGATCTTCTTGGCAGAG 1005

RESULT 32
US-11-342-367-1917
; Sequence 1917, Application US/11342367
; GENERAL INFORMATION:
;  APPLICANT: Hood, Leroy
;  APPLICANT:  Lin, Biaoyang
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING AND
;  TITLE OF INVENTION:  MONITORING DRUG SIDE EFFECTS
;  FILE REFERENCE: 460092.402
;  CURRENT APPLICATION NUMBER: US/11/342,367
;  CURRENT FILING DATE:  2006-01-27
;  PRIOR APPLICATION NUMBER: 60/683,004
;  PRIOR FILING DATE: 2005-05-20
;  PRIOR APPLICATION NUMBER: 60/647,792
;  PRIOR FILING DATE: 2005-01-27
;  NUMBER OF SEQ ID NOS: 1947
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1917
;   LENGTH: 1963
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-342-367-1917

Alignment Scores:
Pred. No.:                999          Length:          1963
Score:                    35.00        Matches:         7
Percent Similarity:       87.5%        Conservative:    0
Best Local Similarity:    87.5%        Mismatches:      1
Query Match:              77.8%        Indels:          0
DB:                       7            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-342-367-1917 (1-1963)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||||||||||   |||
Db        982 CCAGATGTGATCTTCTTGGCAGAG 1005

RESULT 33
US-10-461-673-568
; Sequence 568, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
```

; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 568
; LENGTH: 1988
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(1277)
US-10-461-673-568

Alignment Scores:

Pred. No.:	1.01e+03	Length:	1988
Score:	35.00	Matches:	7
Percent Similarity:	87.5%	Conservative:	0
Best Local Similarity:	87.5%	Mismatches:	1
Query Match:	77.8%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-461-673-568 (1-1988)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||||||   |||
Db       1029 CCAGATGTGATCTTCTTGGCAGAG 1052
```

RESULT 34
US-11-266-748A-32741
; Sequence 32741, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32741
; LENGTH: 1995
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-32741

Alignment Scores:

Pred. No.:	1.02e+03	Length:	1995
Score:	35.00	Matches:	7
Percent Similarity:	87.5%	Conservative:	0
Best Local Similarity:	87.5%	Mismatches:	1
Query Match:	77.8%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-32741 (1-1995)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||||||   |||
Db       1029 CCAGATGTGATCTTCTTGGCAGAG 1052
```

RESULT 35
US-60-751-420-1685
; Sequence 1685, Application US/60751420
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John Verner
; APPLICANT: Bradley, Walter Edward
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Debrus, Sophie
; APPLICANT: Van Eerdewegh, Paul
; APPLICANT: Little, Randall David
; APPLICANT: Keith, Tim
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: GeneMap of the human genes associated with Crohn's disease
; FILE REFERENCE: GENI-011/00US
; CURRENT APPLICATION NUMBER: US/60/751,420
; CURRENT FILING DATE: 2005-12-19
; NUMBER OF SEQ ID NOS: 27266
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1685
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Homosapiens
US-60-751-420-1685

Alignment Scores:

Pred. No.:	1.12e+03	Length:	2163
Score:	35.00	Matches:	7
Percent Similarity:	87.5%	Conservative:	0
Best Local Similarity:	87.5%	Mismatches:	1
Query Match:	77.8%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-60-751-420-1685 (1-2163)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||||||   |||
Db       1028 CCAGATGTGATCTTCTTGGCAGAG 1051
```

RESULT 36
US-11-314-006-3974

```
; Sequence 3974, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 3974
;   LENGTH: 226
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-047-Q1-E1-C9
US-11-314-006-3974

Alignment Scores:
Pred. No.:                142           Length:         226
Score:                    34.00         Matches:        6
Percent Similarity:       100.0%        Conservative:   1
Best Local Similarity:    85.7%         Mismatches:     0
Query Match:              75.6%         Indels:         0
DB:                       9             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-3974 (1-226)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         79 CCGGACGTGCTGTTCCTGCAA 99

RESULT 37
US-11-314-006-1805
; Sequence 1805, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 1805
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-045-Q1-E1-C2
US-11-314-006-1805

Alignment Scores:
Pred. No.:                144           Length:         228
Score:                    34.00         Matches:        6
Percent Similarity:       100.0%        Conservative:   1
Best Local Similarity:    85.7%         Mismatches:     0
Query Match:              75.6%         Indels:         0
DB:                       9             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-1805 (1-228)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         92 CCGGACGTGCTGTTCCTGCAA 112

RESULT 38
US-11-314-006-8
; Sequence 8, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 8
;   LENGTH: 334
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-001-Q1-E1-B3
US-11-314-006-8

Alignment Scores:
Pred. No.:                222           Length:         334
Score:                    34.00         Matches:        6
Percent Similarity:       100.0%        Conservative:   1
Best Local Similarity:    85.7%         Mismatches:     0
Query Match:              75.6%         Indels:         0
DB:                       9             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-8 (1-334)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db        107 CCGGACGTGCTGTTCCTGCAA 127

RESULT 39
US-11-353-150-18282
; Sequence 18282, Application US/11353150
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Scott E.
;  APPLICANT:  Conner, Timothy W.
;  APPLICANT:  Lalgudi, Raghunath V.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.365 - 38-21(15459)D/US
;  CURRENT APPLICATION NUMBER: US/11/353,150
;  CURRENT FILING DATE:  2006-02-14
;  PRIOR APPLICATION NUMBER: 09/565,386
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/132,860
;  PRIOR FILING DATE: 1999-05-07
;  NUMBER OF SEQ ID NOS: 83523
; SEQ ID NO 18282
;   LENGTH: 371
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB189-006-Q1-E1-E7
US-11-353-150-18282

Alignment Scores:
Pred. No.:                251           Length:         371
Score:                    34.00         Matches:        6
Percent Similarity:       100.0%        Conservative:   1
Best Local Similarity:    85.7%         Mismatches:     0
Query Match:              75.6%         Indels:         0
DB:                       7             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-353-150-18282 (1-371)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db        162 CCTGATGTGCTGTTCTTGCAA 182

RESULT 40
```

```
US-11-314-006-3002
; Sequence 3002, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 3002
;   LENGTH: 371
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-021-Q1-E1-A7
US-11-314-006-3002

Alignment Scores:
Pred. No.:                251             Length:         371
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-3002 (1-371)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         78 CCGGACGTGCTGTTCCTGCAA 98

RESULT 41
US-11-314-006-626
; Sequence 626, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 626
;   LENGTH: 374
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-014-Q1-E1-H7
US-11-314-006-626

Alignment Scores:
Pred. No.:                253             Length:         374
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-626 (1-374)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         84 CCGGACGTGCTGTTCCTGCAA 104

RESULT 42
US-11-314-006-598
; Sequence 598, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 598
;   LENGTH: 375
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-014-Q1-E1-C7
US-11-314-006-598

Alignment Scores:
Pred. No.:                254             Length:         375
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-598 (1-375)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         84 CCGGACGTGCTGTTCCTGCAA 104

RESULT 43
US-11-314-006-2239
; Sequence 2239, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 2239
;   LENGTH: 377
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-057-Q1-E1-H3
US-11-314-006-2239

Alignment Scores:
Pred. No.:                255             Length:         377
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-2239 (1-377)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         78 CCGGACGTGCTGTTCCTGCAA 98

RESULT 44
```

US-11-314-006-1145
; Sequence 1145, Application US/11314006
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
; TITLE OF INVENTION: plants
; FILE REFERENCE: 16517.351 - 38-21(15726)C/US
; CURRENT APPLICATION NUMBER: US/11/314,006
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US 09/540,234
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 1145
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB103-026-Q1-E1-B10
US-11-314-006-1145

Alignment Scores:

Pred. No.:	258	Length:	380
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-1145 (1-380)

```
Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         99 CCGGACGTGCTGTTCCTGCAA 119
```

RESULT 45
US-11-314-006-2650
; Sequence 2650, Application US/11314006
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
; TITLE OF INVENTION: plants
; FILE REFERENCE: 16517.351 - 38-21(15726)C/US
; CURRENT APPLICATION NUMBER: US/11/314,006
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US 09/540,234
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 2650
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB103-008-Q1-E1-B7
US-11-314-006-2650

Alignment Scores:

Pred. No.:	259	Length:	382
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-2650 (1-382)

```
Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         73 CCGGACGTGCTGTTCCTGCAA 93
```

RESULT 46
US-11-314-006-3428
; Sequence 3428, Application US/11314006
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
; TITLE OF INVENTION: plants
; FILE REFERENCE: 16517.351 - 38-21(15726)C/US
; CURRENT APPLICATION NUMBER: US/11/314,006
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US 09/540,234
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 3428
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB103-033-Q1-E1-E2
US-11-314-006-3428

Alignment Scores:

Pred. No.:	261	Length:	384
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-3428 (1-384)

```
Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         78 CCGGACGTGCTGTTCCTGCAA 98
```

RESULT 47
US-11-314-006-2731
; Sequence 2731, Application US/11314006
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
; TITLE OF INVENTION: plants
; FILE REFERENCE: 16517.351 - 38-21(15726)C/US
; CURRENT APPLICATION NUMBER: US/11/314,006
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US 09/540,234
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 2731
; LENGTH: 387
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB103-010-Q1-E1-H2
US-11-314-006-2731

Alignment Scores:

Pred. No.:	263	Length:	387
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-2731 (1-387)

```
Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         78 CCGGACGTGCTGTTCCTGCAA 98
```

RESULT 48
US-11-331-019-34748/c
; Sequence 34748, Application US/11331019

```
; GENERAL INFORMATION:
;  APPLICANT: Fincher, Karen L.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  McCarter, David W.
;  APPLICANT:  Pear, Julie R.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.359 - 38-21(51375)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,019
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/637,086
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: US 60/149,881
;  PRIOR FILING DATE: 1999-08-19
;  NUMBER OF SEQ ID NOS: 52949
; SEQ ID NO 34748
;   LENGTH: 388
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3149-040-Q1-K1-G1
US-11-331-019-34748

Alignment Scores:
Pred. No.:                264             Length:         388
Score:                    34.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
Best Local Similarity:    66.7%           Mismatches:     1
Query Match:              75.6%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-331-019-34748 (1-388)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||   |||:::
Db        302 CCGGACATCATCTTCCTTTGAGAGATT 276

RESULT 49
US-11-314-006-45
; Sequence 45, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 45
;   LENGTH: 388
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-001-Q1-E2-A11
US-11-314-006-45

Alignment Scores:
Pred. No.:                264             Length:         388
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-45 (1-388)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         84 CCGGACGTGCTGTTCCTGCAA 104

RESULT 50
US-11-314-006-1103
; Sequence 1103, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 1103
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-025-Q1-E1-A4
US-11-314-006-1103

Alignment Scores:
Pred. No.:                270             Length:         396
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-1103 (1-396)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         92 CCGGACGTGCTGTTCCTGCAA 112

RESULT 51
US-11-314-006-2849
; Sequence 2849, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 2849
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-015-Q1-E1-F2
US-11-314-006-2849

Alignment Scores:
Pred. No.:                273             Length:         400
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-2849 (1-400)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         78 CCGGACGTGCTGTTCCTGCAA 98

RESULT 52
```

```
US-11-314-006-3724
; Sequence 3724, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 3724
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-040-Q1-E1-A2
US-11-314-006-3724

Alignment Scores:
Pred. No.:              274          Length:         401
Score:                  34.00        Matches:        6
Percent Similarity:     100.0%       Conservative:   1
Best Local Similarity:  85.7%        Mismatches:     0
Query Match:            75.6%        Indels:         0
DB:                     9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-3724 (1-401)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         78 CCGGACGTGCTGTTCCTGCAA 98

RESULT 53
US-11-314-006-3916
; Sequence 3916, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 3916
;   LENGTH: 403
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-045-Q1-E1-F2
US-11-314-006-3916

Alignment Scores:
Pred. No.:              276          Length:         403
Score:                  34.00        Matches:        6
Percent Similarity:     100.0%       Conservative:   1
Best Local Similarity:  85.7%        Mismatches:     0
Query Match:            75.6%        Indels:         0
DB:                     9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-3916 (1-403)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         80 CCGGACGTGCTGTTCCTGCAA 100

RESULT 54
US-11-314-006-54
; Sequence 54, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 54
;   LENGTH: 404
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-001-Q1-E2-B3
US-11-314-006-54

Alignment Scores:
Pred. No.:              276          Length:         404
Score:                  34.00        Matches:        6
Percent Similarity:     100.0%       Conservative:   1
Best Local Similarity:  85.7%        Mismatches:     0
Query Match:            75.6%        Indels:         0
DB:                     9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-54 (1-404)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db        106 CCGGACGTGCTGTTCCTGCAA 126

RESULT 55
US-11-314-006-1295
; Sequence 1295, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 1295
;   LENGTH: 408
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-030-Q1-E1-A6
US-11-314-006-1295

Alignment Scores:
Pred. No.:              280          Length:         408
Score:                  34.00        Matches:        6
Percent Similarity:     100.0%       Conservative:   1
Best Local Similarity:  85.7%        Mismatches:     0
Query Match:            75.6%        Indels:         0
DB:                     9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-1295 (1-408)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         85 CCGGACGTGCTGTTCCTGCAA 105

RESULT 56
```

```
US-11-314-006-2829
; Sequence 2829, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 2829
;   LENGTH: 409
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-015-Q1-E1-B12
US-11-314-006-2829

Alignment Scores:
Pred. No.:                280             Length:         409
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-2829 (1-409)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         83 CCGGACGTGCTGTTCCTGCAA 103

RESULT 57
US-11-314-006-2227
; Sequence 2227, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 2227
;   LENGTH: 421
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-057-Q1-E1-F7
US-11-314-006-2227

Alignment Scores:
Pred. No.:                290             Length:         421
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-2227 (1-421)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         97 CCGGACGTGCTGTTCCTGCAA 117

RESULT 58
US-11-314-006-3827
; Sequence 3827, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 3827
;   LENGTH: 422
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-043-Q1-E1-D11
US-11-314-006-3827

Alignment Scores:
Pred. No.:                291             Length:         422
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-3827 (1-422)

Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db        173 CCGGACGTGCTGTTCCTGCAA 193

RESULT 59
US-10-998-197A-31
; Sequence 31, Application US/10998197A
; GENERAL INFORMATION:
;  APPLICANT: Dohrmann, Cord
;  APPLICANT:  Austen, Matthias
;  TITLE OF INVENTION: Pancreas-Specific Proteins
;  FILE REFERENCE: 2923-669
;  CURRENT APPLICATION NUMBER: US/10/998,197A
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: PCT/EP03/05700
;  PRIOR FILING DATE: 2003-05-30
;  PRIOR APPLICATION NUMBER: EP 02020829.4
;  PRIOR FILING DATE: 2002-09-17
;  PRIOR APPLICATION NUMBER: EP 02011963.2
;  PRIOR FILING DATE: 2002-05-29
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
;   LENGTH: 423
;   TYPE: DNA
;   ORGANISM: Gallus gallus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (265)..(265)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-998-197A-31

Alignment Scores:
Pred. No.:                291             Length:         423
Score:                    34.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
Best Local Similarity:    66.7%           Mismatches:     1
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-998-197A-31 (1-423)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
```

```
                ||||||::::::  ||||||||||||
Db        377 CCGGATATCTTGTGCCTGCAAGAAGTC 403
```

RESULT 60
US-11-314-006-1619
; Sequence 1619, Application US/11314006
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
; TITLE OF INVENTION: plants
; FILE REFERENCE: 16517.351 - 38-21(15726)C/US
; CURRENT APPLICATION NUMBER: US/11/314,006
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US 09/540,234
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 1619
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB103-041-Q1-E1-C1
US-11-314-006-1619

Alignment Scores:

Pred. No.:	291	Length:	423
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-1619 (1-423)

```
Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         93 CCGGACGTGCTGTTCCTGCAA 113
```

RESULT 61
US-11-314-006-4009
; Sequence 4009, Application US/11314006
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
; TITLE OF INVENTION: plants
; FILE REFERENCE: 16517.351 - 38-21(15726)C/US
; CURRENT APPLICATION NUMBER: US/11/314,006
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US 09/540,234
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 4009
; LENGTH: 431
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB103-048-Q1-E2-B7
US-11-314-006-4009

Alignment Scores:

Pred. No.:	298	Length:	431
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-4009 (1-431)

```
Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         81 CCGGACGTGCTGTTCCTGCAA 101
```

RESULT 62
US-11-314-006-2315
; Sequence 2315, Application US/11314006
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
; TITLE OF INVENTION: plants
; FILE REFERENCE: 16517.351 - 38-21(15726)C/US
; CURRENT APPLICATION NUMBER: US/11/314,006
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US 09/540,234
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 2315
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB103-061-Q1-E1-A12
US-11-314-006-2315

Alignment Scores:

Pred. No.:	300	Length:	434
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-2315 (1-434)

```
Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db         85 CCGGACGTGCTGTTCCTGCAA 105
```

RESULT 63
US-11-314-006-3781
; Sequence 3781, Application US/11314006
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
; TITLE OF INVENTION: plants
; FILE REFERENCE: 16517.351 - 38-21(15726)C/US
; CURRENT APPLICATION NUMBER: US/11/314,006
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US 09/540,234
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 3781
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB103-041-Q1-E1-G2
US-11-314-006-3781

Alignment Scores:

Pred. No.:	302	Length:	437
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-3781 (1-437)

```
Qy          1 ProAspValIlePheLeuGln 7
```

```
                  |||||||||:::|||||||||
Db         78 CCGGACGTGCTGTTCCTGCAA 98

RESULT 64
US-11-266-748A-270053
; Sequence 270053, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 270053
;   LENGTH: 437
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (418)..(418)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-270053

Alignment Scores:
Pred. No.:                302             Length:         437
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-270053 (1-437)

Qy           1 ProAspValIlePheLeuGln 7
                 |||||||||:::|||||||||
Db         225 CCAGATGTTCTTTTTCTTCAA 245

RESULT 65
US-11-266-748A-330570/c
; Sequence 330570, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 330570
;   LENGTH: 437
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (20)..(20)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-330570

Alignment Scores:
Pred. No.:                302             Length:         437
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-330570 (1-437)

Qy           1 ProAspValIlePheLeuGln 7
                 |||||||||:::|||||||||
Db         213 CCAGATGTTCTTTTTCTTCAA 193

RESULT 66
US-11-314-006-2378
; Sequence 2378, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 2378
;   LENGTH: 442
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-001-Q1-E1-A11
US-11-314-006-2378

Alignment Scores:
Pred. No.:                306             Length:         442
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-2378 (1-442)

Qy           1 ProAspValIlePheLeuGln 7
```

```
                  |||||||||:::|||||||||
Db         81 CCGGACGTGCTGTTCCTGCAA 101

RESULT 67
US-11-314-006-910
; Sequence 910, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 910
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-021-Q1-E1-A8
US-11-314-006-910

Alignment Scores:
Pred. No.:              312          Length:         449
Score:                  34.00        Matches:        6
Percent Similarity:     100.0%       Conservative:   1
Best Local Similarity:  85.7%        Mismatches:     0
Query Match:            75.6%        Indels:         0
DB:                     9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-910 (1-449)

Qy          1 ProAspValIlePheLeuGln 7
                |||||||||:::|||||||||
Db         97 CCGGACGTGCTGTTCCTGCAA 117

RESULT 68
US-11-314-006-915
; Sequence 915, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 915
;   LENGTH: 454
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-021-Q1-E1-B5
US-11-314-006-915

Alignment Scores:
Pred. No.:              316          Length:         454
Score:                  34.00        Matches:        6
Percent Similarity:     100.0%       Conservative:   1
Best Local Similarity:  85.7%        Mismatches:     0
Query Match:            75.6%        Indels:         0
DB:                     9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-915 (1-454)

Qy          1 ProAspValIlePheLeuGln 7
                |||||||||:::|||||||||
Db         96 CCGGACGTGCTGTTCCTGCAA 116

RESULT 69
US-11-314-006-2262
; Sequence 2262, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 2262
;   LENGTH: 454
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-059-Q1-E1-A6
US-11-314-006-2262

Alignment Scores:
Pred. No.:              316          Length:         454
Score:                  34.00        Matches:        6
Percent Similarity:     100.0%       Conservative:   1
Best Local Similarity:  85.7%        Mismatches:     0
Query Match:            75.6%        Indels:         0
DB:                     9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-2262 (1-454)

Qy          1 ProAspValIlePheLeuGln 7
                |||||||||:::|||||||||
Db         97 CCGGACGTGCTGTTCCTGCAA 117

RESULT 70
US-11-314-006-317
; Sequence 317, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 317
;   LENGTH: 456
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-007-Q1-E1-D5
US-11-314-006-317

Alignment Scores:
Pred. No.:              318          Length:         456
Score:                  34.00        Matches:        6
Percent Similarity:     100.0%       Conservative:   1
Best Local Similarity:  85.7%        Mismatches:     0
Query Match:            75.6%        Indels:         0
DB:                     9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-317 (1-456)

Qy          1 ProAspValIlePheLeuGln 7
```

```
                  |||||||||:::|||||||||
Db          96 CCGGACGTGCTGTTCCTGCAA 116

RESULT 71
US-11-314-006-951
; Sequence 951, Application US/11314006
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with
;  TITLE OF INVENTION:  plants
;  FILE REFERENCE: 16517.351 - 38-21(15726)C/US
;  CURRENT APPLICATION NUMBER: US/11/314,006
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: US 09/540,234
;  PRIOR FILING DATE: 2000-04-03
;  NUMBER OF SEQ ID NOS: 11840
; SEQ ID NO 951
;   LENGTH: 473
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB103-021-Q1-E1-G9
US-11-314-006-951

Alignment Scores:
Pred. No.:                331             Length:         473
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-314-006-951 (1-473)

Qy           1 ProAspValIlePheLeuGln 7
                |||||||||:::|||||||||
Db          94 CCGGACGTGCTGTTCCTGCAA 114

RESULT 72
US-11-313-816-17804
; Sequence 17804, Application US/11313816
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Ruan, Yijun G.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.353 - 38-21(15404)C/US
;  CURRENT APPLICATION NUMBER: US/11/313,816
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: 09/333,534
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 18718
; SEQ ID NO 17804
;   LENGTH: 486
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB35-056-Q1-E2-C9
US-11-313-816-17804

Alignment Scores:
Pred. No.:                342             Length:         486
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-313-816-17804 (1-486)

Qy           1 ProAspValIlePheLeuGln 7
                |||||||||:::|||||||||
Db         213 CCGGACGTGCTGTTCCTGCAA 233

RESULT 73
US-11-266-748A-100449/c
; Sequence 100449, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 100449
;   LENGTH: 536
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-100449

Alignment Scores:
Pred. No.:                382             Length:         536
Score:                    34.00           Matches:        7
Percent Similarity:       77.8%           Conservative:   0
Best Local Similarity:    77.8%           Mismatches:     2
Query Match:              75.6%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-100449 (1-536)

Qy           1 ProAspValIlePheLeuGlnGluVal 9
                ||||||   |||||||||   ||||||
Db         398 CCAGACAAAATATTCCTCATGGAAGTG 372

RESULT 74
US-11-266-748A-153260
; Sequence 153260, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
```

```
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 153260
;   LENGTH: 536
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-153260

Alignment Scores:
Pred. No.:              382          Length:          536
Score:                  34.00        Matches:         7
Percent Similarity:     77.8%        Conservative:    0
Best Local Similarity:  77.8%        Mismatches:      2
Query Match:            75.6%        Indels:          0
DB:                     9            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-153260 (1-536)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||   |||||||||   ||||||
Db        139 CCAGACAAAATATTCCTCATGGAAGTG 165

RESULT 75
US-11-266-748A-55838/c
; Sequence 55838, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 55838
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-55838

Alignment Scores:
Pred. No.:              408          Length:          568
Score:                  34.00        Matches:         7
Percent Similarity:     87.5%        Conservative:    0
Best Local Similarity:  87.5%        Mismatches:      1
Query Match:            75.6%        Indels:          0
DB:                     9            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-266-748A-55838 (1-568)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||   |||||||||
Db        215 CCTGATGTAATACACCTTCAAGAG 192


Search completed: March 10, 2006, 23:59:49
Job time : 50.0571 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2006, 14:19:05 ; Search time 8103 Seconds
(without alignments)
8952.555 Million cell updates/sec

Title: US-10-757-745-3
Perfect score: 1312
Sequence: 1 agctattaatgattcgaatt.....tctcaacctttcaggacatc 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUSB_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUSC_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US075_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US076_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US077_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US078_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US079_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US101_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US104_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US105_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US112_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US117_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US600_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US601_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US603_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US605_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US606_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US607_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1312	100.0	1312	61	US-10-757-745-3	Sequence 3, Appli
2	1238	94.4	1944	42	US-10-144-771-826	Sequence 826, App
3	1238	94.4	1944	78	US-60-360-207-826	Sequence 826, App
c 4	682.8	52.0	774	51	US-10-488-619-2153	Sequence 2153, Ap
5	669	51.0	1919	79	US-60-443-566-1795	Sequence 1795, Ap
6	669	51.0	1919	79	US-60-452-680-10426	Sequence 10426, A
7	669	51.0	1919	79	US-60-455-444-3324	Sequence 3324, Ap
8	669	51.0	1919	79	US-60-465-241-3324	Sequence 3324, Ap
9	668.2	50.9	1920	61	US-10-757-745-1	Sequence 1, Appli
10	668.2	50.9	1948	24	US-09-488-725A-1200	Sequence 1200, Ap
11	668.2	50.9	1948	40	US-10-037-270-889	Sequence 889, App
12	668.2	50.9	1948	42	US-10-105-837-889	Sequence 889, App
13	668.2	50.9	1948	42	US-10-117-722-889	Sequence 889, App
14	668.2	50.9	1948	42	US-10-122-851-889	Sequence 889, App
15	668.2	50.9	1948	43	US-10-258-898A-1200	Sequence 1200, Ap
16	668.2	50.9	1948	43	US-10-286-897-1200	Sequence 1200, Ap
17	668.2	50.9	1958	24	US-09-471-275-1314	Sequence 1314, Ap
18	668.2	50.9	1958	24	US-09-488-725A-4772	Sequence 4772, Ap
19	668.2	50.9	1958	24	US-09-488-725B-8295	Sequence 8295, Ap
20	668.2	50.9	1958	27	US-09-552-317-8295	Sequence 8295, Ap
21	668.2	50.9	1958	43	US-10-258-898A-4772	Sequence 4772, Ap
22	668.2	50.9	1958	43	US-10-286-897-4772	Sequence 4772, Ap
23	668.2	50.9	1958	48	US-10-302-689A-59957	Sequence 59957, A

24	668.2	50.9	1958	48	US-10-302-689A-77389	Sequence 77389, A
25	668.2	50.9	2262	28	US-09-606-680-2868	Sequence 2868, Ap
26	668.2	50.9	2262	28	US-09-606-776-4085	Sequence 4085, Ap
27	668.2	50.9	2262	29	US-09-644-869-9498	Sequence 9498, Ap
28	668.2	50.9	2262	29	US-09-652-121-6794	Sequence 6794, Ap
29	668.2	50.9	2262	29	US-09-652-122-3756	Sequence 3756, Ap
30	668.2	50.9	2262	29	US-09-652-128-9329	Sequence 9329, Ap
31	668.2	50.9	2262	29	US-09-652-816-9572	Sequence 9572, Ap
32	668.2	50.9	2262	31	US-09-700-000-6419	Sequence 6419, Ap
33	668.2	50.9	2262	31	US-09-710-280-2045	Sequence 2045, Ap
34	668.2	50.9	2262	31	US-09-710-282-3926	Sequence 3926, Ap
35	668.2	50.9	2262	31	US-09-726-802-2141	Sequence 2141, Ap
36	668.2	50.9	3152	38	US-09-981-353-88	Sequence 88, Appl
37	668.2	50.9	3152	77	US-60-239-841-88	Sequence 88, Appl
38	667	50.8	1296	1	PCT-US00-05883-170	Sequence 170, App
39	667	50.8	1296	36	US-09-925-299-170	Sequence 170, App
40	666.6	50.8	1919	42	US-10-170-235-27345	Sequence 27345, A
41	666.6	50.8	1940	3	PCT-US04-24424-921	Sequence 921, App
42	666.6	50.8	1940	61	US-10-783-271-28	Sequence 28, Appl
43	666.6	50.8	1940	80	US-60-490-890-921	Sequence 921, App
44	666.6	50.8	1940	81	US-60-568-073-582	Sequence 582, App
45	666.6	50.8	1940	82	US-60-671-501-156	Sequence 156, App
46	666.6	50.8	1940	82	US-60-685-372-523	Sequence 523, App
47	666.6	50.8	1940	83	US-60-704-740-28	Sequence 28, Appl
48	666.6	50.8	1940	83	US-60-710-726-741	Sequence 741, App
49	665.6	50.7	3080	27	US-09-585-944-2184	Sequence 2184, Ap
c 50	665.6	50.7	3082	75	US-60-172-360-956	Sequence 956, App
51	665	50.7	1898	28	US-09-629-469A-13209	Sequence 13209, A
52	665	50.7	1898	63	US-10-917-503-13209	Sequence 13209, A
53	658.8	50.2	1930	31	US-09-705-256A-7708	Sequence 7708, Ap
54	658.8	50.2	1930	75	US-60-164-285-7708	Sequence 7708, Ap
55	600.4	45.8	602	51	US-10-488-619-2154	Sequence 2154, Ap
c 56	575.6	43.9	1849	79	US-60-443-566-1789	Sequence 1789, Ap
c 57	575.6	43.9	1849	79	US-60-452-680-10422	Sequence 10422, A
c 58	575.6	43.9	1849	79	US-60-455-444-3323	Sequence 3323, Ap
c 59	575.6	43.9	1849	79	US-60-465-241-3323	Sequence 3323, Ap
c 60	573.2	43.7	1849	42	US-10-170-235-26834	Sequence 26834, A
61	493.2	37.6	1088	1	PCT-US01-08631-22058	Sequence 22058, A
62	493.2	37.6	1088	48	US-10-302-689A-136240	Sequence 136240,
63	493.2	37.6	1088	51	US-10-450-763-22058	Sequence 22058, A
c 64	490	37.3	1079	22	US-09-248-178-63	Sequence 63, Appl
c 65	490	37.3	1079	22	US-09-288-950-63	Sequence 63, Appl
c 66	490	37.3	1079	23	US-09-346-327-63	Sequence 63, Appl
c 67	490	37.3	1079	29	US-09-687-507-63	Sequence 63, Appl
c 68	490	37.3	1079	29	US-09-687-507A-63	Sequence 63, Appl
c 69	490	37.3	1079	31	US-09-745-288-63	Sequence 63, Appl
c 70	490	37.3	1079	32	US-09-778-381-63	Sequence 63, Appl
c 71	490	37.3	1079	51	US-10-453-919-63	Sequence 63, Appl
72	473.6	36.1	1852	29	US-09-644-868-7772	Sequence 7772, Ap
73	473.6	36.1	1852	29	US-09-649-161-8751	Sequence 8751, Ap
74	473.6	36.1	1852	29	US-09-652-109-8165	Sequence 8165, Ap
75	473.6	36.1	1852	29	US-09-652-128-7441	Sequence 7441, Ap

ALIGNMENTS

```
RESULT 1
US-10-757-745-3
; Sequence 3, Application US/10757745
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (122)..(1234)
;   OTHER INFORMATION:
US-10-757-745-3

  Query Match           100.0%;  Score 1312;  DB 61;  Length 1312;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1312;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60

Qy         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120

Qy        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180

Qy        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240

Qy        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300

Qy        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360

Qy        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420

Qy        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480

Qy        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540

Qy        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600

Qy        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660

Qy        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720

Qy        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780

Qy        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840

Qy        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900

Qy        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960

Qy        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020

Qy       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080

Qy       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140

Qy       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200

Qy       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260

Qy       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
```

```
RESULT 2
US-10-144-771-826
; Sequence 826, Application US/10144771
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig
;  TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
;  FILE REFERENCE: CL001321
;  CURRENT APPLICATION NUMBER: US/10/144,771
;  CURRENT FILING DATE:  2002-05-15
;  NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 826
;   LENGTH: 1944
;   TYPE: DNA
;   ORGANISM: HUMAN
US-10-144-771-826

  Query Match            94.4%;  Score 1238;  DB 42;  Length 1944;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1238;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          75 GCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCA 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCA 60

Qy         135 GCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAG 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAG 120

Qy         195 CAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGT 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGT 180

Qy         255 TGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGA 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGA 240

Qy         315 CGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCC 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCC 300

Qy         375 AGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATG 434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATG 360

Qy         435 ATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTT 494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 ATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTT 420

Qy         495 CTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAG 554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAG 480

Qy         555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 540

Qy         615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 600

Qy         675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 660

Qy         735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 720

Qy         795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 780

Qy         855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 840

Qy         915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 900

Qy         975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 960

Qy        1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1020

Qy        1095 TTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTG 1154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 TTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTG 1080

Qy        1155 TTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCA 1214
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 TTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCA 1140

Qy        1215 CCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACTA 1274
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACTA 1200

Qy        1275 GTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
               |||||||||||||||||||||||||||||||||||||
Db        1201 GTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1238
```

```
RESULT 3
US-60-360-207-826
; Sequence 826, Application US/60360207
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig
;  TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
;  FILE REFERENCE: CL001321
;  CURRENT APPLICATION NUMBER: US/60/360,207
;  CURRENT FILING DATE:  2002-03-01
;  NUMBER OF SEQ ID NOS: 47235
```

```
; SEQ ID NO 826
;   LENGTH: 1944
;   TYPE: DNA
;   ORGANISM: HUMAN
US-60-360-207-826

  Query Match            94.4%;  Score 1238;  DB 78;  Length 1944;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1238;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         75 GCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCA 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCA 60

Qy        135 GCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAG 120

Qy        195 CAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGT 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGT 180

Qy        255 TGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGA 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGA 240

Qy        315 CGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCC 300

Qy        375 AGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATG 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATG 360

Qy        435 ATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTT 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTT 420

Qy        495 CTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAG 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAG 480

Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 540

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 600

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 660

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 720

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 840

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 900

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 960

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1020

Qy       1095 TTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTG 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTG 1080

Qy       1155 TTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCA 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCA 1140

Qy       1215 CCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACTA 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACTA 1200

Qy       1275 GTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              |||||||||||||||||||||||||||||||||||||
Db       1201 GTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1238


RESULT 4
US-10-488-619-2153/c
; Sequence 2153, Application US/10488619
; GENERAL INFORMATION:
;  APPLICANT: Greenlee, Winner and Sullivan, P.C.
;  TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
;  TITLE OF INVENTION:  Physiological Conditions, And Genotyping Arrays
;  FILE REFERENCE: 98-01 WO
;  CURRENT APPLICATION NUMBER: US/10/488,619
;  CURRENT FILING DATE:  2004-03-01
;  NUMBER OF SEQ ID NOS: 3040
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2153
;   LENGTH: 774
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-488-619-2153

  Query Match            52.0%;  Score 682.8;  DB 51;  Length 774;
  Best Local Similarity  99.4%;  Pred. No. 4.8e-214;
  Matches  706;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    2;

Qy        593 GTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGT 652
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        768 GTATTTCTACAGGAAG-TATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAG- 711

Qy        653 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        710 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 651

Qy        713 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAAC 772
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        650 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAGC 591

Qy        773 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        590 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 531

Qy        833 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        530 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 471

Qy        893 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 411

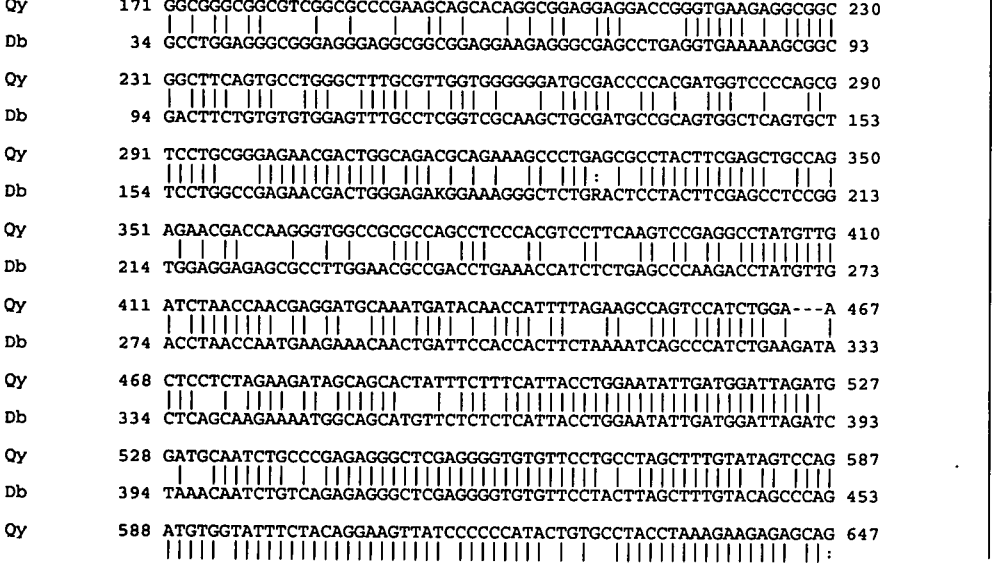
Qy        953 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 351

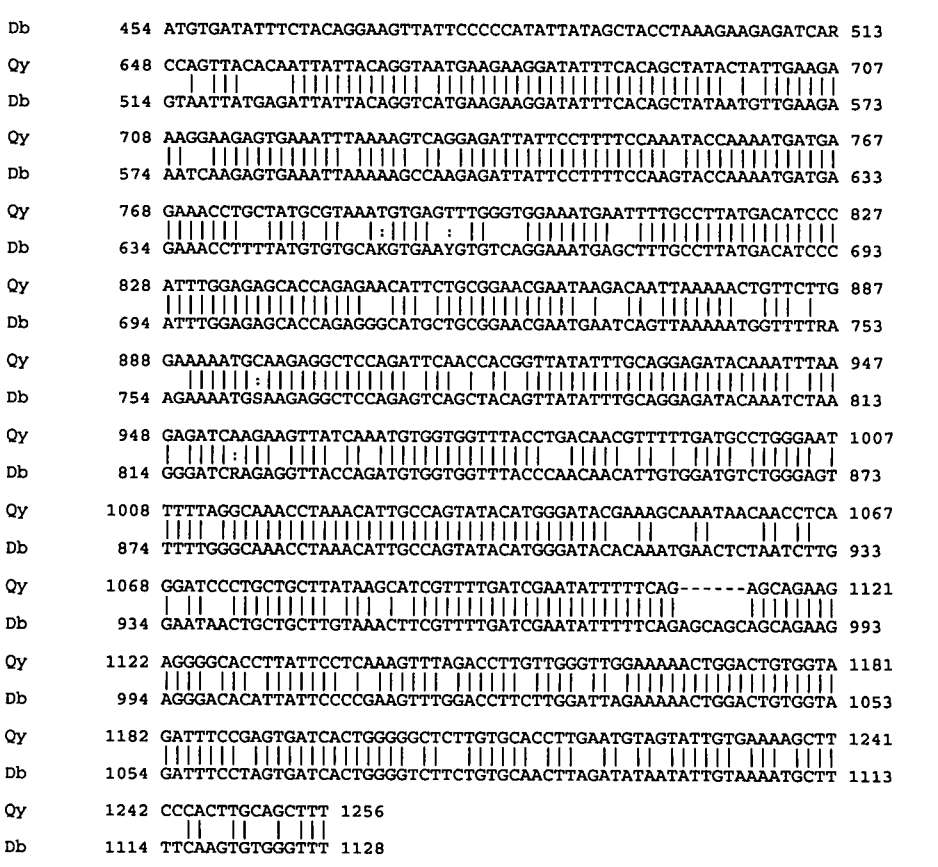
Qy       1013 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 1072
```



RESULT 5
US-60-443-566-1795
; Sequence 1795, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1795
; LENGTH: 1919
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-443-566-1795

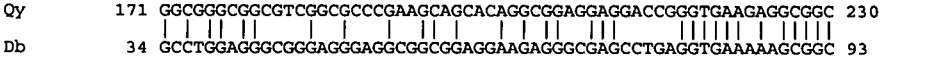
Query Match 51.0%; Score 669; DB 79; Length 1919;
Best Local Similarity 76.6%; Pred. No. 4.9e-209;
Matches 839; Conservative 6; Mismatches 241; Indels 9; Gaps 2;

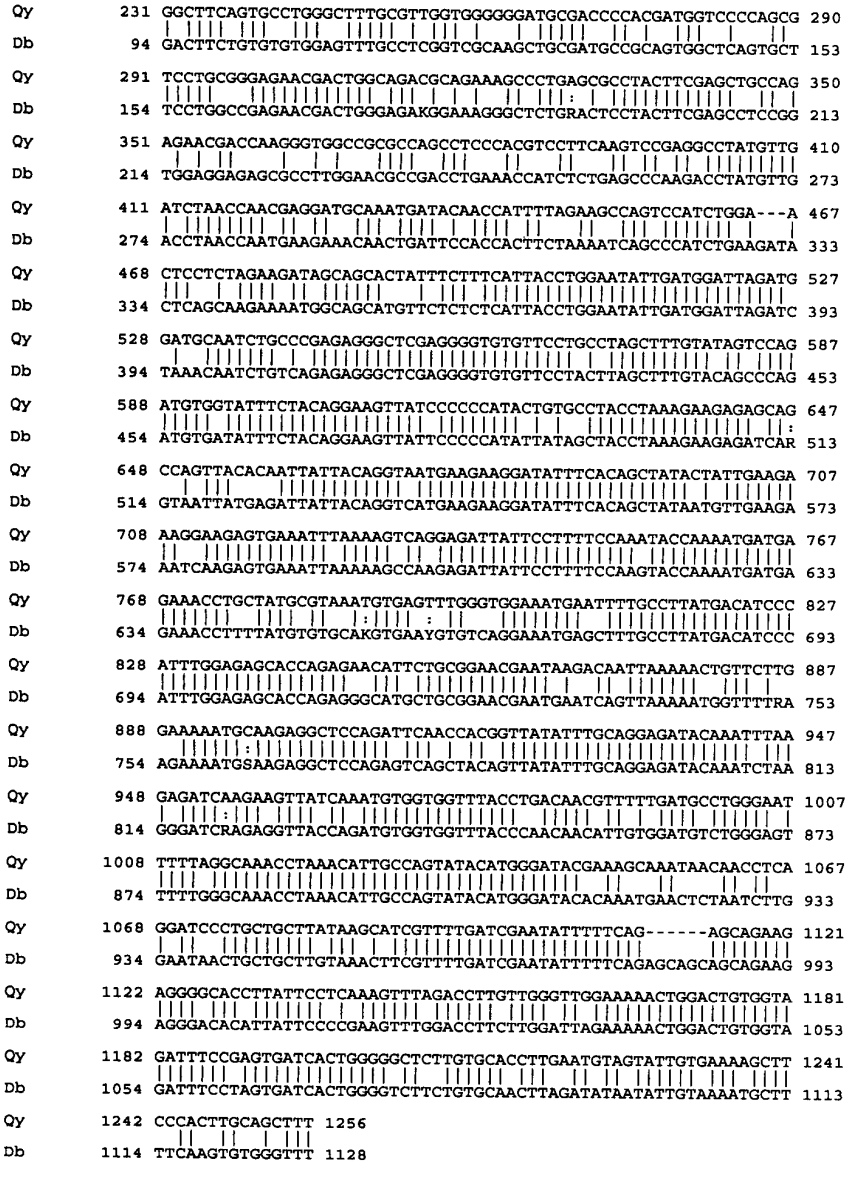




RESULT 6
US-60-452-680-10426
; Sequence 10426, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10426
; LENGTH: 1919
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-452-680-10426

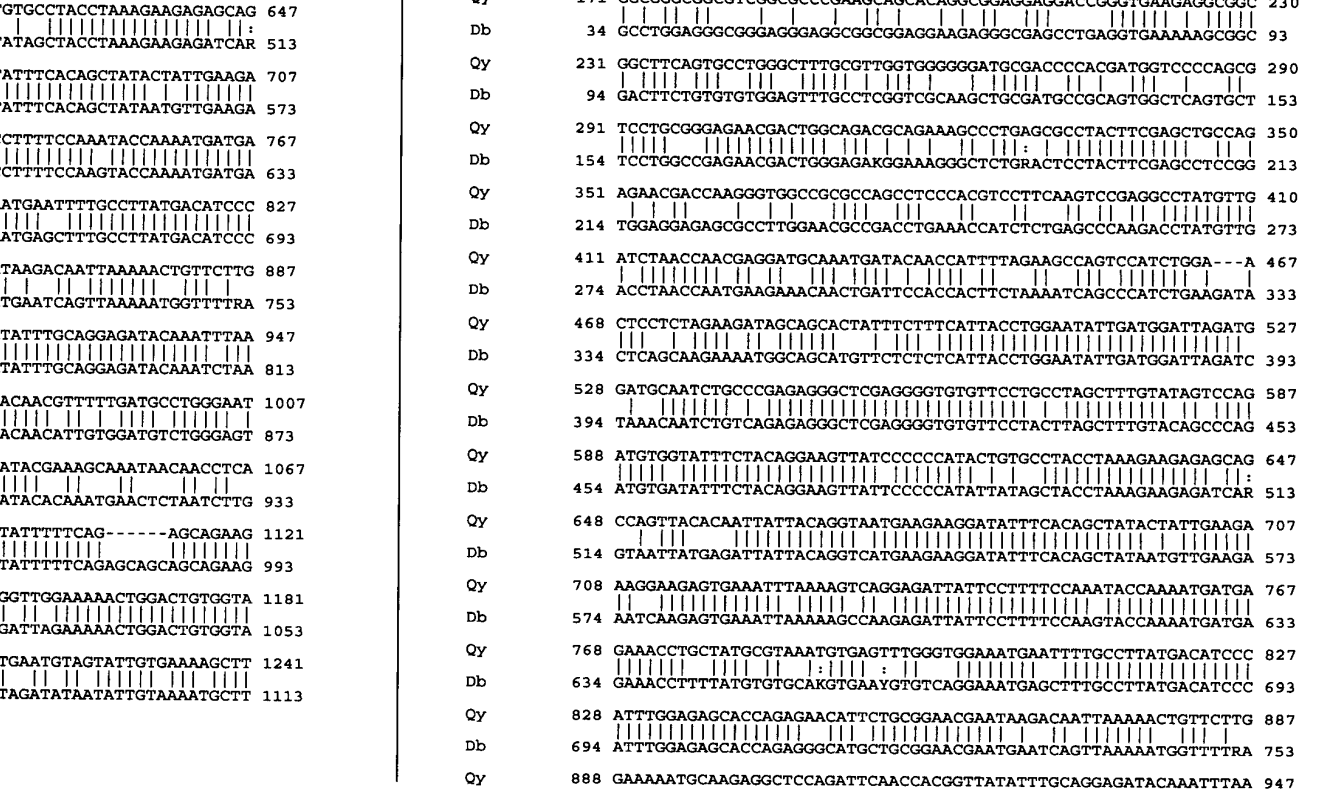
Query Match 51.0%; Score 669; DB 79; Length 1919;
Best Local Similarity 76.6%; Pred. No. 4.9e-209;
Matches 839; Conservative 6; Mismatches 241; Indels 9; Gaps 2;





RESULT 7
US-60-455-444-3324
; Sequence 3324, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3324
; LENGTH: 1919
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-455-444-3324

Query Match 51.0%; Score 669; DB 79; Length 1919;
Best Local Similarity 76.6%; Pred. No. 4.9e-209;
Matches 839; Conservative 6; Mismatches 241; Indels 9; Gaps 2;



```
Db        754 AGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 813

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        814 GGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 873

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        874 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 933

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        934 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 993

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db        994 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1053

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1054 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1113

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1114 TTCAAGTGTGGGTTT 1128


RESULT 8
US-60-465-241-3324
; Sequence 3324, Application US/60465241
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001468
;  CURRENT APPLICATION NUMBER: US/60/465,241
;  CURRENT FILING DATE:  2003-04-23
;  NUMBER OF SEQ ID NOS: 258418
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3324
;   LENGTH: 1919
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-465-241-3324

  Query Match           51.0%;  Score 669;  DB 79;  Length 1919;
  Best Local Similarity   76.6%;  Pred. No. 4.9e-209;
  Matches  839;  Conservative    6;  Mismatches  241;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         34 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 93

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db         94 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 153

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        154 TCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCCTACTTCGAGCCTCCGG 213

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        214 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 273

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        274 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 333

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        334 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 393

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        394 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 453

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        454 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAR 513

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        514 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 573

Qy        708 AAGGAAGAGTGAAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        574 AATCAAGAGTGAAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 633

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        634 GAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 693

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        694 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTRA 753

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        754 AGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 813

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        814 GGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 873

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        874 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 933

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        934 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 993

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db        994 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1053

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1054 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1113

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1114 TTCAAGTGTGGGTTT 1128


RESULT 9
US-10-757-745-1
; Sequence 1, Application US/10757745
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
```

```
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1627)..(1627)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1108)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1849)..(1849)
;   OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

  Query Match           50.9%;  Score 668.2;  DB 61;  Length 1920;
  Best Local Similarity   77.0%;  Pred. No. 9e-209;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   | || |   | | ||  |||     |||||| | |||||
Db         36 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 95

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db         96 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 155

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        156 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 215

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   | | |  ||||  |||   ||   ||   || || || |||||||||
Db        216 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 275

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |   |
Db        276 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 335

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        336 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 395

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||| | ||||||||||| || ||||
Db        396 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 455

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        456 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 515

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db        516 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 575

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        576 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 635

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||  ||||||||  ||||||||||||||||||
Db        636 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 695

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| |||||||||||| |  || |||||||  ||| |
Db        696 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 755

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        756 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 815

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  |||| || | |||| |||||| |
Db        816 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 875

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||  |||   ||    || ||
Db        876 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 935

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||| | |||||||||||||||||||||||      ||||||||
Db        936 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 995

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db        996 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1055

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1056 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1115

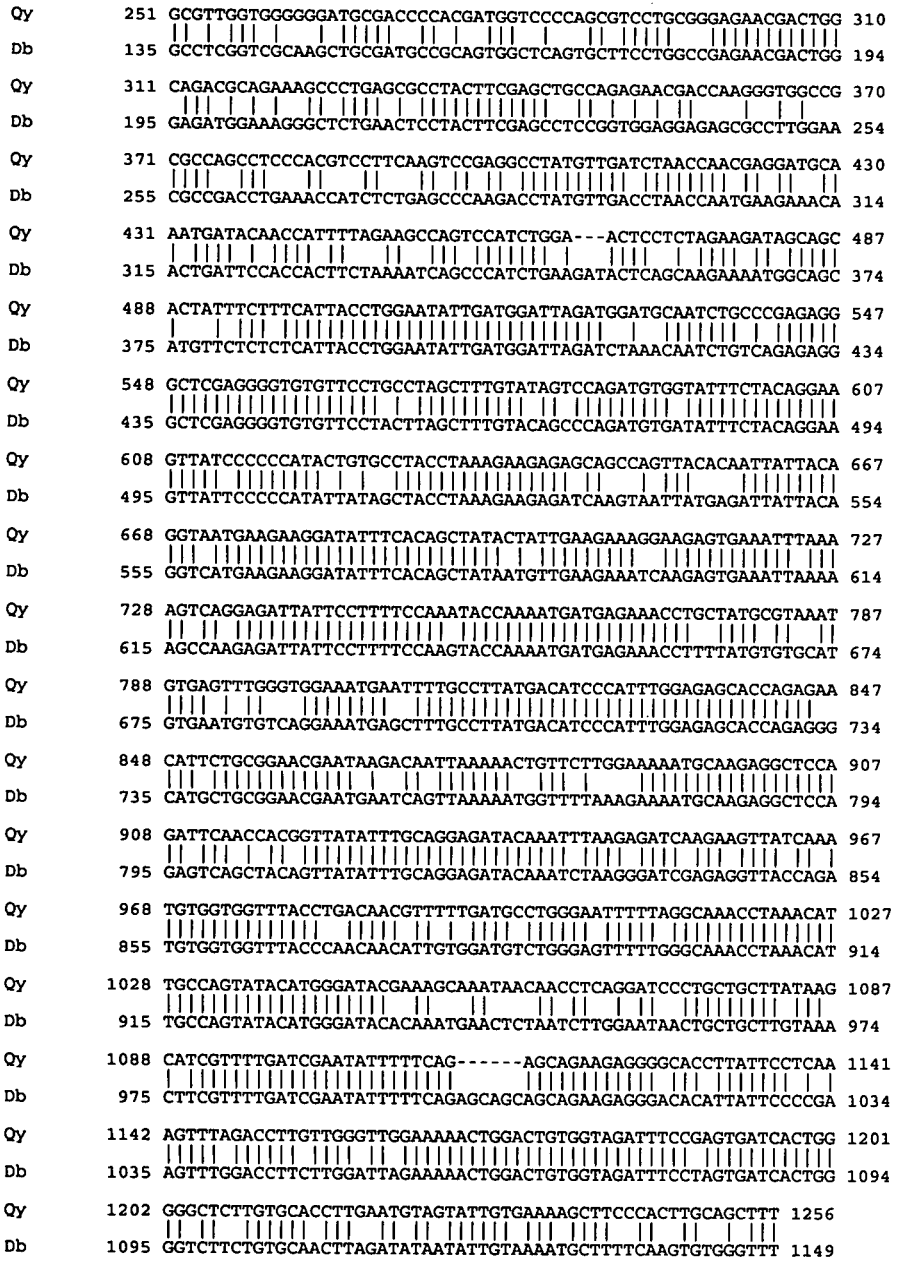
Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1116 TTCAAGTGTGGGTTT 1130


RESULT 10
US-09-488-725A-1200
; Sequence 1200, Application US/09488725A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/09/488,725A
;  CURRENT FILING DATE:  2000-12-22
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7144
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-09-488-725A-1200

  Query Match           50.9%;  Score 668.2;  DB 24;  Length 1948;
  Best Local Similarity   75.6%;  Pred. No. 9.1e-209;
  Matches  858;  Conservative    0;  Mismatches  268;  Indels    9;  Gaps    2;

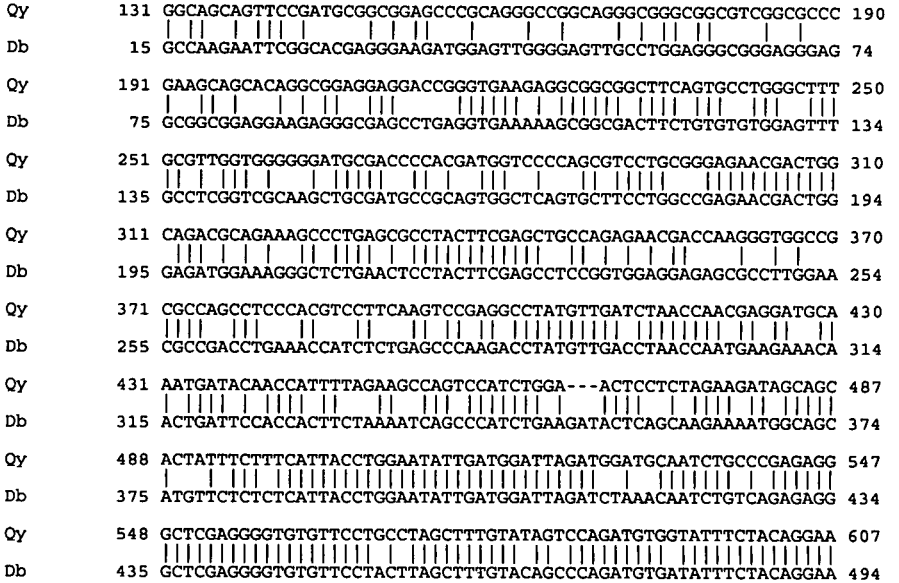
Qy        131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
              | ||  | ||| |     | | ||   |   |  | |  | | || ||    |  |
Db         15 GCCAAGAATTCGGCACGAGGGAAGATGGAGTTGGGAGTTGCCTGGAGGGCGGGAGGGAG 74

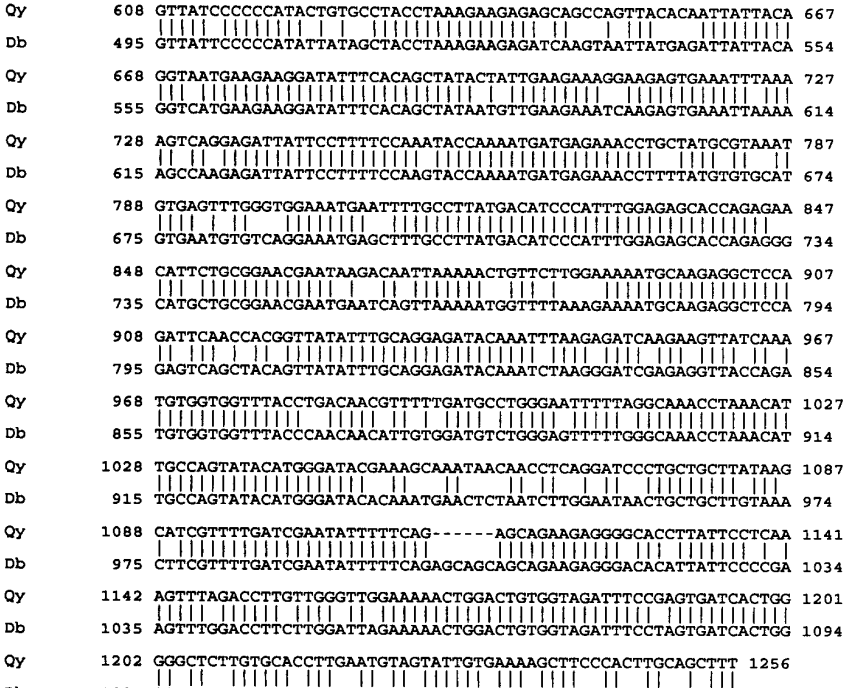
Qy        191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
              |  || |   | | ||  |||      |||||| | |||||| |||| |||  |||  |||
Db         75 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 134
```



RESULT 11
US-10-037-270-889
; Sequence 889, Application US/10037270
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-037-270-889

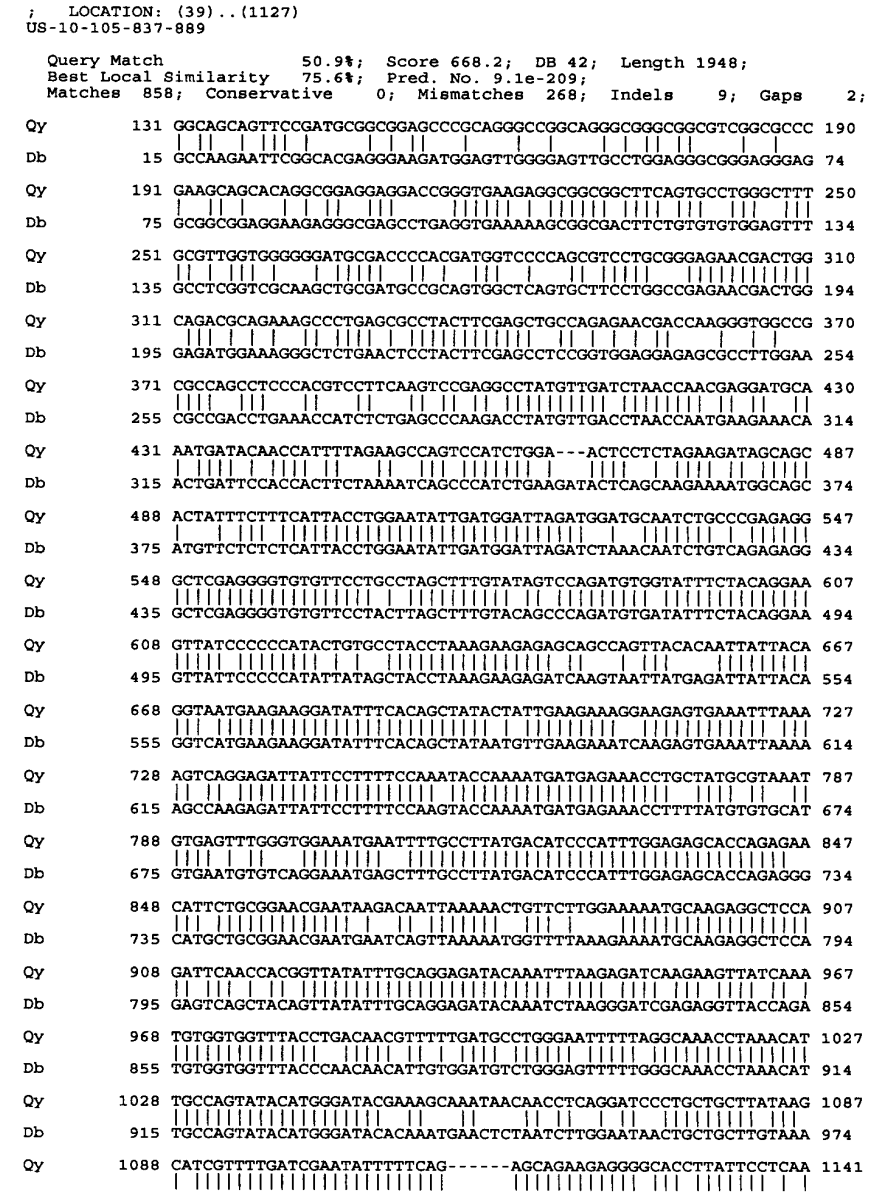
Query Match 50.9%; Score 668.2; DB 40; Length 1948;
Best Local Similarity 75.6%; Pred. No. 9.1e-209;
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;





RESULT 12
US-10-105-837-889
; Sequence 889, Application US/10105837
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BDIVA
; CURRENT APPLICATION NUMBER: US/10/105,837
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-105-837-889

Query Match 50.9%; Score 668.2; DB 42; Length 1948;
Best Local Similarity 75.6%; Pred. No. 9.1e-209;
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;



```
Db        975 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 1034

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
             ||||| |||||| |||| || ||||||||||||||||||||||||||| ||||||||||||
Db       1035 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 1094

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              || ||  |||||| ||| || || |||||| ||| |||| || || | |||
Db       1095 GGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 1149


RESULT 13
US-10-117-722-889
; Sequence 889, Application US/10117722
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BCIP
;  CURRENT APPLICATION NUMBER: US/10/117,722
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-117-722-889

  Query Match             50.9%;  Score 668.2;  DB 42;  Length 1948;
  Best Local Similarity   75.6%;  Pred. No. 9.1e-209;
  Matches  858;  Conservative    0;  Mismatches  268;  Indels    9;  Gaps    2;

Qy        131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
              | ||  | ||| |     | | ||   |    |  |   | | || ||    |  |
Db         15 GCCAAGAATTCGGCACGAGGGAAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAG 74

Qy        191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
              |  || |   | | ||  |||      ||||||| | |||||| |||| |||  |||  |||
Db         75 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 134

Qy        251 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 310
              || | ||| |   | |||||  || |  |||  |   || |||||   ||||||||||||
Db        135 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 194

Qy        311 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 370
              |||  | | |  || |||| | ||||||||||||  || | | | ||    | | |
Db        195 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 254

Qy        371 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 430
              ||||  |||   ||   ||  || || || |||||||||| |||||||| || ||  ||
Db        255 CGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 314

Qy        431 AATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGC 487
              | |||| | |||| ||   ||  ||| ||||||| |   |||| | |||| || |||||
Db        315 ACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGC 374

Qy        488 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 547
              |   | ||| |||||||||||||||||||||||||||||  | ||||||| | ||||||
Db        375 ATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGG 434

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db        435 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 494

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||   | |||    |||||||||
Db        495 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 554

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||||||| | |||||||||  ||||||||||| |||
Db        555 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 614

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        615 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 674

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||| ||||||||||||||||||||||||||||||||||||
Db        675 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 734

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| ||||||||||||| |  || |||||||  ||| |    ||||||||||||||||||
Db        735 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 794

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || |||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        795 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 854

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||
Db        855 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 914

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||||  ||   ||   || ||  | ||  ||||||||| |||
Db        915 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 974

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
              | |||||||||||||||||||||||      |||||||||||| ||| ||||||| | |
Db        975 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 1034

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
              ||||| |||||| |||| || ||||||||||||||||||||||||||| ||||||||||||
Db       1035 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 1094

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              || ||  |||||| ||| || || |||||| ||| |||| || || | |||
Db       1095 GGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 1149


RESULT 14
US-10-122-851-889
; Sequence 889, Application US/10122851
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BDV3
;  CURRENT APPLICATION NUMBER: US/10/122,851
;  CURRENT FILING DATE:  2002-04-12
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
```

```
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-122-851-889
```

Query Match 50.9%; Score 668.2; DB 42; Length 1948;
Best Local Similarity 75.6%; Pred. No. 9.1e-209;
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;

```
Qy        131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
              | ||  | ||| |       | | |||  |   |  |   | | || ||   |  |
Db         15 GCCAAGAATTCGGCACGAGGGAAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAG 74

Qy        191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
              |   ||  |    | | ||  |||      |||||| | |||||| |||| |||   |||  |||
Db         75 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 134

Qy        251 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 310
              || | |||  |   | |||||  || |  |||  |   || |||||    ||||||||||||
Db        135 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 194

Qy        311 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 370
               ||| | | |  || |||| | ||||||||||||  || | | | ||     | | |
Db        195 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 254

Qy        371 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 430
              ||||  |||   ||   ||  || || || |||||||||| ||||||||  || ||  ||
Db        255 CGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 314

Qy        431 AATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGC 487
              | |||| | |||| ||   ||  ||| ||||||| |   ||||  | |||| || |||||
Db        315 ACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGC 374

Qy        488 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 547
              |   | ||| ||||||||||||||||||||||||||||||  |  |||||| | ||||||
Db        375 ATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGG 434

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||||
Db        435 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 494

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  |||||||||||||||| ||    | |||   |||||||||
Db        495 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 554

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        555 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 614

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        615 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 674

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||| |||||||||||||||||||||||||||||||||||||
Db        675 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 734

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||| |  || |||||||  ||| |    ||||||||||||||||||
Db        735 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 794

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || |||||||||||||||||||||| |||| |||| ||| |||| || |
Db        795 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 854

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||  ||||| || | |||| |||||| |||| ||||||||||||||||
Db        855 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 914

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||  ||   ||   || ||  | ||  ||||||||| |||
Db        915 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 974

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
              |_|||||||||||||||||||||||      |||||||||||| ||| ||||||| | |
Db         975 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 1034

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
              ||||| |||||| |||| || ||||||||||||||||||||||||||| ||||||||||||
Db        1035 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 1094

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              || ||  |||||| |||  || || |||||| ||| ||||  ||  ||  | |||
Db        1095 GGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 1149
```

```
RESULT 15
US-10-258-898A-1200
; Sequence 1200, Application US/10258898A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/258,898A
;  CURRENT FILING DATE:  2002-10-29
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-258-898A-1200
```

Query Match 50.9%; Score 668.2; DB 43; Length 1948;
Best Local Similarity 75.6%; Pred. No. 9.1e-209;
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;

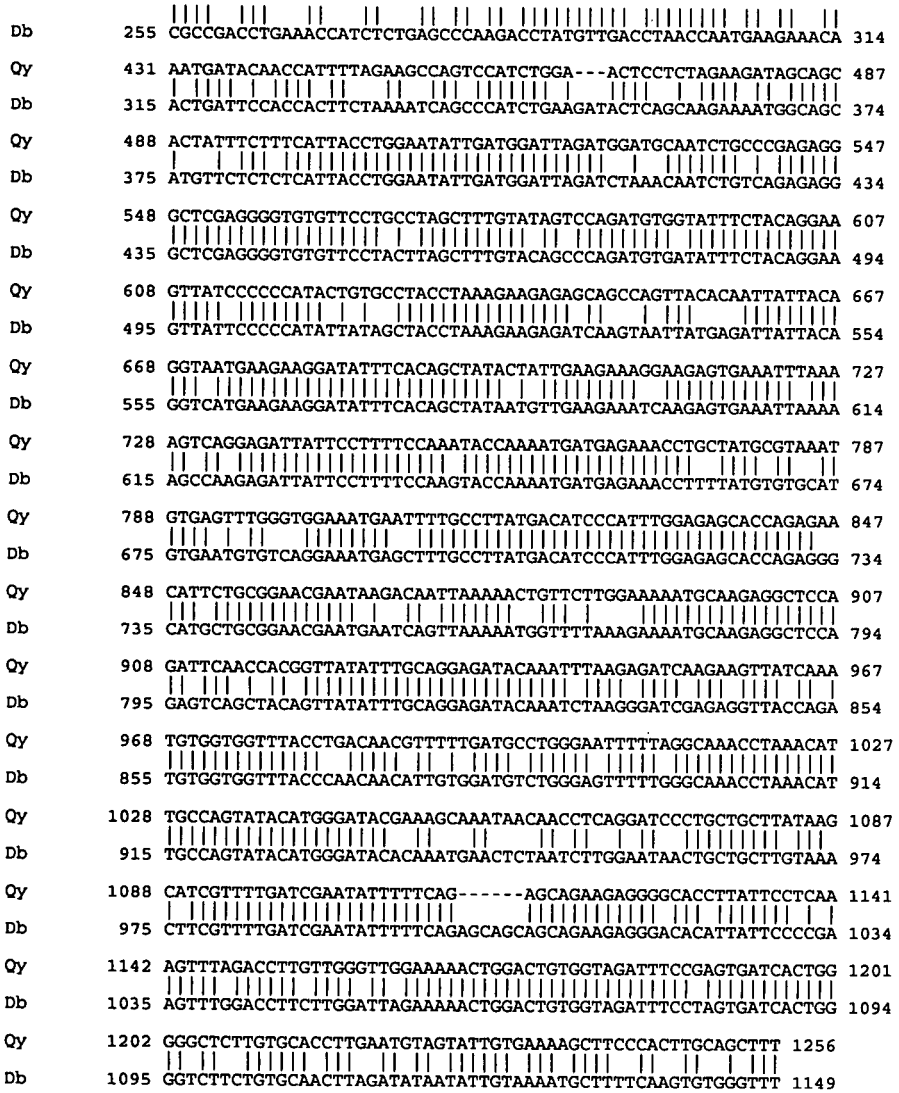
```
Qy        131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
              | ||  | ||| |       | | |||  |   |  |   | | || ||   |  |
Db         15 GCCAAGAATTCGGCACGAGGGAAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAG 74

Qy        191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
              |   ||  |    | | ||  |||      |||||| | |||||| |||| |||   |||  |||
Db         75 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 134

Qy        251 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 310
              || | |||  |   | |||||  || |  |||  |   || |||||    ||||||||||||
Db        135 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 194

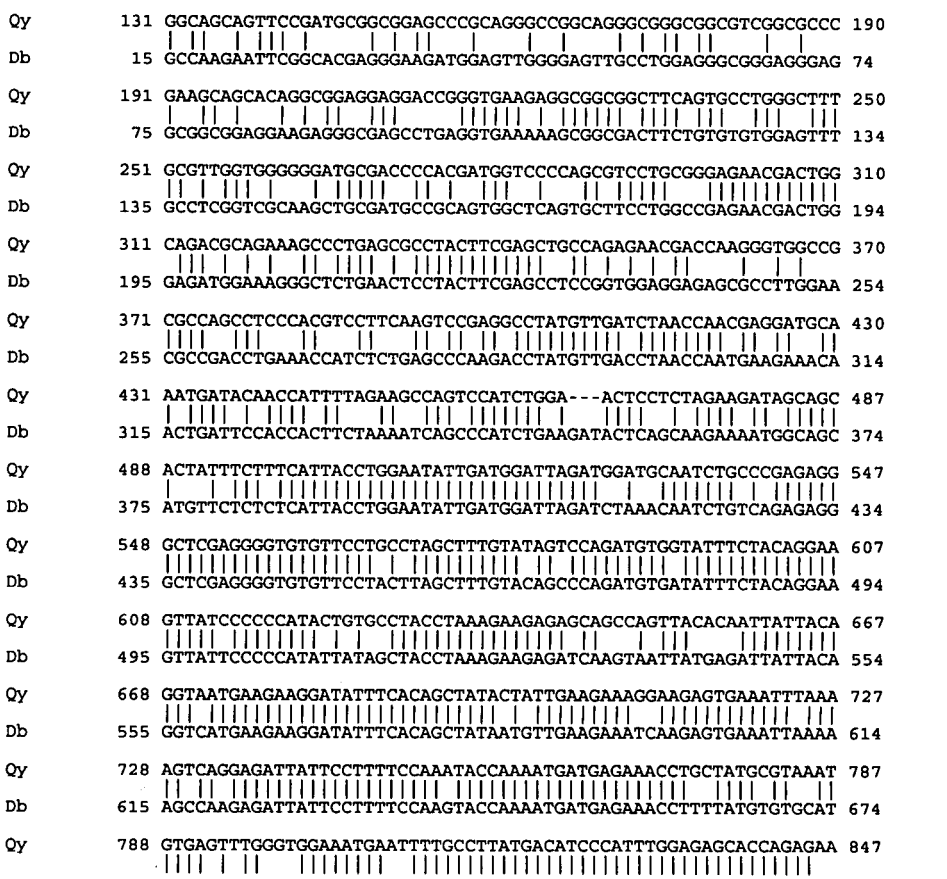
Qy        311 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 370
               ||| | | |  || |||| | ||||||||||||  || | | | ||     | | |
Db        195 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 254

Qy        371 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 430
```



```
RESULT 16
US-10-286-897-1200
; Sequence 1200, Application US/10286897
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/286,897
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US/09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US/09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US/09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US/09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US/09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US/09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US/09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-286-897-1200

  Query Match            50.9%;  Score 668.2;  DB 43;  Length 1948;
  Best Local Similarity  75.6%;  Pred. No. 9.1e-209;
  Matches  858;  Conservative    0;  Mismatches  268;  Indels    9;  Gaps    2;
```



```
Db        675 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 734

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907

Db        735 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 794

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967

Db        795 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 854

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027

Db        855 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 914

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087

Db        915 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 974

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141

Db        975 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 1034

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201

Db       1035 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 1094

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256

Db       1095 GGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 1149


RESULT 17
US-09-471-275-1314
; Sequence 1314, Application US/09471275
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Contigs Obtained
;  TITLE OF INVENTION: From Various Libraries
;  FILE REFERENCE: 782
;  CURRENT APPLICATION NUMBER: US/09/471,275
;  CURRENT FILING DATE: 1999-12-23
;  EARLIER APPLICATION NUMBER: US 09/235,076
;  EARLIER FILING DATE: 1999-01-20
;  EARLIER APPLICATION NUMBER: US 09/234,611
;  EARLIER FILING DATE: 1999-01-22
;  EARLIER APPLICATION NUMBER: US 09/240,371
;  EARLIER FILING DATE: 1999-01-29
;  EARLIER APPLICATION NUMBER: US 09/277,227
;  EARLIER FILING DATE: 1999-03-25
;  EARLIER APPLICATION NUMBER: US 09/271,490
;  EARLIER FILING DATE: 1999-03-18
;  EARLIER APPLICATION NUMBER: US 09/293,972
;  EARLIER FILING DATE: 1999-04-15
;  EARLIER APPLICATION NUMBER: US 09/274,861
;  EARLIER FILING DATE: 1999-03-23
;  EARLIER APPLICATION NUMBER: US 60/125,453
;  EARLIER FILING DATE: 1999-03-19
;  EARLIER APPLICATION NUMBER: US 60/126,605
;  EARLIER FILING DATE: 1999-03-26
;  EARLIER APPLICATION NUMBER: US 09/306,350
;  EARLIER FILING DATE: 1999-05-07
;  EARLIER APPLICATION NUMBER: US 09/399,720
;  EARLIER FILING DATE: 1999-09-21
;  EARLIER APPLICATION NUMBER: US 09/404,284
;  EARLIER FILING DATE: 1999-09-21
;  EARLIER APPLICATION NUMBER:
;  EARLIER FILING DATE: 1999-12-16
;  NUMBER OF SEQ ID NOS: 10451
;  SOFTWARE: pt_CT_genes Version 1.0
; SEQ ID NO 1314
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens

;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-471-275-1314

  Query Match           50.9%;  Score 668.2;  DB 24;  Length 1958;
  Best Local Similarity 77.0%;  Pred. No. 9.2e-209;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         67 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 126

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        127 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 186

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        187 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 246

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        247 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 306

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        307 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 366

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        367 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 426

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        427 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 486

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        487 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 546

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        547 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 606

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        607 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 666

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        667 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 726

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        727 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 786

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        787 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 846

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        847 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 906

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        907 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 966

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
```

```
Db        967 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1026

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       1027 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1086

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1087 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1146

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1147 TTCAAGTGTGGGTTT 1161


RESULT 18
US-09-488-725A-4772
; Sequence 4772, Application US/09488725A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/09/488,725A
;  CURRENT FILING DATE:  2000-12-22
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7144
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-488-725A-4772

  Query Match             50.9%;  Score 668.2;  DB 24;  Length 1958;
  Best Local Similarity   77.0%;  Pred. No. 9.2e-209;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         67 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 126

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        127 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 186

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        187 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 246

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        247 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 306

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        307 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 366

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        367 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 426

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        427 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 486

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        487 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 546

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        547 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 606

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        607 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 666

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        667 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 726

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        727 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 786

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        787 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 846

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        847 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 906

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        907 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 966

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        967 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1026

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       1027 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1086

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1087 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1146

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1147 TTCAAGTGTGGGTTT 1161


RESULT 19
US-09-488-725B-8295
; Sequence 8295, Application US/09488725B
; GENERAL INFORMATION:
;  APPLICANT: Yuanhua T. Tang
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Ankura Sinku
;  APPLICANT:  Chenghua Liu
;  APPLICANT:  Radoje T. Drmanac
;  TITLE OF INVENTION: Novel Contigs Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 784
;  CURRENT APPLICATION NUMBER: US/09/488,725B
;  CURRENT FILING DATE:  2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/004,182
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: US 09/034,341
```

```
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 09/045,400
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: US 09/321,214
;  PRIOR FILING DATE: 1999-05-26
;  PRIOR APPLICATION NUMBER: US 09/131,598
;  PRIOR FILING DATE: 1998-08-10
;  PRIOR APPLICATION NUMBER: US 09/170,294
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: US 09/179,473
;  PRIOR FILING DATE: 1998-10-27
;  PRIOR APPLICATION NUMBER: US 09/181,430
;  PRIOR FILING DATE: 1998-10-28
;  PRIOR APPLICATION NUMBER: US 09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  PRIOR APPLICATION NUMBER: US 09/234,611
;  PRIOR FILING DATE: 1999-01-22
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10289
;  SOFTWARE: pt_CT_genes Version 1.01
; SEQ ID NO 8295
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-488-725B-8295

  Query Match             50.9%;  Score 668.2;  DB 24;  Length 1958;
  Best Local Similarity   77.0%;  Pred. No. 9.2e-209;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||      |||||| | |||||
Db         67 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 126

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        127 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 186

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        187 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 246

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   |||| |||   ||   ||  || || || |||||||||
Db        247 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 306

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||| |   |
Db        307 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 366

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        367 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 426

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||||| | ||||||||||||||||||||||||||| | ||||||||| || ||||
Db        427 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 486

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        487 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 546

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db        547 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 606

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
```

```
Db        607 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 666

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   ||||||||  ||||||||||||||||||
Db        667 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 726

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||||  ||| ||||||||||||| |  || |||||||  ||| |
Db        727 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 786

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||| ||| | || ||||||||||||||||||||||| |||
Db        787 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 846

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || |||||||||||||||  ||||| || | |||| ||||||
Db        847 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 906

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| |||||||||||||||||||||||||||||||||||||||||  ||   ||   || ||
Db        907 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 966

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  ||||||||| ||| | |||||||||||||||||||||||      ||||||||
Db        967 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1026

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db       1027 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1086

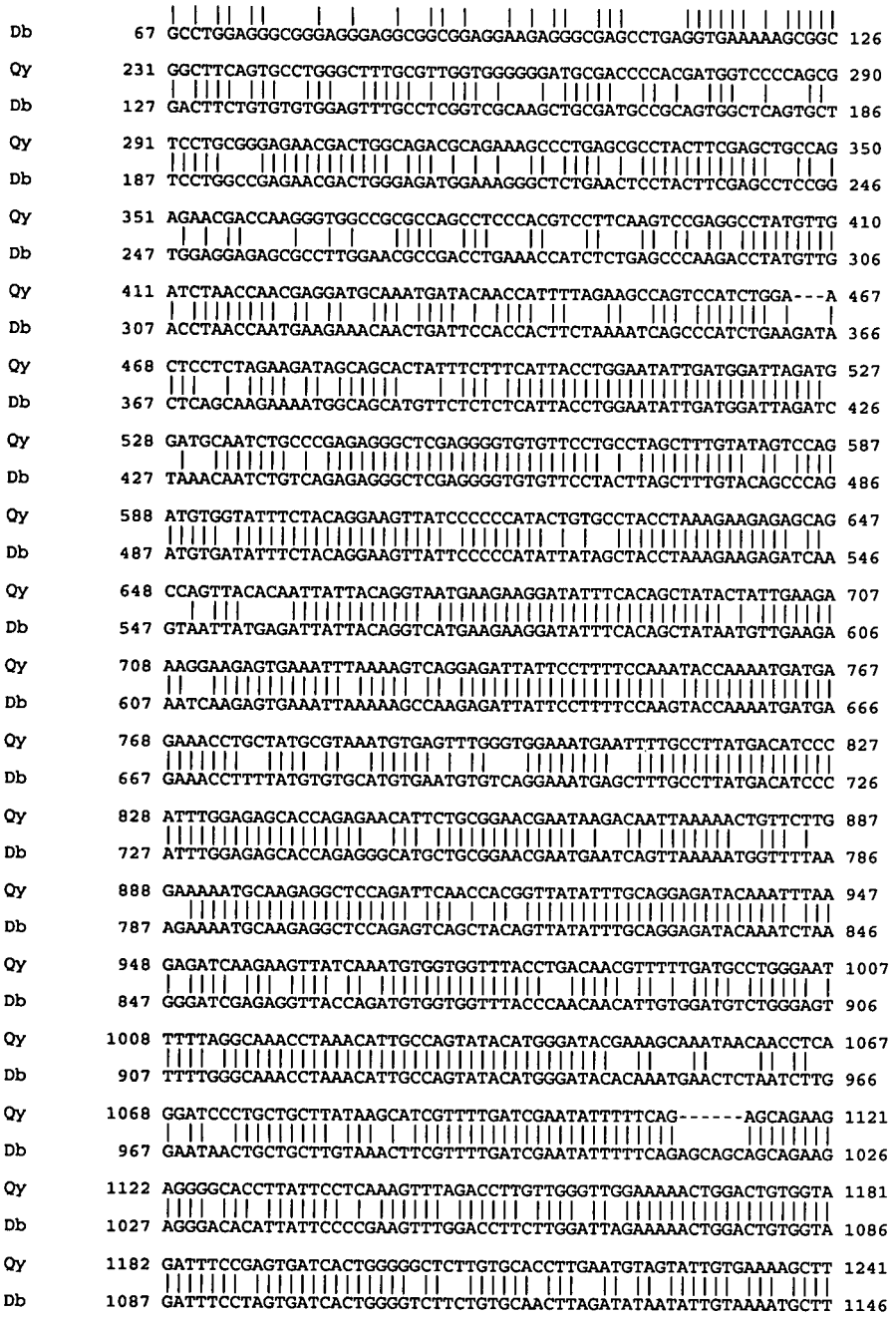
Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1087 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1146

Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1147 TTCAAGTGTGGGTTT 1161
```

```
RESULT 20
US-09-552-317-8295
; Sequence 8295, Application US/09552317
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Sinku, Ankura
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids
;  TITLE OF INVENTION:  and Polypeptides
;  FILE REFERENCE: 784CIP
;  CURRENT APPLICATION NUMBER: US/09/552,317
;  CURRENT FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 10289
;  SOFTWARE: pt_CT_genes Version 1.01
; SEQ ID NO 8295
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-552-317-8295

  Query Match             50.9%;  Score 668.2;  DB 27;  Length 1958;
  Best Local Similarity   77.0%;  Pred. No. 9.2e-209;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

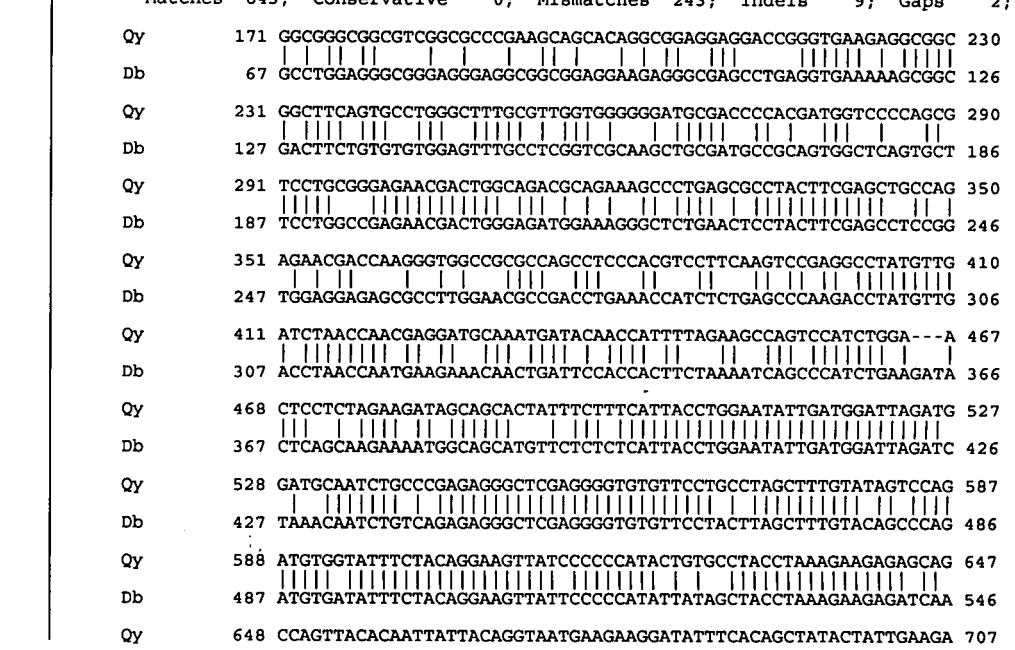
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
```

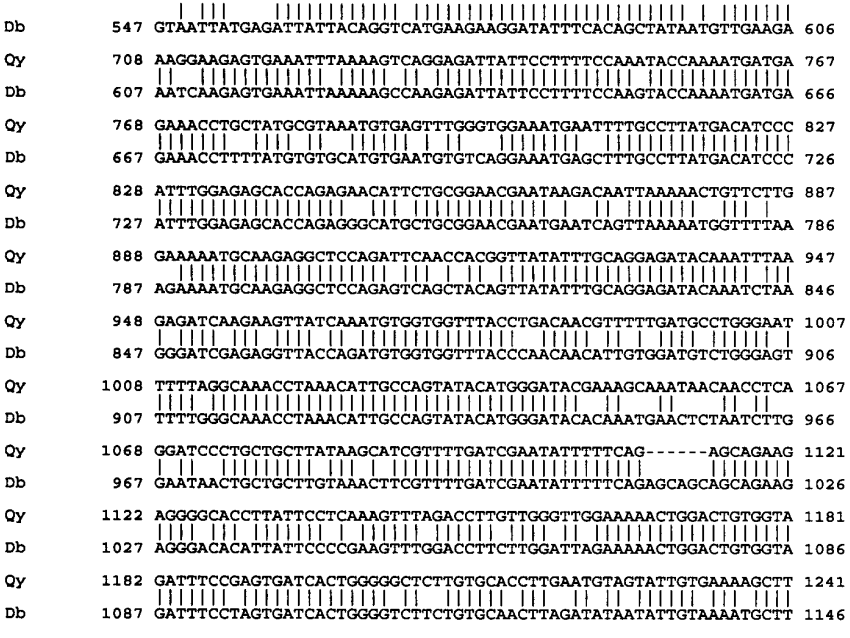


Db 1147 TTCAAGTGTGGGTTT 1161

RESULT 21
US-10-258-898A-4772
; Sequence 4772, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/258,898A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
; LENGTH: 1958
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-258-898A-4772

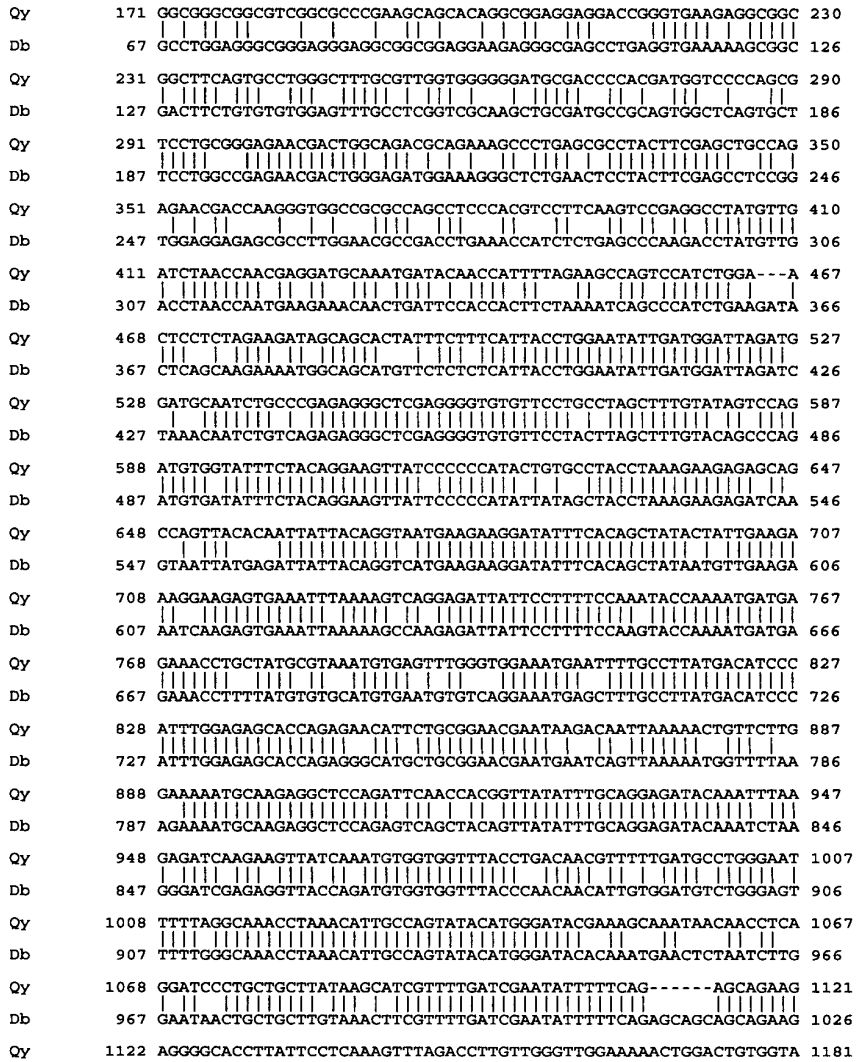
Query Match 50.9%; Score 668.2; DB 43; Length 1958;
Best Local Similarity 77.0%; Pred. No. 9.2e-209;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;





RESULT 22
US-10-286-897-4772
; Sequence 4772, Application US/10286897
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
; LENGTH: 1958
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-286-897-4772

Query Match 50.9%; Score 668.2; DB 43; Length 1958;
Best Local Similarity 77.0%; Pred. No. 9.2e-209;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;



```
Db       1027 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1086

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1087 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1146

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1147 TTCAAGTGTGGGTTT 1161


RESULT 23
US-10-302-689A-59957
; Sequence 59957, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ballinger, Dennis
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Liu, Jin
;  APPLICANT:  Loeb, Deborah
;  APPLICANT:  Montgomery, Julia, R.
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 59957
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-302-689A-59957

  Query Match             50.9%;  Score 668.2;  DB 48;  Length 1958;
  Best Local Similarity   77.0%;  Pred. No. 9.2e-209;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         67 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 126

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        127 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 186

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        187 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 246

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        247 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 306

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        307 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 366

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        367 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 426

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        427 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 486

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        487 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 546

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        547 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 606

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        607 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 666

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        667 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 726

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        727 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 786

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        787 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 846

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        847 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 906

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        907 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 966

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        967 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1026

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       1027 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1086

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1087 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1146

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1147 TTCAAGTGTGGGTTT 1161


RESULT 24
US-10-302-689A-77389
; Sequence 77389, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
```

```
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ballinger, Dennis
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Liu, Jin
;  APPLICANT:  Loeb, Deborah
;  APPLICANT:  Montgomery, Julia, R.
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 77389
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-302-689A-77389

  Query Match             50.9%;  Score 668.2;  DB 48;  Length 1958;
  Best Local Similarity   77.0%;  Pred. No. 9.2e-209;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     |||||| | |||||
Db         67 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 126

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |    ||
Db        127 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 186

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||   || |
Db        187 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 246

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |   ||||  |||   ||   ||   || || || |||||||||
Db        247 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 306

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | ||||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||| |   |
Db        307 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 366

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||    | ||| ||||||||||||||||||||||||||||
Db        367 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 426

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
                |  ||||||| | |||||||||||||||||||||||||| | |||||||||| || ||||

Db        427 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 486

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              |||||  |||||||||||||||||||  |||||||| | |  |||||||||||||| ||
Db        487 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 546

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
                | |||     |||||||||||| |||||||||||||||||||||||||| | |||||||
Db        547 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 606

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        607 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 666

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   ||||||||  ||||||||||||||||||
Db        667 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 726

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||||  ||| ||||||||||||| |  || |||||||  ||| |
Db        727 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 786

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        787 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 846

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || |||||||||||||  ||||| || | |||| |||||| |
Db        847 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 906

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||||||||||  ||    || ||
Db        907 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 966

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              |||   |||||||| ||| | |||||||||||||||||||||||      ||||||||
Db        967 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1026

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db       1027 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1086

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1087 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1146

Qy       1242 CCCACTTGCAGCTTT 1256
                ||  ||  | |||
Db       1147 TTCAAGTGTGGGTTT 1161


RESULT 25
US-09-606-680-2868
; Sequence 2868, Application US/09606680
; GENERAL INFORMATION:
;  APPLICANT: Lloyd, Clare M.
;  APPLICANT:  Williamson, Mark
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 1600.1131-001
;  CURRENT APPLICATION NUMBER: US/09/606,680
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER: 60/141,227
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 60/141,226
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 4394
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2868
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-606-680-2868

  Query Match             50.9%;  Score 668.2;  DB 28;  Length 2262;
  Best Local Similarity   77.0%;  Pred. No. 1.1e-208;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||   |  |  |  || |   | ||  |||     ||||||| | |||||
Db        331 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 390

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        391 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 450

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        451 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 510

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   |  | |   |||| |||   ||   ||  || || || |||||||||
Db        511 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 570

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        571 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 630

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||
Db        631 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 690

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||||| | ||||||||||| || ||||
Db        691 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 750

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        751 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 810

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db        811 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 870

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| |||||| || |||||||||||||||||||| ||||||||||||||
Db        871 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 930

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   ||||||||  ||||||||||||||||||
Db        931 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 990

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||  |  || |||||||  ||| |
Db        991 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1050

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || |||||||||||||||||||||||| |||
Db       1051 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 1110

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||||  |||||  || | |||| |||||| |
Db       1111 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 1170

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| |||||||||||||||||||||||||||||||||||  ||   ||   || ||
Db       1171 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 1230

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||| | |||||||||||||||||||||||      ||||||||
Db       1231 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1290

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db       1291 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1350

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1351 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1410

Qy       1242 CCCACTTGCAGCTTT 1256
              ||  ||  | |||
Db       1411 TTCAAGTGTGGGTTT 1425


RESULT 26
US-09-606-776-4085
; Sequence 4085, Application US/09606776
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  APPLICANT:  Myers, Paul
;  APPLICANT:  Gearing, David P.
;  APPLICANT:  Pan, Yang
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 1600.1129-001
;  CURRENT APPLICATION NUMBER: US/09/606,776
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER: 60/141,578
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 60/141,379
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: 60/141,138
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: 60/141,581
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 5415
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4085
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-606-776-4085

  Query Match             50.9%;  Score 668.2;  DB 28;  Length 2262;
  Best Local Similarity   77.0%;  Pred. No. 1.1e-208;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||   |  |  |  || |   | ||  |||     ||||||| | |||||
Db        331 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 390

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        391 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 450

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        451 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 510

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   |  | |   |||| |||   ||   ||  || || || |||||||||
Db        511 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 570

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        571 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 630

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||
```

```
Db        631 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 690

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        691 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 750

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        751 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 810

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        811 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 870

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        871 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 930

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        931 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 990

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        991 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1050

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db       1051 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 1110

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db       1111 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 1170

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db       1171 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 1230

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db       1231 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1290

Qy       1122 AGGGGCACCTTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       1291 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1350

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1351 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1410

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1411 TTCAAGTGTGGGTTT 1425


RESULT 27
US-09-644-869-9498
; Sequence 9498, Application US/09644869
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  APPLICANT:  McCarthy, Sean A.
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Monahan, John
;  APPLICANT:  Richardson, Jennifer
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1182-001
;  CURRENT APPLICATION NUMBER: US/09/644,869
;  CURRENT FILING DATE:  2000-08-28
;  PRIOR APPLICATION NUMBER:  60/151,062
;  PRIOR FILING DATE:  1999-08-27
;  NUMBER OF SEQ ID NOS:  9708
;  SOFTWARE:  FastSEQ for Windows Version 4.0

; SEQ ID NO 9498
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-644-869-9498

  Query Match           50.9%;  Score 668.2;  DB 29;  Length 2262;
  Best Local Similarity  77.0%;  Pred. No. 1.1e-208;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db        331 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 390

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        391 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 450

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        451 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 510

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        511 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 570

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        571 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 630

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        631 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 690

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        691 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 750

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        751 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 810

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        811 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 870

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        871 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 930

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        931 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 990

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        991 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1050

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db       1051 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 1110

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db       1111 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 1170

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db       1171 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 1230
```

```
Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db       1231 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1290

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db       1291 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1350

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db       1351 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1410

Qy       1242 CCCACTTGCAGCTTT 1256
Db       1411 TTCAAGTGTGGGTTT 1425
```

RESULT 28
US-09-652-121-6794
; Sequence 6794, Application US/09652121
; GENERAL INFORMATION:
; APPLICANT: DiStefano, Peter
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1188-001
; CURRENT APPLICATION NUMBER: US/09/652,121
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,129
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 7615
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6794
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2262)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-121-6794

Query Match 50.9%; Score 668.2; DB 29; Length 2262;
Best Local Similarity 77.0%; Pred. No. 1.1e-208;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db        331 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 390

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db        391 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 450

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db        451 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 510

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db        511 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 570

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db        571 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 630

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db        631 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 690

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db        691 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 750
```

```
Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db        751 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 810

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db        811 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 870

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db        871 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 930

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db        931 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 990

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db        991 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1050

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db       1051 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 1110

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db       1111 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 1170

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db       1171 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 1230

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db       1231 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1290

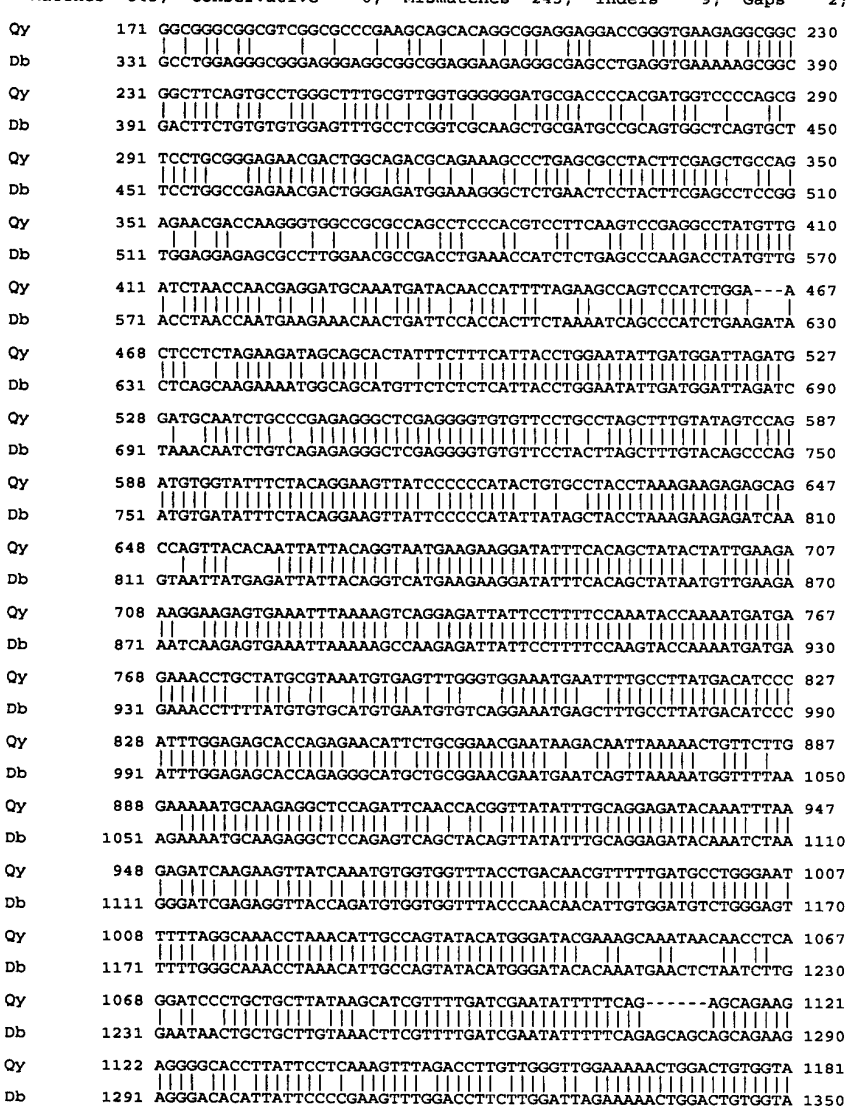
Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db       1291 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1350

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db       1351 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1410

Qy       1242 CCCACTTGCAGCTTT 1256
Db       1411 TTCAAGTGTGGGTTT 1425
```

RESULT 29
US-09-652-122-3756
; Sequence 3756, Application US/09652122
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1194-001
; CURRENT APPLICATION NUMBER: US/09/652,122
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,421
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 5020
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3756
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2262)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-122-3756

Query Match 50.9%; Score 668.2; DB 29; Length 2262;
Best Local Similarity 77.0%; Pred. No. 1.1e-208;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

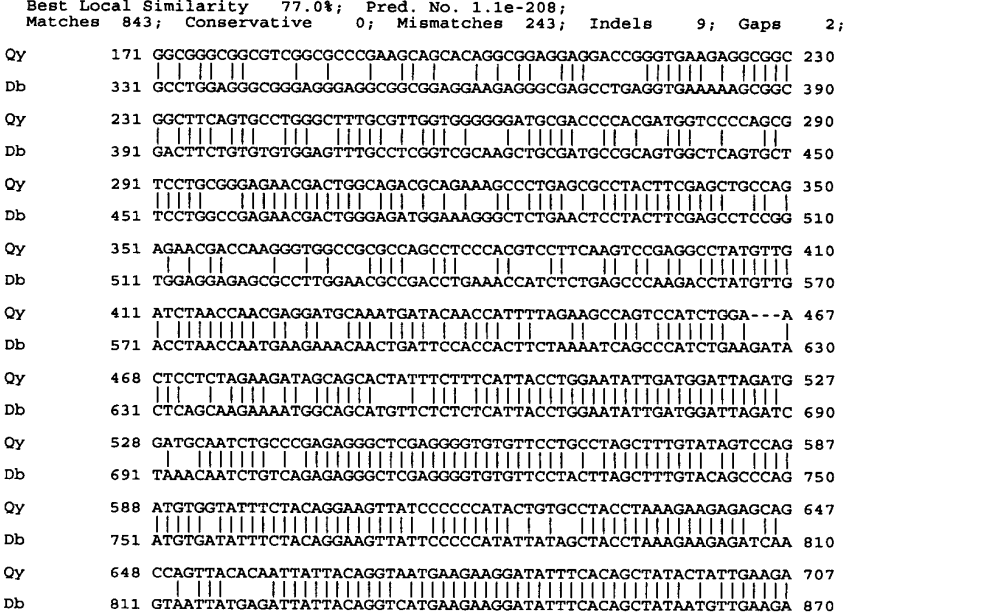


Query 1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db 1351 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1410

Query 1242 CCCACTTGCAGCTTT 1256
Db 1411 TTCAAGTGTGGGTTT 1425

RESULT 30
US-09-652-128-9329
; Sequence 9329, Application US/09652128
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1171-001
; CURRENT APPLICATION NUMBER: US/09/652,128
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,133
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10265
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9329
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2262)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-128-9329

Query Match 50.9%; Score 668.2; DB 29; Length 2262;
Best Local Similarity 77.0%; Pred. No. 1.1e-208;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;



```
Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db       871 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 930

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db       931 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 990

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db       991 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1050

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db      1051 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 1110

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db      1111 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 1170

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db      1171 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 1230

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db      1231 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1290

Qy      1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db      1291 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1350

Qy      1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db      1351 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1410

Qy      1242 CCCACTTGCAGCTTT 1256
Db      1411 TTCAAGTGTGGGTTT 1425
```

```
RESULT 31
US-09-652-816-9572
; Sequence 9572, Application US/09652816
; GENERAL INFORMATION:
;  APPLICANT:  Gutierrez-Ramos, Jose-Carlos
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1177-001
;  CURRENT APPLICATION NUMBER: US/09/652,816
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER:  60/152,111
;  PRIOR FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  9647
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9572
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-816-9572

  Query Match           50.9%;  Score 668.2;  DB 29;  Length 2262;
  Best Local Similarity  77.0%;  Pred. No. 1.1e-208;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy       171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db       331 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 390

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db       391 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 450

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db       451 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 510

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db       511 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 570

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db       571 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 630

Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db       631 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 690

Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db       691 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 750

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db       751 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 810

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db       811 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 870

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db       871 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 930

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db       931 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 990

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db       991 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1050

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db      1051 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 1110

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db      1111 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 1170

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db      1171 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 1230

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db      1231 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1290

Qy      1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db      1291 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1350

Qy      1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db      1351 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1410

Qy      1242 CCCACTTGCAGCTTT 1256
Db      1411 TTCAAGTGTGGGTTT 1425
```

```
RESULT 32
US-09-700-000-6419
; Sequence 6419, Application US/09700000
; GENERAL INFORMATION:
;  APPLICANT:  Richardson, Jennifer
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2022-001
;  CURRENT APPLICATION NUMBER: US/09/700,000
;  CURRENT FILING DATE:  2000-10-30
;  PRIOR APPLICATION NUMBER:  60/162,619
;  PRIOR FILING DATE:  1999-10-29
;  NUMBER OF SEQ ID NOS:  7171
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6419
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-700-000-6419

  Query Match             50.9%;  Score 668.2;  DB 31;  Length 2262;
  Best Local Similarity   77.0%;  Pred. No. 1.1e-208;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||   | |  |  || |  | | ||  |||    ||||| | |||||
Db        331 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 390

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        391 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 450

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        451 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 510

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |   |||| |||   ||   ||   ||| || || |||||||||
Db        511 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 570

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |   |
Db        571 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 630

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||
Db        631 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 690

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||||| | ||||||||| || ||||
Db        691 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 750

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        751 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 810

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              |  |||     |||||||||||| |||||||||||||||||||||||| | |||||||
Db        811 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 870

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        871 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 930

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   ||||||||  ||||||||||||||||||
Db        931 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 990

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| |||||||||||| |  || |||||||  ||| |
Db        991 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1050

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||||||||||||||||| ||| | || |||||||||||||||||||||||||||||
Db       1051 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 1110

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||   ||||| || | |||| ||||||
Db       1111 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 1170

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||||||||||  ||    ||    || ||
Db       1171 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 1230

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  ||||||||| ||| | |||||||||||||||||||||||       ||||||||
Db       1231 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1290

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db       1291 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1350

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1351 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1410

Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1411 TTCAAGTGTGGGTTT 1425

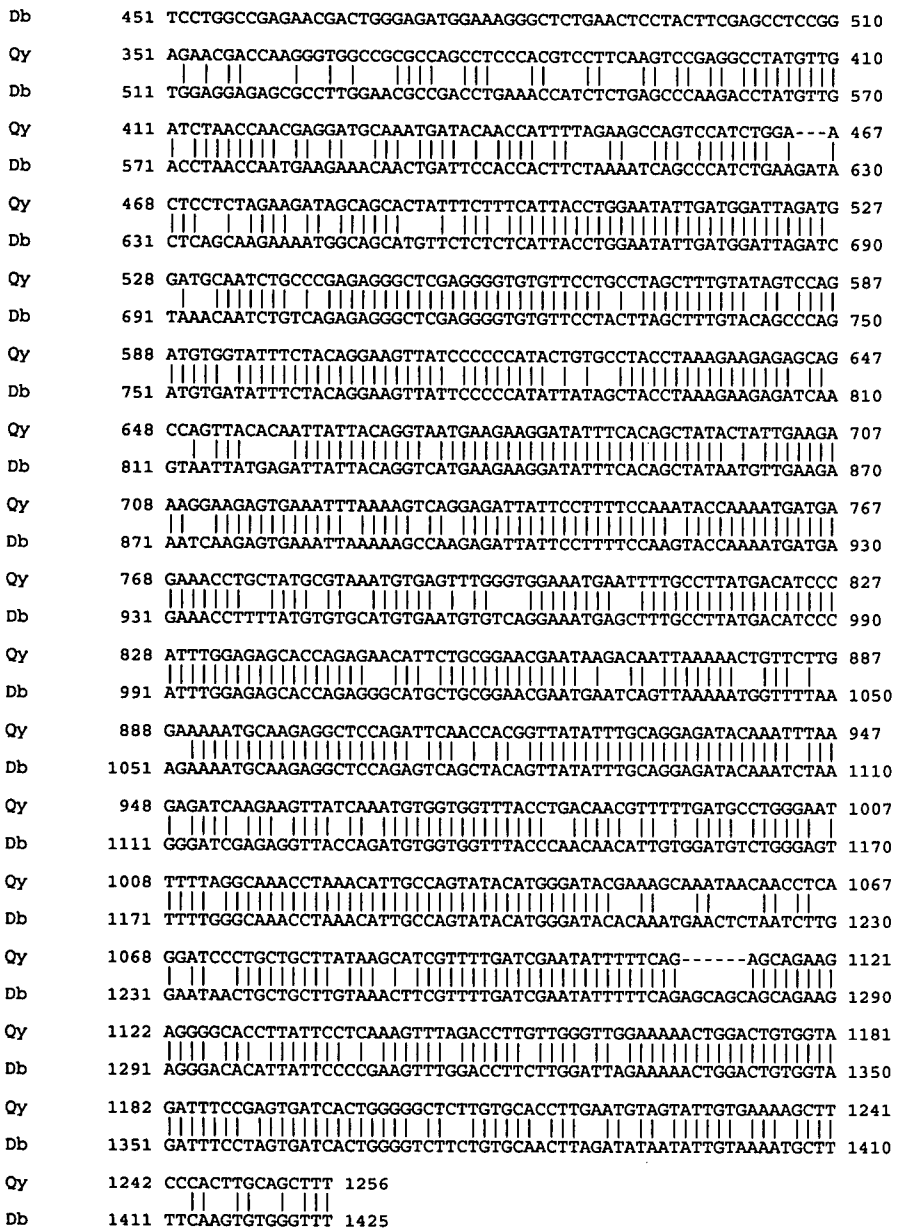

RESULT 33
US-09-710-280-2045
; Sequence 2045, Application US/09710280
; GENERAL INFORMATION:
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Gearing, David P.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2007-001
;  CURRENT APPLICATION NUMBER: US/09/710,280
;  CURRENT FILING DATE:  2000-11-10
;  PRIOR APPLICATION NUMBER:  60/164,537
;  PRIOR FILING DATE:  1999-11-10
;  NUMBER OF SEQ ID NOS:  2763
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2045
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-710-280-2045

  Query Match             50.9%;  Score 668.2;  DB 31;  Length 2262;
  Best Local Similarity   77.0%;  Pred. No. 1.1e-208;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||   |  |  |  || |  | | ||  |||    ||||| | |||||
Db        331 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 390

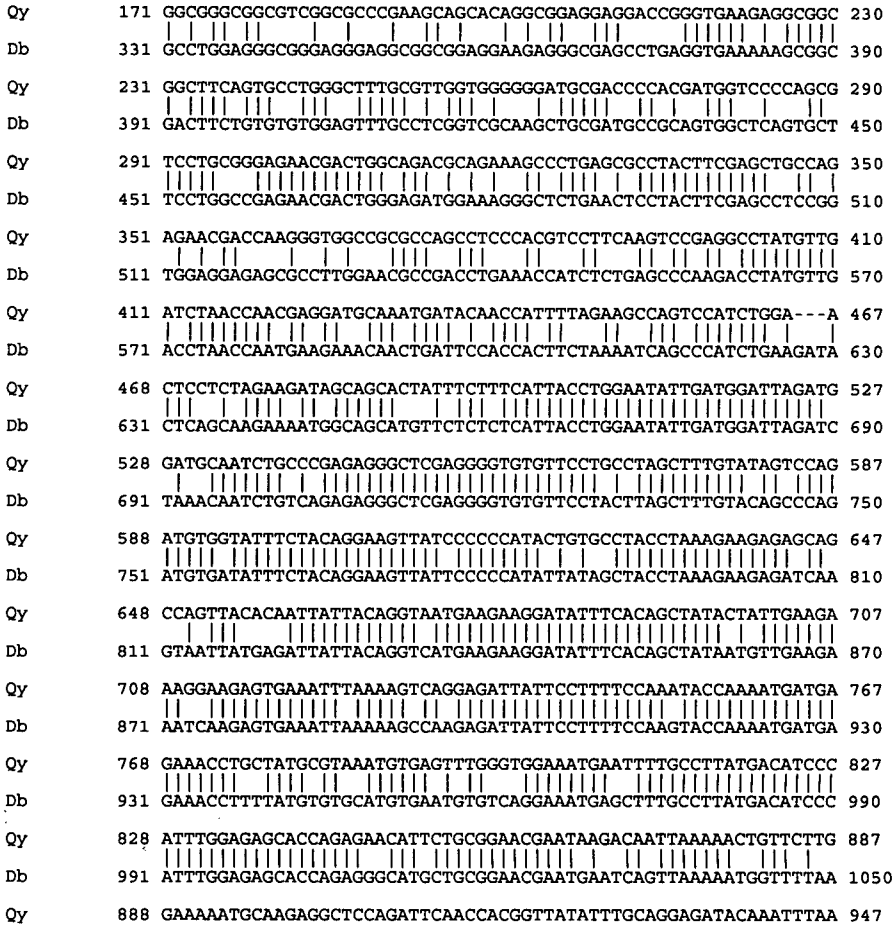
Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        391 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 450

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
```



RESULT 34
US-09-710-282-3926
; Sequence 3926, Application US/09710282
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Kingsbury, Gillian A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2004-001
; CURRENT APPLICATION NUMBER: US/09/710,282
; CURRENT FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 60/164,256
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4662
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3926
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2262)
; OTHER INFORMATION: n = A,T,C or G
US-09-710-282-3926

Query Match 50.9%; Score 668.2; DB 31; Length 2262;
Best Local Similarity 77.0%; Pred. No. 1.1e-208;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;



```
Db       1051 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 1110

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db       1111 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 1170

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db       1171 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 1230

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db       1231 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1290

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       1291 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1350

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1351 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1410

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1411 TTCAAGTGTGGGTTT 1425


RESULT 35
US-09-726-802-2141
; Sequence 2141, Application US/09726802
; GENERAL INFORMATION:
;  APPLICANT:  Acton, Susan L.
;  APPLICANT:  Falb, Dean A.
;  APPLICANT:  .
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2025-001
;  CURRENT APPLICATION NUMBER: US/09/726,802
;  CURRENT FILING DATE:  2000-11-30
;  PRIOR APPLICATION NUMBER:  60/168,012
;  PRIOR FILING DATE:  1999-11-30
;  NUMBER OF SEQ ID NOS:  2872
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2141
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-726-802-2141

  Query Match             50.9%;  Score 668.2;  DB 31;  Length 2262;
  Best Local Similarity   77.0%;  Pred. No. 1.1e-208;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db        331 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 390

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        391 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 450

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        451 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 510

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        511 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 570

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        571 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 630

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        631 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 690

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        691 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 750

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        751 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 810

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        811 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 870

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        871 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 930

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        931 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 990

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        991 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1050

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db       1051 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 1110

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db       1111 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 1170

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db       1171 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 1230

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db       1231 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1290

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       1291 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1350

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1351 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1410

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1411 TTCAAGTGTGGGTTT 1425


RESULT 36
US-09-981-353-88
; Sequence 88, Application US/09981353
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE: PA-0038 US
;  CURRENT APPLICATION NUMBER: US/09/981,353
;  CURRENT FILING DATE:  2001-10-11
;  NUMBER OF SEQ ID NOS: 194
```

```
;  SOFTWARE: PERL Program
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 232992.1
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

  Query Match           50.9%;  Score 668.2;  DB 38;  Length 3152;
  Best Local Similarity 77.0%;  Pred. No. 1.5e-208;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     ||||| | |||||
Db       1268 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 1327

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db       1328 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 1387

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||   || |
Db       1388 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 1447

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   |||| |||   ||   ||  || || || ||||||||||
Db       1448 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 1507

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| | |
Db       1508 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 1567

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||
Db       1568 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 1627

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||||  | |||||||||||||||||||||||||||| | ||||||||| || ||||
Db       1628 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 1687

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db       1688 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 1747

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db       1748 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 1807

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db       1808 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 1867

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   |||||||| ||||||||||||||||||
Db       1868 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 1927

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||||  |  || |||||||  ||| |
Db       1928 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1987

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||||| |||
Db       1988 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 2047

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  ||||| || | |||| |||||| |
Db       2048 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 2107

Qy        1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
               |||| |||||||||||||||||||||||||||||||||||||||||  ||   ||  || ||
Db        2108 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 2167

Qy        1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
               | ||  ||||||||| ||| | |||||||||||||||||||||||      ||||||||
Db        2168 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 2227

Qy        1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
               |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||||
Db        2228 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 2287

Qy        1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
               ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db        2288 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 2347

Qy        1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db        2348 TTCAAGTGTGGGTTT 2362


RESULT 37
US-60-239-841-88
; Sequence 88, Application US/60239841
; GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION:  GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE:  PA-0038 P
;  CURRENT APPLICATION NUMBER: US/60/239,841
;  CURRENT FILING DATE:  2000-10-11
;  NUMBER OF SEQ ID NOS:  194
;  SOFTWARE:  PERL Program
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 232992.1
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-60-239-841-88

  Query Match           50.9%;  Score 668.2;  DB 77;  Length 3152;
  Best Local Similarity 77.0%;  Pred. No. 1.5e-208;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     ||||| | |||||
Db       1268 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 1327

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db       1328 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 1387

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||   || |
Db       1388 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 1447

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   |||| |||   ||   ||  || || || ||||||||||
Db       1448 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 1507

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| | |
Db       1508 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 1567

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||
Db       1568 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 1627
```

```
Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db      1628 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 1687

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db      1688 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 1747

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db      1748 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 1807

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db      1808 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 1867

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db      1868 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 1927

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db      1928 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1987

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db      1988 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 2047

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db      2048 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 2107

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db      2108 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 2167

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db      2168 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 2227

Qy      1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db      2228 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 2287

Qy      1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db      2288 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 2347

Qy      1242 CCCACTTGCAGCTTT 1256
Db      2348 TTCAAGTGTGGGTTT 2362

RESULT 38
PCT-US00-05883-170
; Sequence 170, Application PC/TUS0005883
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT: Steve Ruben
;  TITLE OF INVENTION: Human Colon Cancer Associated Gene Sequences and Polypeptides
;  FILE REFERENCE: PA102PCT
;  CURRENT APPLICATION NUMBER: PCT/US00/05883
;  CURRENT FILING DATE: 2000-03-08
;  EARLIER APPLICATION NUMBER: 60/124,270
;  EARLIER FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
PCT-US00-05883-170

  Query Match             50.8%;  Score 667;  DB 1;  Length 1296;
  Best Local Similarity   76.7%;  Pred. No. 1.5e-208;
  Matches  840;  Conservative    3;  Mismatches  243;  Indels    9;  Gaps    2;

Qy       171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db        40 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 99

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db       100 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 159

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db       160 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 219

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db       220 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 279

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db       280 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 339

Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db       340 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 399

Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db       400 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 459

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db       460 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 519

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db       520 GTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACAGCTATAATGTTGAAGA 579

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db       580 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 639

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db       640 GAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 699

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db       700 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 759

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db       760 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 819

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db       820 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 879

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db       880 TTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 939

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
```

```
Db        940 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 999

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       1000 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1059

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1060 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1119

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1120 TTCAAGTGTGGGTTT 1134


RESULT 39
US-09-925-299-170
; Sequence 170, Application US/09925299
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

  Query Match           50.8%;  Score 667;  DB 36;  Length 1296;
  Best Local Similarity   76.7%;  Pred. No. 1.5e-208;
  Matches  840;  Conservative    3;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         40 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 99

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        100 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 159

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        160 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 219

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        220 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 279

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        280 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 339

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        340 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 399

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        400 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 459

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        460 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 519

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        520 GTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACAGCTATAATGTTGAAGA 579

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        580 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 639

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        640 GAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 699

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        700 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 759

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        760 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 819

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        820 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 879

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        880 TTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 939

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        940 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 999

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       1000 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1059

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1060 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1119

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1120 TTCAAGTGTGGGTTT 1134


RESULT 40
US-10-170-235-27345
; Sequence 27345, Application US/10170235
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig
;  TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
;  TITLE OF INVENTION:  TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
;  FILE REFERENCE: CL001380
;  CURRENT APPLICATION NUMBER: US/10/170,235
;  CURRENT FILING DATE:  2003-03-17
;  NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 27345
;   LENGTH: 1919
;   TYPE: DNA
;   ORGANISM: HUMAN
US-10-170-235-27345

  Query Match           50.8%;  Score 666.6;  DB 42;  Length 1919;
  Best Local Similarity   76.9%;  Pred. No. 3.1e-208;
  Matches  842;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    2;
```

```
Qy       171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db        34 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 93

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db        94 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 153

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db       154 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 213

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db       214 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 273

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db       274 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 333

Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db       334 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 393

Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db       394 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 453

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db       454 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 513

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db       514 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 573

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db       574 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 633

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db       634 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 693

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db       694 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 753

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db       754 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 813

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db       814 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 873

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db       874 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 933

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db       934 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 993

Qy      1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db       994 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1053

Qy      1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db      1054 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1113

Qy      1242 CCCACTTGCAGCTTT 1256
```

```
Db      1114 TTCAAGTGTGGGTTT 1128
```

RESULT 41
PCT-US04-24424-921
; Sequence 921, Application PC/TUS0424424
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PCT
; CURRENT APPLICATION NUMBER: PCT/US04/24424
; CURRENT FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US 60/490,890
; PRIOR FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2786
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-24424-921

Query Match 50.8%; Score 666.6; DB 3; Length 1940;
Best Local Similarity 76.9%; Pred. No. 3.1e-208;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;

```
Qy       171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db        42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db       102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db       162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db       222 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db       282 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 341

Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db       342 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 401

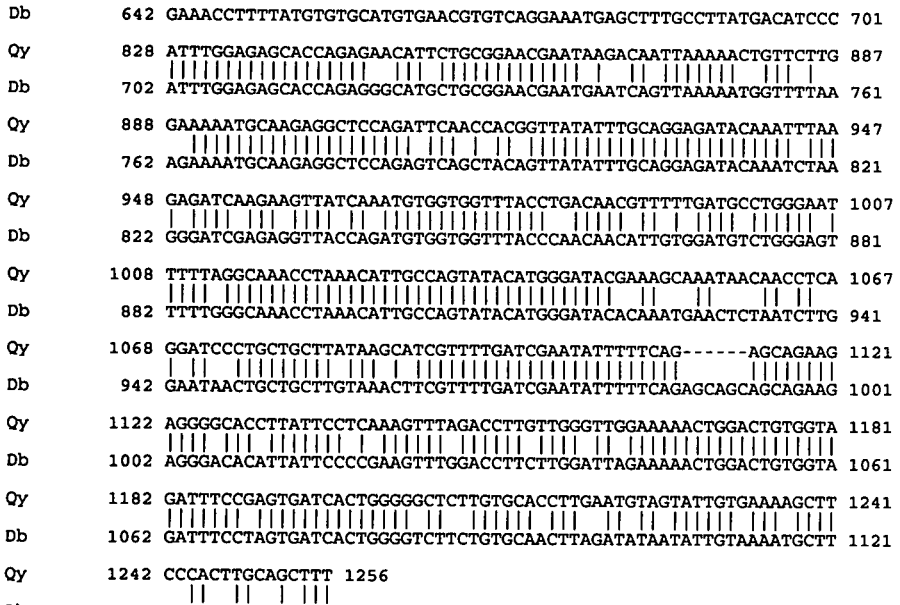
Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db       402 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 461

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db       462 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 521

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db       522 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 581

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db       582 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 641

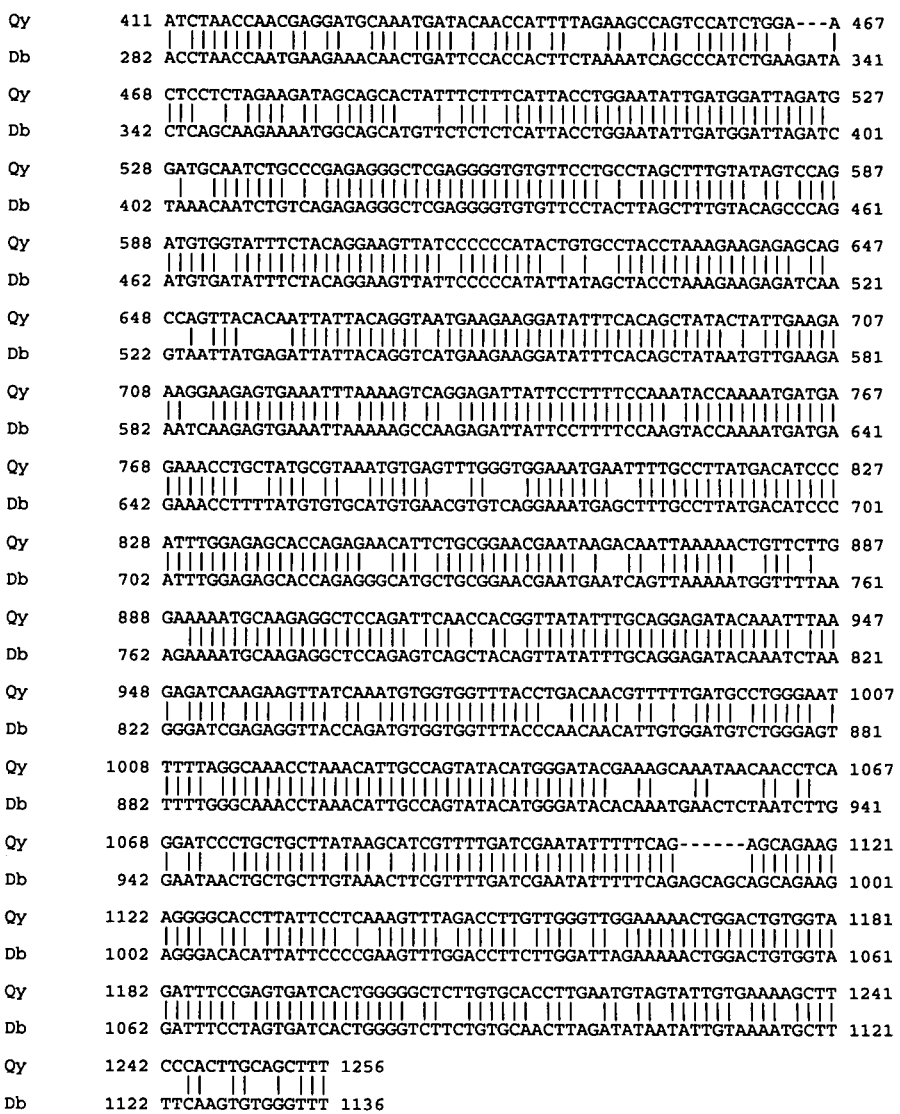
Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
```



RESULT 42
US-10-783-271-28
; Sequence 28, Application US/10783271
; GENERAL INFORMATION:
; APPLICANT: Veridex, LLC
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
; FILE REFERENCE: VDX-5003 USNP
; CURRENT APPLICATION NUMBER: US/10/783,271
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: human
US-10-783-271-28

Query Match 50.8%; Score 666.6; DB 61; Length 1940;
Best Local Similarity 76.9%; Pred. No. 3.1e-208;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;





RESULT 43
US-60-490-890-921
; Sequence 921, Application US/60490890
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PSP
; CURRENT APPLICATION NUMBER: US/60/490,890

```
;  CURRENT FILING DATE:  2003-07-29
;  NUMBER OF SEQ ID NOS: 2779
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-490-890-921

  Query Match            50.8%;  Score 666.6;  DB 80;  Length 1940;
  Best Local Similarity  76.9%;  Pred. No. 3.1e-208;
  Matches  842;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |    |  || |   | | ||  |||      |||||| | |||||
Db         42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| ||| |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |  |||| |||   ||   ||   || || || ||||||||||
Db        222 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        282 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 341

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        342 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 401

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||||| | |||||||||||||||||||||||||| | |||||||||| || ||||
Db        402 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 461

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        462 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 521

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||     |||||||||||| |||||||||||||||||||||||| | |||||||
Db        522 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 581

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||||| ||||||||||||
Db        582 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 641

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  ||||||   ||   ||||||||  ||||||||||||||||||
Db        642 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 701

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||||  ||| ||||||||||||| |   || |||||||  ||| |
Db        702 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 761

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||||| |||
Db        762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || |||||||||||||||  ||||| || |  |||| ||||||  |
Db        822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||  ||   ||   || ||
Db        882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||| | |||||||||||||||||||||||      ||||||||
Db        942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||||
Db       1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1122 TTCAAGTGTGGGTTT 1136


RESULT 44
US-60-568-073-582
; Sequence 582, Application US/60568073
; GENERAL INFORMATION:
;  APPLICANT: Huang, Fei
;  APPLICANT:  Han, Xia
;  APPLICANT:  Shaw, Peter
;  APPLICANT:  Clark, Edwin
;  TITLE OF INVENTION: COLON CANCER BIOMARKERS
;  FILE REFERENCE: 10205 PSP
;  CURRENT APPLICATION NUMBER: US/60/568,073
;  CURRENT FILING DATE:  2004-05-04
;  NUMBER OF SEQ ID NOS: 1305
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 582
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-568-073-582

  Query Match            50.8%;  Score 666.6;  DB 81;  Length 1940;
  Best Local Similarity  76.9%;  Pred. No. 3.1e-208;
  Matches  842;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |    |  || |   | | ||  |||      |||||| | |||||
Db         42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| ||| |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |  |||| |||   ||   ||   || || || ||||||||||
Db        222 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        282 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 341

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        342 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 401

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||||| | |||||||||||||||||||||||||| | |||||||||| || ||||
Db        402 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 461

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        462 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 521
```

```
Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db       522 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 581

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db       582 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 641

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db       642 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 701

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db       702 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 761

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db       762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db       822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db       882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db       942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy      1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db      1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy      1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db      1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

Qy      1242 CCCACTTGCAGCTTT 1256
Db      1122 TTCAAGTGTGGGTTT 1136
```

RESULT 45
US-60-671-501-156
; Sequence 156, Application US/60671501
; GENERAL INFORMATION:
; APPLICANT: Wim Van Criekinge
; APPLICANT: Josef Straub
; TITLE OF INVENTION: METHYLATION MARKERS FOR DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF CANCERS
; FILE REFERENCE: 006791.00001
; CURRENT APPLICATION NUMBER: US/60/671,501
; CURRENT FILING DATE: 2005-04-15
; NUMBER OF SEQ ID NOS: 420
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 156
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-671-501-156

Query Match 50.8%; Score 666.6; DB 82; Length 1940;
Best Local Similarity 76.9%; Pred. No. 3.1e-208;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;

```
Qy       171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db        42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db       102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db       162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db       222 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db       282 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 341

Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db       342 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 401

Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db       402 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 461

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db       462 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 521

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db       522 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 581

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db       582 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 641

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db       642 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 701

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db       702 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 761

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db       762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db       822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db       882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db       942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy      1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db      1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy      1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db      1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

Qy      1242 CCCACTTGCAGCTTT 1256
Db      1122 TTCAAGTGTGGGTTT 1136
```

RESULT 46

```
US-60-685-372-523
; Sequence 523, Application US/60685372
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John V
;  APPLICANT:  Bradley, Walter E
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall D
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul V
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: GeneMap of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 059908-5005-PR
;  CURRENT APPLICATION NUMBER: US/60/685,372
;  CURRENT FILING DATE:  2005-05-31
;  NUMBER OF SEQ ID NOS: 2738
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 523
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-685-372-523

  Query Match             50.8%;  Score 666.6;  DB 82;  Length 1940;
  Best Local Similarity   76.9%;  Pred. No. 3.1e-208;
  Matches  842;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||      |||||| | |||||
Db         42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |    ||
Db        102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||  |  | |  |||| ||| ||  ||  || || || |||||||||
Db        222 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |   |
Db        282 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 341

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||
Db        342 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 401

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||||| | | |||||||||| || ||||
Db        402 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 461

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||||| |||||||| | |  ||||||||||||||||| ||
Db        462 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 521

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              |  |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db        522 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 581

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        582 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 641

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  ||||||  ||  |||||||| ||||||||||||||||||||
Db        642 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 701

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| |||||||||||| |  || |||||||  ||| |
Db        702 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 761

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  |||||  || | |||| ||||||
Db        822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||||||||  ||    ||    || ||
Db        882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  ||||||||| ||| | |||||||||||||||||||||||      ||||||||
Db        942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db       1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1122 TTCAAGTGTGGGTTT 1136

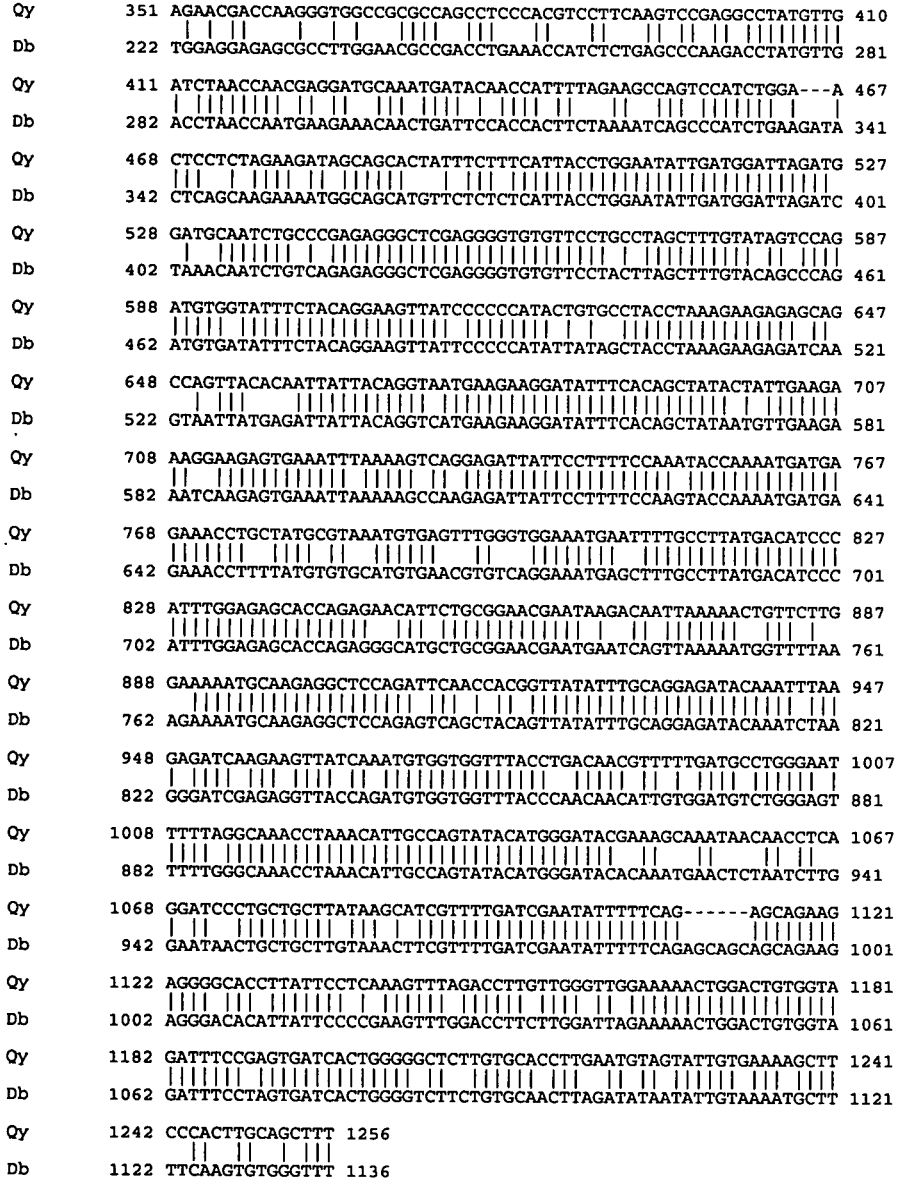

RESULT 47
US-60-704-740-28
; Sequence 28, Application US/60704740
; GENERAL INFORMATION:
;  APPLICANT: Wang, Yixin
;  APPLICANT:  Zhang, Yi
;  APPLICANT:  Atkins, David
;  APPLICANT:  Sieuwerts, Anieta M.
;  APPLICANT:  Smid, Marcel
;  APPLICANT:  Klijn, Jan G.M.
;  APPLICANT:  Martens, John W.M.
;  APPLICANT:  Foekens, John A.
;  TITLE OF INVENTION: Predicting Bone Relapse of Breast Cancer
;  FILE REFERENCE: VDX 5026 USPSP
;  CURRENT APPLICATION NUMBER: US/60/704,740
;  CURRENT FILING DATE:  2005-08-02
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: human
US-60-704-740-28

  Query Match             50.8%;  Score 666.6;  DB 83;  Length 1940;
  Best Local Similarity   76.9%;  Pred. No. 3.1e-208;
  Matches  842;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||      |||||| | |||||
Db         42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

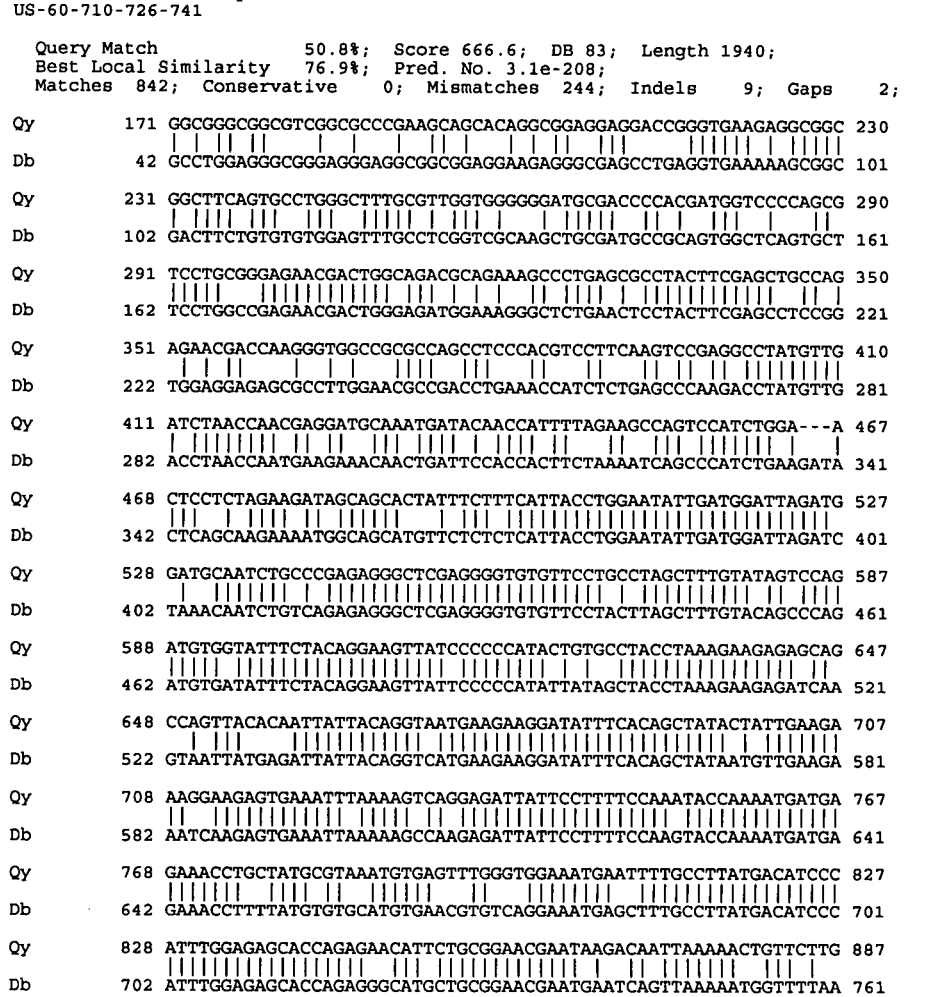
Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |    ||
Db        102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221
```



RESULT 48
US-60-710-726-741
; Sequence 741, Application US/60710726
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John V
; APPLICANT: Bradley, Walter E
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Serre, Valerie
; APPLICANT: Eerdewegh, Paul V
; APPLICANT: Little, Randall D
; APPLICANT: Keith , Tim
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: GeneMap of the human genes associated with Crohn's disease
; FILE REFERENCE: 059908-5009-PR
; CURRENT APPLICATION NUMBER: US/60/710,726
; CURRENT FILING DATE: 2005-08-24
; NUMBER OF SEQ ID NOS: 20166
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 741
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homosapiens
US-60-710-726-741

Query Match 50.8%; Score 666.6; DB 83; Length 1940;
Best Local Similarity 76.9%; Pred. No. 3.1e-208;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;



```
Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || |||||||||||||||  ||||| || | |||| ||||||  |
Db        822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              ||||  ||||||||||||||||||||||||||||||||||   ||    ||    || ||
Db        882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||| ||| | ||||||||||||||||||||||||      ||||||||
Db        942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db       1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1122 TTCAAGTGTGGGTTT 1136
```

RESULT 49
US-09-585-944-2184
; Sequence 2184, Application US/09585944
; GENERAL INFORMATION:
; APPLICANT: Hodgson, David M.
; APPLICANT: Lincoln, Stephen E.
; APPLICANT: Russo, Frank D.
; APPLICANT: Spiro, Peter A.
; APPLICANT: Banville, Steve C.
; APPLICANT: Bratcher, Shawn R.
; APPLICANT: Dufour, Gerard E.
; APPLICANT: Cohen, Howard J.
; APPLICANT: Rosen, Bruce
; APPLICANT: Chalup, Michael S.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Jones, Anissa L.
; APPLICANT: Yu, Jimmy Y.
; APPLICANT: Greenawalt, Lila B.
; APPLICANT: Panzer, Scott R.
; APPLICANT: Roseberry, Ann M.
; APPLICANT: Wright, Rachel J.
; APPLICANT: Daniels, Susan E.
; TITLE OF INVENTION: SECRETORY MOLECULES
; FILE REFERENCE: PT-0064 US
; CURRENT APPLICATION NUMBER: US/09/585,944
; CURRENT FILING DATE: 2000-06-02
; EARLIER APPLICATION NUMBER: US 60/137,412
; EARLIER FILING DATE: 1999-06-03
; EARLIER APPLICATION NUMBER: US 60/147,542
; EARLIER FILING DATE: 1999-08-05
; EARLIER APPLICATION NUMBER: US 60/147,501
; EARLIER FILING DATE: 1999-08-05
; EARLIER APPLICATION NUMBER: US 60/147,500
; EARLIER FILING DATE: 1999-08-05
; NUMBER OF SEQ ID NOS: 2515
; SOFTWARE: PERL Program
; SEQ ID NO 2184
; LENGTH: 3080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 232992.7

; FEATURE:
; NAME/KEY: unsure
; LOCATION: 125, 1098
; OTHER INFORMATION: a, t, c, g, or other
US-09-585-944-2184

Query Match 50.7%; Score 665.6; DB 27; Length 3080;
Best Local Similarity 76.8%; Pred. No. 1.1e-207;
Matches 842; Conservative 0; Mismatches 244; Indels 10; Gaps 2;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||   |  |  |  || | | | || |||    ||||||| | ||||||
Db       1195 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 1254

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db       1255 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 1314

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db       1315 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 1374

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   | | |  |||| |||   ||   ||  || || || |||||||||
Db       1375 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 1434

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTC 470
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |  |
Db       1435 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 1494

Qy        471 CT----CTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGAT 526
              ||    | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db       1495 CTCCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGAT 1554

Qy        527 GGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCA 586
              |  |||||||  | |||||||||||||||||||||||||| | |||||||||| || |||
Db       1555 CTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCA 1614

Qy        587 GATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCA 646
              |||||| |||||||||||||||||||| |||||||| | |  |||||||||||||||| ||
Db       1615 GATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCA 1674

Qy        647 GCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAG 706
              |  |||    |||||||||||| ||||||||||||||||||||||||||| | |||||||
Db       1675 AGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAG 1734

Qy        707 AAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATG 766
              |||  |||||||||||| ||||| || |||||||||||||||||||||||||||||||||
Db       1735 AAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATG 1794

Qy        767 AGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 826
              ||||||||  |||| ||  |||||| | ||   ||||||||  |||||||||||||||||
Db       1795 AGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCC 1854

Qy        827 CATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTT 886
              |||||||||||||||||  ||| ||||||||||||  | ||  |||||||  ||| |
Db       1855 CATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTA 1914

Qy        887 GGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTA 946
              ||||||||||||||||||| ||| | || ||||||||||||||||||||||||||||| ||
Db       1915 AAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTA 1974

Qy        947 AGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAA 1006
              || |||| ||| |||| || ||||||||||||||  |||||  || | |||| ||||||
Db       1975 AGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAG 2034

Qy       1007 TTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTC 1066
              ||||| ||||||||||||||||||||||||||||||||||||  ||    ||    || ||
Db       2035 TTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTT 2094

Qy       1067 AGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAA 1120
              | ||  |||||||| ||| | ||||||||||||||||||||||||      |||||||
Db       2095 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 2154
```

```
Qy       1121 GAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGT 1180
             ||||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db       2155 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 2214

Qy       1181 AGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCT 1240
             |||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| |||
Db       2215 AGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCT 2274

Qy       1241 TCCCACTTGCAGCTTT 1256
             |  ||  ||  | |||
Db       2275 TTTCAAGTGTGGGTTT 2290


RESULT 50
US-60-172-360-956/c
; Sequence 956, Application US/60172360
; GENERAL INFORMATION:
;  APPLICANT: Morris, MacDonald
;  APPLICANT: Lal, Preeti
;  APPLICANT: Diep, Dinh
;  TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
;  TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymorp
;  FILE REFERENCE: GX-0007 P
;  CURRENT APPLICATION NUMBER: US/60/172,360
;  CURRENT FILING DATE: 1999-12-16
;  NUMBER OF SEQ ID NOS: 29838
;  SOFTWARE: PERL Program
; SEQ ID NO 956
;   LENGTH: 3082
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 232992.7c
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 1983, 2958
;   OTHER INFORMATION: a, t, c, g, or other
US-60-172-360-956

  Query Match             50.7%;  Score 665.6;  DB 75;  Length 3082;
  Best Local Similarity   76.8%;  Pred. No. 1.1e-207;
  Matches  842;  Conservative    0;  Mismatches  244;  Indels   10;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     |||||| | |||||
Db       1886 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 1827

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db       1826 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 1767

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db       1766 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 1707

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |   |||| |||   ||   ||  || || || |||||||||
Db       1706 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 1647

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTC 470
              | ||||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |  |
Db       1646 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 1587

Qy        471 CT----CTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGAT 526
              ||    | |||| || ||||||    | ||| ||||||||||||||||||||||||||||
Db       1586 CTCCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGAT 1527

Qy        527 GGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCA 586
              |  ||||||| | ||||||||||||||||||||||| | |||||||||| || |||
Db       1526 CTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCA 1467
```

```
Qy        587 GATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCA 646
              |||||| |||||||||||||||||| |||||||| | |  ||||||||||||||||| ||
Db       1466 GATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCA 1407

Qy        647 GCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAG 706
              | |||     |||||||||||| |||||||||||||||||||||||||| | ||||||
Db       1406 AGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAG 1347

Qy        707 AAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATG 766
              |||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||
Db       1346 AAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATG 1287

Qy        767 AGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 826
              |||||||  |||| ||  |||||| | ||  ||||||||  ||||||||||||||||||
Db       1286 AGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCC 1227

Qy        827 CATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTT 886
              |||||||||||||||||||  ||| |||||||||||||| |  || |||||||  ||| |
Db       1226 CATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTA 1167

Qy        887 GGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTA 946
              |||||||||||||||||||||| ||| | || |||||||||||||||||||||||||| ||
Db       1166 AAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTA 1107

Qy        947 AGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAA 1006
              || |||| ||| |||| || ||||||||||||||  ||||| || | |||| ||||||
Db       1106 AGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAG 1047

Qy       1007 TTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTC 1066
              ||||| ||||||||||||||||||||||||||||||||||  ||    ||    || ||
Db       1046 TTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTT 987

Qy       1067 AGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAA 1120
              | ||  |||||||||| ||| | |||||||||||||||||||||||      |||||||
Db        986 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 927

Qy       1121 GAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGT 1180
              ||||| ||| |||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db        926 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 867

Qy       1181 AGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCT 1240
              |||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| |||
Db        866 AGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCT 807

Qy       1241 TCCCACTTGCAGCTTT 1256
              |  ||  ||  | |||
Db        806 TTTCAAGTGTGGGTTT 791


RESULT 51
US-09-629-469A-13209
; Sequence 13209, Application US/09629469A
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/09/629,469A
;  CURRENT FILING DATE:  2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
```

```
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
;   LENGTH: 1898
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1086)
US-09-629-469A-13209

  Query Match            50.7%;  Score 665;  DB 28;  Length 1898;
  Best Local Similarity  76.8%;  Pred. No. 1e-207;
  Matches  841;  Conservative    0;  Mismatches  245;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | ||| ||    |  |   |  || |   | | ||  |||     |||||| | |||||
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| ||| |||  ||||| | ||| |   | |||||  || ||  |||  |   ||
Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||  |||||||||||| ||| | | |  || |||| | ||| ||||||||  || |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   ||||  |||   ||   ||  || || || |||||||||
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | ||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||| |   |
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||
Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||||| | ||||||||||||||||||||| |||| | |||||||||| || ||||
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| ||||||||||||||||||||||| | |||||||
Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        557 AATCAAGAGTGAAATTAAAAAGCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              ||||||  ||||  ||  ||||||   ||   |||||||  ||||||||||||||||||
Db        617 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 676

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||  ||| ||||||||||||| |  || |||||||  ||| |
Db        677 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 736

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        737 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 796

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  ||||| || | |||| |||||| |
Db        797 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 856

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||  ||   ||    || ||
Db        857 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 916

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||| | |||||||||||||||||||||||      ||||||||
Db        917 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 976

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| |||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db        977 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1036

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1037 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1096

Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1097 TTCAAGTGTGGGTTT 1111


RESULT 52
US-10-917-503-13209
; Sequence 13209, Application US/10917503
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
;   LENGTH: 1898
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1086)
US-10-917-503-13209
```

Query Match 50.7%; Score 665; DB 63; Length 1898;
Best Local Similarity 76.8%; Pred. No. 1e-207;
Matches 841; Conservative 0; Mismatches 245; Indels 9; Gaps 2;

```
Qy      171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db       17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC  76

Qy      231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db       77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy      291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db      137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy      351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db      197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy      411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db      257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy      468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db      317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy      528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db      377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy      588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db      437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy      648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db      497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy      708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db      557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy      768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db      617 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 676

Qy      828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db      677 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 736

Qy      888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db      737 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 796

Qy      948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db      797 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 856

Qy     1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db      857 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 916

Qy     1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db      917 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 976

Qy     1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db      977 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1036
```

```
Qy     1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db     1037 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1096

Qy     1242 CCCACTTGCAGCTTT 1256
Db     1097 TTCAAGTGTGGGTTT 1111
```

RESULT 53
US-09-705-256A-7708
; Sequence 7708, Application US/09705256A
; GENERAL INFORMATION:
; APPLICANT: Ma, Xiao-Jun; Dotson, Stanton B.: Monsanto Company
; TITLE OF INVENTION: Tumor Associated Molecules (TAMs): Targets for diagnosis, treatm
; TITLE OF INVENTION: and prevention of cancer
; FILE REFERENCE: 3214
; CURRENT APPLICATION NUMBER: US/09/705,256A
; CURRENT FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: US 60/164,285
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 8259
; SEQ ID NO 7708
; LENGTH: 1930
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-705-256A-7708

Query Match 50.2%; Score 658.8; DB 31; Length 1930;
Best Local Similarity 77.0%; Pred. No. 1.2e-205;
Matches 844; Conservative 0; Mismatches 242; Indels 10; Gaps 3;

```
Qy      171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db       42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy      231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db      102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy      291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db      162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy      351 AGAACGACCAAGGGTGGCCGCGCC-AGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTT 409
Db      222 TGGAGGAGAGCGCCTTGGAACGCCGAGCCTGAAACCATCTCTGAGCCCAAGACCTATGTT 281

Qy      410 GATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA--- 466
Db      282 GACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGAT 341

Qy      467 ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGAT 526
Db      342 ACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGAT 401

Qy      527 GGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCA 586
Db      402 CTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCA 461

Qy      587 GATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCA 646
Db      462 GATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCA 521

Qy      647 GCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAG 706
Db      522 AGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAG 581

Qy      707 AAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATG 766
Db      582 AAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATG 641

Qy      767 AGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 826
```

```
Db        642 AGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCC 701

Qy        827 CATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTT 886

Db        702 CATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTA 761

Qy        887 GGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTA 946

Db        762 AAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTA 821

Qy        947 AGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAA 1006

Db        822 AGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAG 881

Qy       1007 TTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTC 1066

Db        882 TTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTT 941

Qy       1067 AGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAA 1120

Db        942 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 1001

Qy       1121 GAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGT 1180

Db       1002 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 1061

Qy       1181 AGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCT 1240

Db       1062 AGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCT 1121

Qy       1241 TCCCACTTGCAGCTTT 1256

Db       1122 TTTCAAGTGTGGGTTT 1137


RESULT 54
US-60-164-285-7708
; Sequence 7708, Application US/60164285
; GENERAL INFORMATION:
;  APPLICANT: Ma, Xiao-Jun
;  TITLE OF INVENTION: Tumor Associated Molecules (TAMs): Targets for diagnosis, treatme
;  FILE REFERENCE: 3214
;  CURRENT APPLICATION NUMBER: US/60/164,285
;  CURRENT FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 8259
; SEQ ID NO 7708
;   LENGTH: 1930
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-164-285-7708

  Query Match             50.2%;  Score 658.8;  DB 75;  Length 1930;
  Best Local Similarity   77.0%;  Pred. No. 1.2e-205;
  Matches  844;  Conservative    0;  Mismatches  242;  Indels   10;  Gaps    3;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy        351 AGAACGACCAAGGGTGGCCGCGCC-AGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTT 409

Db        222 TGGAGGAGAGCGCCTTGGAACGCCGAGCCTGAAACCATCTCTGAGCCCAAGACCTATGTT 281

Qy        410 GATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA--- 466

Db        282 GACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGAT 341

Qy        467 ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGAT 526

Db        342 ACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGAT 401

Qy        527 GGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCA 586

Db        402 CTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCA 461

Qy        587 GATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCA 646

Db        462 GATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCA 521

Qy        647 GCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAG 706

Db        522 AGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAG 581

Qy        707 AAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATG 766

Db        582 AAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATG 641

Qy        767 AGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 826

Db        642 AGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCC 701

Qy        827 CATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTT 886

Db        702 CATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTA 761

Qy        887 GGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTA 946

Db        762 AAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTA 821

Qy        947 AGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAA 1006

Db        822 AGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAG 881

Qy       1007 TTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTC 1066

Db        882 TTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTT 941

Qy       1067 AGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAA 1120

Db        942 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 1001

Qy       1121 GAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGT 1180

Db       1002 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 1061

Qy       1181 AGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCT 1240

Db       1062 AGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCT 1121

Qy       1241 TCCCACTTGCAGCTTT 1256

Db       1122 TTTCAAGTGTGGGTTT 1137


RESULT 55
US-10-488-619-2154
; Sequence 2154, Application US/10488619
; GENERAL INFORMATION:
;  APPLICANT: Greenlee, Winner and Sullivan, P.C.
;  TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
;  TITLE OF INVENTION:  Physiological Conditions, And Genotyping Arrays
;  FILE REFERENCE: 98-01 WO
;  CURRENT APPLICATION NUMBER: US/10/488,619
;  CURRENT FILING DATE:  2004-03-01
;  NUMBER OF SEQ ID NOS: 3040
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2154
```

```
;   LENGTH: 602
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-488-619-2154

  Query Match             45.8%;  Score 600.4;  DB 51;  Length 602;
  Best Local Similarity   99.8%;  Pred. No. 9.5e-187;
  Matches  601;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        357 ACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAA 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAA 60

Qy        417 CCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAG 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAG 120

Qy        477 AAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATC 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATC 180

Qy        537 TGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTAT 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTAT 240

Qy        597 TTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACA 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACA 300

Qy        657 CAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAG 360

Qy        717 TGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGC 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        361 TGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAGCCTGC 420

Qy        777 TATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGA 836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGA 480

Qy        837 GCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGC 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGC 540

Qy        897 AAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAG 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAG 600

Qy        957 AA 958
              ||
Db        601 AA 602


RESULT 56
US-60-443-566-1789/c
; Sequence 1789, Application US/60443566
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001447
;  CURRENT APPLICATION NUMBER: US/60/443,566
;  CURRENT FILING DATE:  2003-01-30
;  NUMBER OF SEQ ID NOS: 25102
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1789
;   LENGTH: 1849
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-443-566-1789

  Query Match             43.9%;  Score 575.6;  DB 79;  Length 1849;
  Best Local Similarity   76.1%;  Pred. No. 5.5e-178;
  Matches  714;  Conservative    6;  Mismatches  215;  Indels    3;  Gaps    1;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    | |  | || |  | | ||  |||     |||||| | |||||
Db        938 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 879

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        878 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 819

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | |  || |||: | ||||||||||||  || |
Db        818 TCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCCTACTTCGAGCCTCCGG 759

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |  ||||  |||   ||   ||   || || || |||||||||
Db        758 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 699

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||    ||  ||| ||||||| |   |
Db        698 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 639

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        638 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 579

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||| | |||||||||| || ||||
Db        578 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 519

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||||| |||||||| | |  ||||||||||||||||| ||:
Db        518 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAR 459

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||   |||||||||||| |||||||||||||||||||||||| | |||||||
Db        458 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 399

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        398 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 339

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |:|||| : ||   |||||||  ||||||||||||||||||
Db        338 GAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 279

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||  |  || |||||||  ||| |
Db        278 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTRA 219

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||:|||||||||||||| ||| | || |||||||||||||||||||||||||| |||
Db        218 AGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 159

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | ||||:||| |||| || ||||||||||||||  |||| || | |||| |||||| |
Db        158 GGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 99

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||||  ||    ||  || ||
Db         98 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 39

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATA 1105
              | ||  |||||||||| ||| | ||||||||||||||||
Db         38 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 1


RESULT 57
US-60-452-680-10422/c
; Sequence 10422, Application US/60452680
; GENERAL INFORMATION:
```

```
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  GRUPE, Andrew
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001450
;  CURRENT APPLICATION NUMBER: US/60/452,680
;  CURRENT FILING DATE:  2003-03-07
;  NUMBER OF SEQ ID NOS: 116213
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10422
;   LENGTH: 1849
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-452-680-10422

  Query Match             43.9%;  Score 575.6;  DB 79;  Length 1849;
  Best Local Similarity   76.1%;  Pred. No. 5.5e-178;
  Matches  714;  Conservative    6;  Mismatches  215;  Indels    3;  Gaps    1;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   | || |  | | ||  |||     ||||||| | |||||
Db        938 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 879

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |  | |||||  || |  |||  |  ||
Db        878 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 819

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||: | ||||||||||||  || |
Db        818 TCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCCTACTTCGAGCCTCCGG 759

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   |  | |   |||| |||   ||   ||  || || || |||||||||
Db        758 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 699

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |   |
Db        698 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 639

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        638 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 579

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | ||||||||||||||||||||||||||| | ||||||||||| || ||||
Db        578 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 519

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||||| ||:
Db        518 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAR 459

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||  |||||||||||| |||||||||||||||||||||||||||| | |||||||
Db        458 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 399

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        398 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 339

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |:|||| : ||   |||||||| ||||||||||||||||||
Db        338 GAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 279

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||  ||| |||||||||||||  || |||||||  ||| |
Db        278 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTRA 219

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||:|||||||||||| ||| | || ||||||||||||||||||||||| |||
Db        218 AGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 159

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | ||||:||| |||| || ||||||||||||||  ||||| || | |||| ||||||
Db        158 GGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 99

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||||  ||   ||   || ||
Db         98 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 39

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATA 1105
              | ||  |||||||||| ||| | ||||||||||||||||
Db         38 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 1


RESULT 58
US-60-455-444-3323/c
; Sequence 3323, Application US/60455444
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele
;  APPLICANT:  BEGOVICH, Ann
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001455
;  CURRENT APPLICATION NUMBER: US/60/455,444
;  CURRENT FILING DATE:  2003-03-18
;  NUMBER OF SEQ ID NOS: 50986
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3323
;   LENGTH: 1849
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-455-444-3323

  Query Match             43.9%;  Score 575.6;  DB 79;  Length 1849;
  Best Local Similarity   76.1%;  Pred. No. 5.5e-178;
  Matches  714;  Conservative    6;  Mismatches  215;  Indels    3;  Gaps    1;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   | || |  | | ||  |||     ||||||| | |||||
Db        938 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 879

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |  | |||||  || |  |||  |  ||
Db        878 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 819

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||: | ||||||||||||  || |
Db        818 TCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCCTACTTCGAGCCTCCGG 759

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   |  | |   |||| |||   ||   ||  || || || |||||||||
Db        758 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 699

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        698 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 639

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        638 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 579

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | ||||||||||||||||||||||||||| | ||||||||||| || ||||
Db        578 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 519

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||||| ||:
Db        518 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAR 459

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||  |||||||||||| |||||||||||||||||||||||||||| | |||||||
Db        458 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 399

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        398 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 339
```

```
Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |:|||| : ||   ||||||||  ||||||||||||||||||
Db        338 GAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 279

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||| |  || |||||||  ||| |
Db        278 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTRA 219

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||:|||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        218 AGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 159

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | ||||:||| |||| || ||||||||||||||  |||||  || | ||||  |||||| |
Db        158 GGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 99

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||  ||   ||    || ||
Db         98 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 39

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATA 1105
              | ||  |||||||| ||| | ||||||||||||||||
Db         38 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 1
```

RESULT 59
US-60-465-241-3323/c
; Sequence 3323, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3323
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-465-241-3323

Query Match 43.9%; Score 575.6; DB 79; Length 1849;
Best Local Similarity 76.1%; Pred. No. 5.5e-178;
Matches 714; Conservative 6; Mismatches 215; Indels 3; Gaps 1;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||      |||||| | |||||
Db        938 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 879

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        878 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 819

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||  |||||||||||| ||| | | |  || |||: | ||||||||||||  || |
Db        818 TCCTGGCCGAGAACGACTGGGAGAKGGAAAGGGCTCTGRACTCCTACTTCGAGCCTCCGG 759

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   | | |  |||| |||   ||  ||  || || || |||||||||
Db        758 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 699

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||  ||  ||| |||||||| |   |
Db        698 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 639

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        638 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 579
```

```
Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||||  | |||||||||||||||||||||||||||  | ||||||||||  || ||||
Db        578 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 519

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||||  |||||||  | |  ||||||||||||||| ||:
Db        518 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAR 459

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||     |||||||||||| |||||||||||||||||||||||||| | |||||||
Db        458 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 399

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        398 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 339

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |:|||| : ||   ||||||||  ||||||||||||||||||
Db        338 GAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 279

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||| |  || |||||||  ||| |
Db        278 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTRA 219

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||:|||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        218 AGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 159

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | ||||:||| |||| || ||||||||||||||  |||||  || | ||||  |||||| |
Db        158 GGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 99

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||  ||   ||    || ||
Db         98 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 39

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATA 1105
              | ||  |||||||| ||| | ||||||||||||||||
Db         38 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 1
```

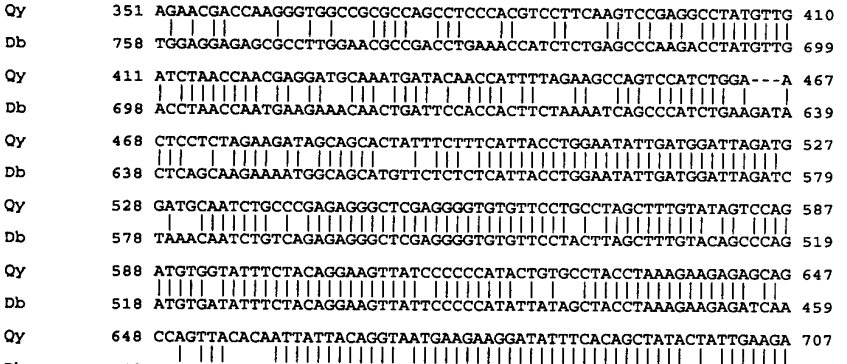
RESULT 60
US-10-170-235-26834/c
; Sequence 26834, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 26834
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-26834

Query Match 43.7%; Score 573.2; DB 42; Length 1849;
Best Local Similarity 76.4%; Pred. No. 3.5e-177;
Matches 717; Conservative 0; Mismatches 218; Indels 3; Gaps 1;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||      |||||| | |||||
Db        938 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 879

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        878 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 819

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||  |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        818 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 759
```



RESULT 61
PCT-US01-08631-22058
; Sequence 22058, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 22058
; LENGTH: 1088
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1)..(687)
; OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein
; OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
; OTHER INFORMATION: Waterman Score=1199.
PCT-US01-08631-22058

Query Match 37.6%; Score 493.2; DB 1; Length 1088;
Best Local Similarity 82.9%; Pred. No. 8.2e-151;
Matches 577; Conservative 0; Mismatches 113; Indels 6; Gaps 1;

RESULT 62
US-10-302-689A-136240
; Sequence 136240, Application US/10302689A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Asundi, Vinod
; APPLICANT: Ballinger, Dennis
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; APPLICANT: Loeb, Deborah
; APPLICANT: Montgomery, Julia, R.

```
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 136240
;   LENGTH: 1088
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-302-689A-136240

  Query Match            37.6%;  Score 493.2;  DB 48;  Length 1088;
  Best Local Similarity  82.9%;  Pred. No. 8.2e-151;
  Matches  577;  Conservative    0;  Mismatches  113;  Indels    6;  Gaps    1;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||    | |||    |||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| | |
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||||  |||| ||  ||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   ||||||||  ||||||||||||||||||||||||||||||||||| ||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||  || |||||||  | | |    ||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || |||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||  ||   ||    || ||  | ||  ||||||||| |||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
              | |||||||||||||||||||||||      |||||||||||| ||| ||||||| | |
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
              ||||| |||||| |||| || ||||||||||||||||||||||||||||| ||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 660

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAA 1237
              || ||  |||||| |||  || || ||||   ||||
Db        661 GGTCTTCTGTGCAACTTAGATATAATATTTGTAAAA 696


RESULT 63
US-10-450-763-22058
; Sequence 22058, Application US/10450763
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22058
;   LENGTH: 1088
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1)..(687)
;   OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=1199.
US-10-450-763-22058

  Query Match            37.6%;  Score 493.2;  DB 51;  Length 1088;
  Best Local Similarity  82.9%;  Pred. No. 8.2e-151;
  Matches  577;  Conservative    0;  Mismatches  113;  Indels    6;  Gaps    1;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||    | |||    |||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| | |
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||||  |||| ||  ||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   ||||||||  ||||||||||||||||||||||||||||||||||| ||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||  || |||||||  | | |    ||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360
```

```
Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || ||||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||  ||    ||     || ||  | ||  |||||||| |||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
              | ||||||||||||||||||||||      |||||||||||| ||| ||||||| | |
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
              |||||  |||||| |||| || ||||||||||||||||||||||||||| ||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 660

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAA 1237
              || ||  |||||| |||  || || ||||  ||||
Db        661 GGTCTTCTGTGCAACTTAGATATAATATTTGTAAAA 696
```

RESULT 64
US-09-248-178-63/c
; Sequence 63, Application US/09248178
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF BREAST CANCER AND
; TITLE OF INVENTION: METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C2
; CURRENT APPLICATION NUMBER: US/09/248,178
; CURRENT FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-248-178-63

Query Match 37.3%; Score 490; DB 22; Length 1079;
Best Local Similarity 82.8%; Pred. No. 9.4e-150;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              |||||||||||| | |||||||||| || |||||||| ||||||||||||||||||| |
Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              ||||||| | |  |||||||||||||||| ||   | |||    |||||||||||| |||
Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              |||||||||||||||||||||| | |||||||||  |||||||||||| ||||| || |
Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              |||||||||||||||||| |||||||||||||||||||  |||| ||  |||||| |
Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||  |||||||| ||||||||||||||||||||||||||||||||||  ||| |||
Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              |||||||||| |  || |||||||  ||| |    |||||||||||||||||||| |||
Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720
```

```
Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || |||||||||||||||||||||||| |||| |||| ||| |||| || |||||||||
Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              |||||||  ||||| || | |||| |||||| ||||| |||||||||||||||||||||
Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              ||||||||||||||  ||    ||     || ||  | ||  |||||||| ||| | |||||
Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
              |||||||||||||||||||      |||||||||||| ||| ||||||| | ||||||| |
Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
              |||||  |||| || ||||||||||||||||||||||||||| |||||||||||||| ||
Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              |||||| |||  || || |||||| ||| ||||  ||  ||  | |||
Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

RESULT 65
US-09-288-950-63/c
; Sequence 63, Application US/09288950
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C3
; CURRENT APPLICATION NUMBER: US/09/288,950
; CURRENT FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-288-950-63

Query Match 37.3%; Score 490; DB 22; Length 1079;
Best Local Similarity 82.8%; Pred. No. 9.4e-150;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              |||||||||||| | |||||||||| || |||||||| ||||||||||||||||||| |
Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

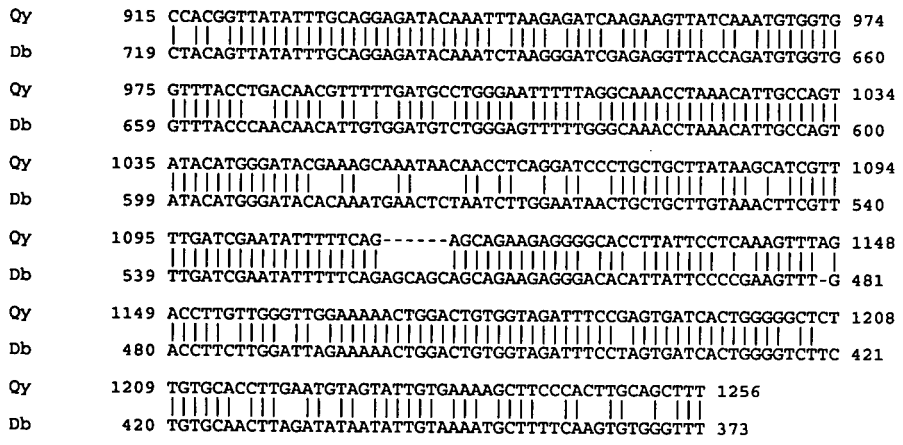
Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              ||||||| | |  |||||||||||||||| ||   | |||    |||||||||||| |||
Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              |||||||||||||||||||||| | |||||||||  |||||||||||| ||||| || |
Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              |||||||||||||||||| |||||||||||||||||||  |||| ||  |||||| |
Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

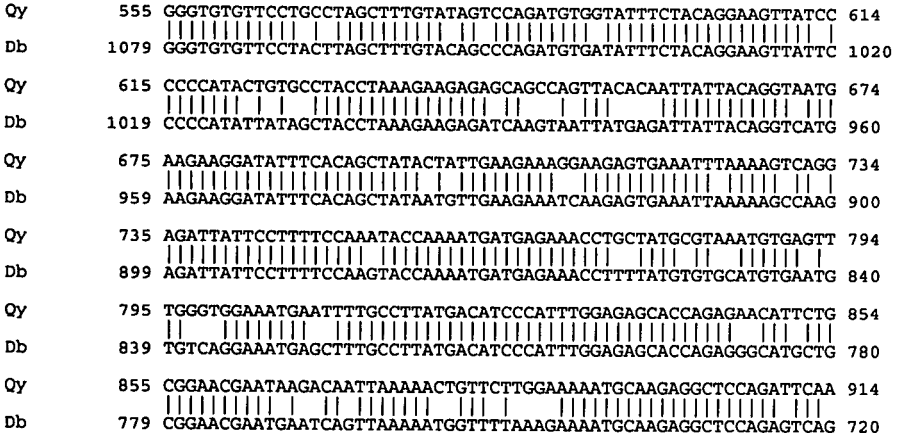
Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||  |||||||| ||||||||||||||||||||||||||||||||||  ||| |||
Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

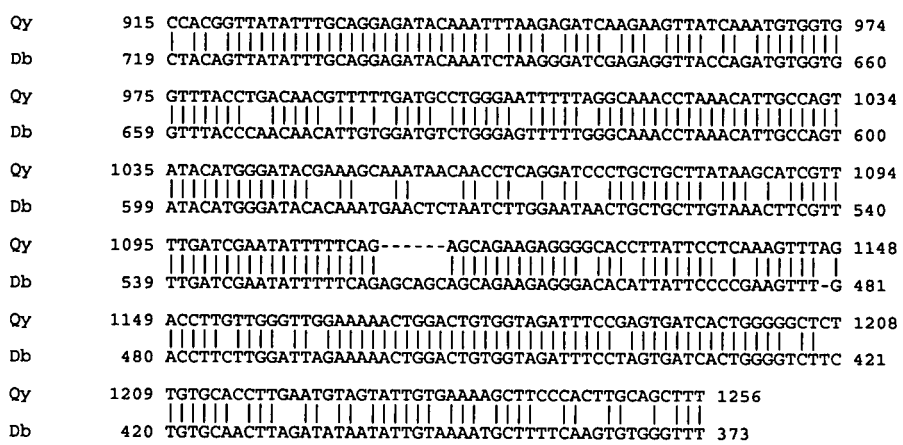
Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              |||||||||| |  || |||||||  ||| |    |||||||||||||||||||| |||
Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720
```



RESULT 66
US-09-346-327-63/c
; Sequence 63, Application US/09346327
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C4
; CURRENT APPLICATION NUMBER: US/09/346,327
; CURRENT FILING DATE: 1999-07-02
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-346-327-63

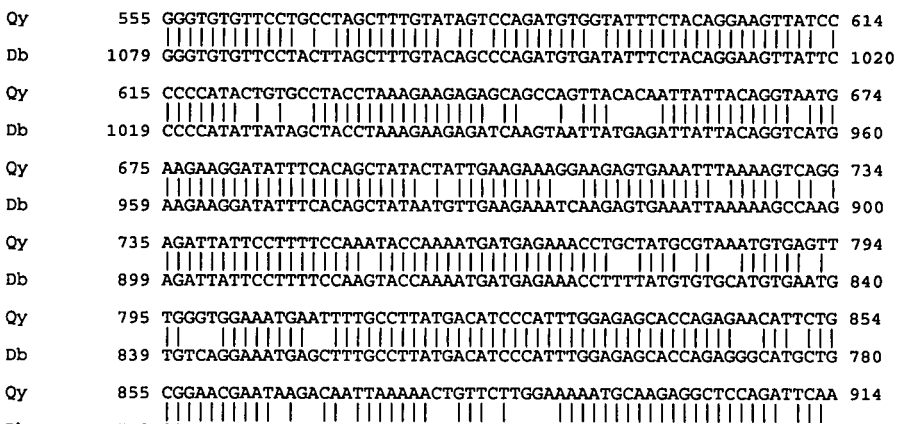
Query Match 37.3%; Score 490; DB 23; Length 1079;
Best Local Similarity 82.8%; Pred. No. 9.4e-150;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;





RESULT 67
US-09-687-507-63/c
; Sequence 63, Application US/09687507
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C6
; CURRENT APPLICATION NUMBER: US/09/687,507
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-687-507-63

Query Match 37.3%; Score 490; DB 29; Length 1079;
Best Local Similarity 82.8%; Pred. No. 9.4e-150;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;



```
Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || ||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              ||||||| ||||| || | |||| |||||| ||||| ||||||||||||||||||||
Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              |||||||||||||  ||   ||    || ||  | ||  |||||||| ||| | |||||
Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
              |||||||||||||||||||      |||||||||||| ||| ||||||| | ||||||| |
Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
              ||||| |||| |||||||||||||||||||||||||||||||| ||||||||||||| ||
Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              |||||| |||  || || |||||| ||| ||||  ||  ||  | |||
Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

RESULT 68
US-09-687-507A-63/c
; Sequence 63, Application US/09687507A
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C6
; CURRENT APPLICATION NUMBER: US/09/687,507A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-687-507A-63

Query Match 37.3%; Score 490; DB 29; Length 1079;
Best Local Similarity 82.8%; Pred. No. 9.4e-150;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              |||||||||||| | |||||||||| || |||||||| |||||||||||||||||| |
Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              ||||||| | |  ||||||||||||||||  ||    | |||   |||||||||||| |||
Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              |||||||||||||||||||||||| | |||||||||  |||||||||||| ||||| || |
Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              |||||||||||||||||| |||||||||||||||||||  |||| ||  |||||| |
Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||   |||||||  |||||||||||||||||||||||||||||||||  ||| |||
Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              |||||||||| |  || |||||||  ||| |    |||||||||||||||||||| |||
```

```
Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || ||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              ||||||| ||||| || | |||| |||||| ||||| ||||||||||||||||||||
Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              |||||||||||||  ||   ||    || ||  | ||  |||||||| ||| | |||||
Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
              |||||||||||||||||||      |||||||||||| ||| ||||||| | ||||||| |
Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
              ||||| |||| || ||||||||||||||||||||||||||||| ||||||||||||| ||
Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              |||||| |||  || || |||||| ||| ||||  ||  ||  | |||
Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

RESULT 69
US-09-745-288-63/c
; Sequence 63, Application US/09745288
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446D1
; CURRENT APPLICATION NUMBER: US/09/745,288
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-745-288-63

Query Match 37.3%; Score 490; DB 31; Length 1079;
Best Local Similarity 82.8%; Pred. No. 9.4e-150;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              |||||||||||| | |||||||||| || |||||||| |||||||||||||||||| |
Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              ||||||| | |  ||||||||||||||||  ||    | |||   |||||||||||| |||
Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              |||||||||||||||||||||||| | |||||||||  |||||||||||| ||||| || |
Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              |||||||||||||||||| |||||||||||||||||||  |||| ||  |||||| |
Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||   |||||||  |||||||||||||||||||||||||||||||||  ||| |||
Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              |||||||||| |  || |||||||  ||| |    |||||||||||||||||||| |||
```

```
Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974

Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034

Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094

Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148

Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208

Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256

Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373


RESULT 70
US-09-778-381-63/c
; Sequence 63, Application US/09778381
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.446C7
;  CURRENT APPLICATION NUMBER: US/09/778,381
;  CURRENT FILING DATE:  2001-02-06
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-778-381-63

  Query Match             37.3%;  Score 490;  DB 32;  Length 1079;
  Best Local Similarity   82.8%;  Pred. No. 9.4e-150;
  Matches  586;  Conservative    0;  Mismatches  115;  Indels    7;  Gaps    2;

Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614

Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674

Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734

Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794

Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854

Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780


Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914

Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974

Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034

Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094

Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148

Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208

Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256

Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373


RESULT 71
US-10-453-919-63/c
; Sequence 63, Application US/10453919
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.446C7
;  CURRENT APPLICATION NUMBER: US/10/453,919
;  CURRENT FILING DATE:  2003-06-03
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-453-919-63

  Query Match             37.3%;  Score 490;  DB 51;  Length 1079;
  Best Local Similarity   82.8%;  Pred. No. 9.4e-150;
  Matches  586;  Conservative    0;  Mismatches  115;  Indels    7;  Gaps    2;

Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614

Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674

Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734

Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794

Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
```

```
Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914

Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974

Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034

Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094

Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148

Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208

Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256

Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

RESULT 72
US-09-644-868-7772
; Sequence 7772, Application US/09644868
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1168-001
; CURRENT APPLICATION NUMBER: US/09/644,868
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,063
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 10075
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7772
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1852)
; OTHER INFORMATION: n = A,T,C or G
US-09-644-868-7772

Query Match 36.1%; Score 473.6; DB 29; Length 1852;
Best Local Similarity 75.8%; Pred. No. 4.7e-144;
Matches 600; Conservative 0; Mismatches 189; Indels 3; Gaps 1;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         49 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 108

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        109 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 168

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        169 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 228

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        229 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 288

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        289 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 348

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        349 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 408

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        409 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 468

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        469 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 528

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        529 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 588

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        589 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 648

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        649 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 708

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        709 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 768

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        769 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 828

Qy        948 GAGATCAAGAAG 959

Db        829 GGGATCGAGAGG 840
```

RESULT 73
US-09-649-161-8751
; Sequence 8751, Application US/09649161
; GENERAL INFORMATION:
; APPLICANT: Gutierrez-Ramos, Jose-Carlos
; APPLICANT: Villeval, Jean-Luc M.G.
; APPLICANT: Fraser, Christopher C.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1178-001
; CURRENT APPLICATION NUMBER: US/09/649,161
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,060
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 10788
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8751
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1852)
; OTHER INFORMATION: n = A,T,C or G
US-09-649-161-8751

```
  Query Match             36.1%;  Score 473.6;  DB 29;  Length 1852;
  Best Local Similarity   75.8%;  Pred. No. 4.7e-144;
  Matches  600;  Conservative    0;  Mismatches  189;  Indels    3;  Gaps    1;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |  | | ||  |||     ||||| | |||||
Db         49 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 108

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| ||| ||| |||||| | ||| |   | |||||  || |  |||  |   ||
Db        109 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 168

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | | || |||| | ||||||||||||  || |
Db        169 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 228

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   |  | |  |||| |||   ||   ||  || || || |||||||||
Db        229 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 288

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        289 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 348

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| |||||||||||||||||||||||||||||
Db        349 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 408

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||| | |||||||||| || ||||
Db        409 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 468

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||||| ||
Db        469 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 528

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db        529 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 588

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        589 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 648

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   ||||||||  ||||||||||||||||||
Db        649 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 708

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||  |  || |||||||  ||| |
Db        709 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 768

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        769 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 828

Qy        948 GAGATCAAGAAG 959
              | |||| ||| |
Db        829 GGGATCGAGAGG 840


RESULT 74
US-09-652-109-8165
; Sequence 8165, Application US/09652109
; GENERAL INFORMATION:
;  APPLICANT:  McCarthy, Sean A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1180-001
;  CURRENT APPLICATION NUMBER: US/09/652,109
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,128
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  10105
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8165
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-109-8165

  Query Match             36.1%;  Score 473.6;  DB 29;  Length 1852;
  Best Local Similarity   75.8%;  Pred. No. 4.7e-144;
  Matches  600;  Conservative    0;  Mismatches  189;  Indels    3;  Gaps    1;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |  | | ||  |||     ||||| | |||||
Db         49 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 108

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| ||| ||| |||||| | ||| |   | |||||  || |  |||  |   ||
Db        109 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 168

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | | || |||| | ||||||||||||  || |
Db        169 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 228

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   |  | |  |||| |||   ||   ||  || || || |||||||||
Db        229 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 288

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        289 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 348

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| |||||||||||||||||||||||||||||
Db        349 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 408

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||| | |||||||||| || ||||
Db        409 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 468

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||||| ||
Db        469 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 528

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db        529 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 588

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        589 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 648

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   ||||||||  ||||||||||||||||||
Db        649 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 708

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||  |  || |||||||  ||| |
Db        709 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 768

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        769 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 828

Qy        948 GAGATCAAGAAG 959
              | |||| ||| |
Db        829 GGGATCGAGAGG 840
```

```
RESULT 75
US-09-652-128-7441
; Sequence 7441, Application US/09652128
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  APPLICANT:  Holtzman, Douglas A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1171-001
;  CURRENT APPLICATION NUMBER: US/09/652,128
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,133
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  10265
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 7441
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-128-7441

  Query Match             36.1%;  Score 473.6;  DB 29;  Length 1852;
  Best Local Similarity   75.8%;  Pred. No. 4.7e-144;
  Matches  600;  Conservative    0;  Mismatches  189;  Indels    3;  Gaps    1;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     |||||| | |||||
Db         49 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 108

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        109 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 168

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        169 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 228

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   ||||  |||   ||   ||   || || || |||||||||
Db        229 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 288

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||  ||  ||| ||||||| |   |
Db        289 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 348

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||    | ||| ||||||||||||||||||||||||||||||
Db        349 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 408

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||||| | ||||||||||||||||||||||||||| | |||||||||| || ||||
Db        409 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 468

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        469 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 528

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db        529 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 588

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        589 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 648

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   ||||||||  ||||||||||||||||||
Db        649 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 708

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||||  ||| ||||||||||||| | || |||||||  ||| |
Db        709 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 768

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        769 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 828

Qy        948 GAGATCAAGAAG 959
              | |||| ||| |
Db        829 GGGATCGAGAGG 840


Search completed: March 11, 2006, 16:34:36
Job time : 8123 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2006, 14:55:32 ; Search time 443 Seconds
(without alignments)
7022.099 Million cell updates/sec

Title: US-10-757-745-3
Perfect score: 1312
Sequence: 1 agctattaatgattcgaatt.....tctcaacctttcaggacatc 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5007642 seqs, 1185514400 residues

Total number of hits satisfying chosen parameters: 10015284

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq1:*
9: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	666.6	50.8	1940	6	US-10-567-867-921	Sequence 921, App
	2	666.6	50.8	1940	11	US-60-742-219-2071	Sequence 2071, Ap
	3	665	50.7	1898	6	US-10-917-503B-13209	Sequence 13209, A
	4	618.2	47.1	1288	9	US-11-266-748A-226184	Sequence 226184,
	5	377.8	28.8	752	6	US-10-917-503B-4908	Sequence 4908, Ap
	6	281.8	21.5	849	9	US-11-266-748A-166646	Sequence 166646,
	7	265.6	20.2	1417	6	US-10-276-817B-5136	Sequence 5136, Ap
	8	236.8	18.0	1179	9	US-11-266-748A-19660	Sequence 19660, A
	9	180.8	13.8	1205	9	US-11-266-748A-184613	Sequence 184613,
	10	180.8	13.8	1205	9	US-11-266-748A-192113	Sequence 192113,
	11	121.8	9.3	1000	9	US-11-266-748A-117422	Sequence 117422,
c	12	121.8	9.3	1000	9	US-11-266-748A-159586	Sequence 159586,
	13	121.8	9.3	1000	9	US-11-266-748A-289180	Sequence 289180,
c	14	121.8	9.3	1000	9	US-11-266-748A-340609	Sequence 340609,
	15	58.2	4.4	1007	9	US-11-266-748A-79844	Sequence 79844, A
c	16	58.2	4.4	1007	9	US-11-266-748A-132655	Sequence 132655,
c	17	57.8	4.4	617	9	US-11-266-748A-219847	Sequence 219847,
	18	57.8	4.4	617	9	US-11-266-748A-239642	Sequence 239642,
c	19	57.6	4.4	315	6	US-10-515-325-731	Sequence 731, App
c	20	55.8	4.3	614	9	US-11-266-748A-219858	Sequence 219858,
c	21	55.2	4.2	678	9	US-11-266-748A-219918	Sequence 219918,
	22	55.2	4.2	678	9	US-11-266-748A-239685	Sequence 239685,
c	23	54.8	4.2	61	6	US-10-199-820A-142	Sequence 142, App
c	24	54.8	4.2	61	6	US-10-199-820B-142	Sequence 142, App
c	25	54.8	4.2	653	9	US-11-266-748A-220017	Sequence 220017,
c	26	54.8	4.2	738	9	US-11-266-748A-45850	Sequence 45850, A
	27	54.8	4.2	1044	5	US-09-831-621A-14	Sequence 14, Appl
	28	54.8	4.2	3289	6	US-10-461-673-1880	Sequence 1880, Ap
c	29	53.8	4.1	661	9	US-11-266-748A-219976	Sequence 219976,
	30	53.8	4.1	661	9	US-11-266-748A-239716	Sequence 239716,
c	31	53.8	4.1	2300	11	US-60-762-056-35503	Sequence 35503, A
	32	53.6	4.1	2300	11	US-60-762-056-35504	Sequence 35504, A
c	33	52.2	4.0	600	9	US-11-266-748A-219957	Sequence 219957,
	34	52.2	4.0	2300	11	US-60-762-056-20736	Sequence 20736, A
	35	52.2	4.0	2968	7	US-11-283-329-179	Sequence 179, App
	36	51.4	3.9	971	9	US-11-266-748A-267412	Sequence 267412,
c	37	51.4	3.9	971	9	US-11-266-748A-327929	Sequence 327929,
c	38	51.2	3.9	890	9	US-11-266-748A-102287	Sequence 102287,
	39	51.2	3.9	890	9	US-11-266-748A-155098	Sequence 155098,
	40	51	3.9	2954	6	US-10-531-147-349	Sequence 349, App
	41	51	3.9	2954	6	US-10-531-147-1529	Sequence 1529, Ap
c	42	49.6	3.8	602	9	US-11-227-183A-4003	Sequence 4003, Ap
	43	49.4	3.8	555	7	US-11-329-388-9794	Sequence 9794, Ap
c	44	49	3.7	2300	11	US-60-762-056-6363	Sequence 6363, Ap
	45	49	3.7	2737	9	US-11-266-748A-30544	Sequence 30544, A
	46	48.6	3.7	420	9	US-11-292-078-5998	Sequence 5998, Ap
c	47	48.6	3.7	2300	11	US-60-762-056-4866	Sequence 4866, Ap
	48	48.4	3.7	627	7	US-11-329-388-8161	Sequence 8161, Ap
	49	48.2	3.7	398	7	US-11-329-388-8142	Sequence 8142, Ap
	50	48.2	3.7	799	9	US-11-311-886-3337	Sequence 3337, Ap
	51	48.2	3.7	1136	9	US-11-266-748A-40236	Sequence 40236, A
	52	47.8	3.6	1109	6	US-10-953-349-34914	Sequence 34914, A
	53	47.6	3.6	1183	9	US-11-266-748A-69859	Sequence 69859, A
	54	47.6	3.6	1183	9	US-11-266-748A-105667	Sequence 105667,
c	55	47.6	3.6	1183	9	US-11-266-748A-122670	Sequence 122670,
	56	47.6	3.6	2080	9	US-11-266-748A-69138	Sequence 69138, A
c	57	47.6	3.6	2080	9	US-11-266-748A-121949	Sequence 121949,
	58	47.2	3.6	452	7	US-11-353-150-5643	Sequence 5643, Ap
c	59	47	3.6	748	9	US-11-266-748A-219917	Sequence 219917,
	60	47	3.6	748	9	US-11-266-748A-239684	Sequence 239684,
	61	46.8	3.6	1077	9	US-11-266-748A-82366	Sequence 82366, A
	62	46.8	3.6	1077	9	US-11-266-748A-111289	Sequence 111289,
c	63	46.8	3.6	1077	9	US-11-266-748A-135177	Sequence 135177,
	64	46.8	3.6	2300	11	US-60-762-056-19219	Sequence 19219, A
c	65	46.6	3.6	1171	9	US-11-266-748A-70423	Sequence 70423, A
c	66	46.6	3.6	1171	9	US-11-266-748A-106000	Sequence 106000,
	67	46.6	3.6	1171	9	US-11-266-748A-123234	Sequence 123234,
	68	46.2	3.5	1088	9	US-11-266-748A-81563	Sequence 81563, A
c	69	46.2	3.5	1088	9	US-11-266-748A-134374	Sequence 134374,
	70	45.8	3.5	823	9	US-11-266-748A-47816	Sequence 47816, A
c	71	45.6	3.5	409	7	US-11-330-082-8448	Sequence 8448, Ap
c	72	45.6	3.5	671	7	US-11-330-364-11771	Sequence 11771, A
c	73	45.6	3.5	2300	11	US-60-762-056-22874	Sequence 22874, A
	74	45.4	3.5	611	9	US-11-266-748A-92915	Sequence 92915, A
c	75	45.4	3.5	611	9	US-11-266-748A-145726	Sequence 145726,

ALIGNMENTS

RESULT 1
US-10-567-867-921
; Sequence 921, Application US/10567867
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PCT
; CURRENT APPLICATION NUMBER: US/10/567,867
; CURRENT FILING DATE: 2006-01-27
; PRIOR APPLICATION NUMBER: US 60/490,890
; PRIOR FILING DATE: 2003-07-29

```
;  NUMBER OF SEQ ID NOS: 2786
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-567-867-921

  Query Match            50.8%;  Score 666.6;  DB 6;  Length 1940;
  Best Local Similarity  76.9%;  Pred. No. 6.4e-174;
  Matches  842;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    2;

Qy       171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db        42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db       102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db       162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db       222 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db       282 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 341

Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db       342 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 401

Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db       402 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 461

Qy       588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db       462 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 521

Qy       648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db       522 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 581

Qy       708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db       582 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 641

Qy       768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db       642 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 701

Qy       828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db       702 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 761

Qy       888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db       762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy       948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db       822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy      1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db       882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy      1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db       942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy      1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db      1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy      1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db      1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

Qy      1242 CCCACTTGCAGCTTT 1256
Db      1122 TTCAAGTGTGGGTTT 1136


RESULT 2
US-60-742-219-2071
; Sequence 2071, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2071
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-2071

  Query Match            50.8%;  Score 666.6;  DB 11;  Length 1940;
  Best Local Similarity  76.9%;  Pred. No. 6.4e-174;
  Matches  842;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    2;

Qy       171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db        42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy       231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db       102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy       291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db       162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy       351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db       222 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281

Qy       411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db       282 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 341

Qy       468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db       342 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 401

Qy       528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
```

```
Db        402 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 461

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        462 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 521

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        522 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 581

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        582 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 641

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        642 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 701

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        702 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 761

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1122 TTCAAGTGTGGGTTT 1136
```

```
RESULT 3
US-10-917-503B-13209
; Sequence 13209, Application US/10917503B
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
;   LENGTH: 1898
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1086)
US-10-917-503B-13209

  Query Match             50.7%;  Score 665;  DB 6;  Length 1898;
  Best Local Similarity   76.8%;  Pred. No. 1.8e-173;
  Matches  841;  Conservative    0;  Mismatches  245;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        617 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 676
```

```
Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||||  ||| ||||||||||||||  |  || |||||||  ||| |
Db        677 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 736

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        737 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 796

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  ||||| || | |||| |||||| |
Db        797 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 856

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||||  ||    ||   || ||
Db        857 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 916

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||| ||| | |||||||||||||||||||||||||      ||||||||
Db        917 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 976

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| |||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db        977 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1036

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1037 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1096

Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1097 TTCAAGTGTGGGTTT 1111
```

RESULT 4
US-11-266-748A-226184
; Sequence 226184, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 226184
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-226184

Query Match 47.1%; Score 618.2; DB 9; Length 1288;
Best Local Similarity 79.8%; Pred. No. 1.3e-160;
Matches 757; Conservative 0; Mismatches 183; Indels 9; Gaps 2;

```
Qy        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376
              ||||  || |||| | ||||||||||||  || | | | ||     | | |   ||||
Db         15 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 74

Qy        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436
              |||   ||   ||   || || || ||||||||||| |||||||| || ||  ||| ||||
Db         75 CCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 134

Qy        437 ACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATT 493
              | |||| ||   ||  ||| ||||||| |   |||| | |||| || ||||||   |
Db        135 TCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTC 194

Qy        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553
              ||| |||||||||||||||||||||||||||||  | ||||||| | ||||||||||||
Db        195 TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 254

Qy        554 GGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATC 613
              |||||||||||||| | |||||||||| || |||||||| ||||||||||||||||||||
Db        255 GGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATT 314

Qy        614 CCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAAT 673
              |||||||| | |  ||||||||||||||||| ||   | |||   ||||||||||| ||
Db        315 CCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCAT 374

Qy        674 GAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAG 733
              |||||||||||||||||||||||| | |||||||||  |||||||||||| ||||| ||
Db        375 GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAA 434

Qy        734 GAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT 793
              |||||||||||||||||| ||||||||||||||||||||  |||| ||  |||||| |
Db        435 GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAT 494

Qy        794 TTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCT 853
              ||   |||||||  |||||||||||||||||||||||||||||||||  ||| ||
Db        495 GTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCT 554

Qy        854 GCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAAATGCAAGAGGCTCCAGATTCA 913
              |||||||||| |  || |||||||  ||| |    ||||||||||||||||||| |||
Db        555 GCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCA 614

Qy        914 ACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGT 973
              | || |||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db        615 GCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGT 674

Qy        974 GGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAG 1033
              ||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||||||
Db        675 GGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAG 734

Qy       1034 TATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGT 1093
              |||||||||||||  ||   ||   || ||  | ||  ||||||||| ||| | ||||
Db        735 TATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGT 794

Qy       1094 TTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTA 1147
              ||||||||||||||||||      |||||||||||| ||| ||||||| | ||||||
Db        795 TTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 854

Qy       1148 GACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTC 1207
              |||||| |||| || |||||||||||||||||||||||||||||| ||||||||||||| ||
Db        855 GACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTT 914

Qy       1208 TTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              |||||| |||   || || |||||| ||| ||||  ||  ||  | |||
Db        915 CTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 963
```

RESULT 5
US-10-917-503B-4908
; Sequence 4908, Application US/10917503B
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI

```
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4908
;   LENGTH: 752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (685)..(685)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (694)..(694)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (740)..(740)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-10-917-503B-4908

  Query Match             28.8%;  Score 377.8;  DB 6;  Length 752;
  Best Local Similarity   72.5%;  Pred. No. 3.3e-94;
  Matches  528;  Conservative    0;  Mismatches  195;  Indels    5;  Gaps    3;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    | |   |  || |   | | ||  |||     ||||| | |||||
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC  76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || || |||  |   ||
Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||| ||||||||  || |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |   |||| |||   ||   ||   || || || |||||||||
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |   |
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||

Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||| |||| | |||||||||| || ||||
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||||| | |||||||
Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy        768 G-AAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 826
              | |||||  |||| ||  ||||||   ||   ||||||||  |||||||||  |||||||
Db        617 GAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTAAGACATCC 676

Qy        827 C-ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCT 885
              | ||| || || ||||| ||  |||      ||||||||| |  || || ||||  | | |
Db        677 CAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATCAGTTTAAAATGGGTTT 736

Qy        886 TGGAAAAA 893
                  |||
Db        737 AAANGAAA 744


RESULT 6
US-11-266-748A-166646
; Sequence 166646, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 166646
;   LENGTH: 849
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-166646

  Query Match             21.5%;  Score 281.8;  DB 9;  Length 849;
  Best Local Similarity   76.9%;  Pred. No. 1.2e-67;
  Matches  399;  Conservative    0;  Mismatches  107;  Indels   13;  Gaps    4;

Qy        751 AAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGT---GGAAATGA 807
                ||  ||||||| |||     | ||   ||| ||  ||| |    || ||   |  |||||
```

```
Db          4 AAGAACCAAAAAGATAGAGGAACTTGGATGTGTGCATGAGCAACTGTGTCAGGACAATGA 63

Qy        808 ATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAG 867

Db         64 GCTATGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAA 123

Qy        868 ACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATT 927

Db        124 TCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATT 183

Qy        928 TGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAA--ATGTGGTGGTTTACCTGAC 985

Db        184 TGCAGGAGATACAAATCTAAGGGATCGAGCAGGTTACCAGCATGTGGTGGTTTACCCAAC 243

Qy        986 AACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGAT 1045

Db        244 AACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGCA 303

Qy       1046 AC--GAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAA 1103

Db        304 TACACAAATGCAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAA 363

Qy       1104 TATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTG 1157

Db        364 TATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTG 423

Qy       1158 GGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCT 1217

Db        424 GATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACT 483

Qy       1218 TGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256

Db        484 TAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 522
```

RESULT 7
US-10-276-817B-5136
; Sequence 5136, Application US/10276817B
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: US/10/276,817B
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US01/14827
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 5136
; LENGTH: 1417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (48)..(422)
; OTHER INFORMATION: 79% homologous to Homo sapiens dJ30M3.3 (novel protein
; OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
; OTHER INFORMATION: Waterman Score=494.
US-10-276-817B-5136

Query Match 20.2%; Score 265.6; DB 6; Length 1417;
Best Local Similarity 80.2%; Pred. No. 4.4e-63;
Matches 377; Conservative 0; Mismatches 79; Indels 14; Gaps 5;

```
Qy        786 ATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAG 845

Db         52 AGGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAG 111

Qy        846 AACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTC 905

Db        112 GGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTC 171

Qy        906 CAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCA 965

Db        172 CAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCA 231

Qy        966 AATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAAC 1025

Db        232 GATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAAC 291

Qy       1026 ATTGCCA-GTATACATGGG---ATACGAAAGCAAATAACAACCTCAGGA--TCCCTGCTG 1079

Db        292 ATTGCCACGTATACATTGGGATACACAAACTGAACTCCTAATCTTGGGAATAACTGCCTG 351

Qy       1080 CTTATAAGCATCGTTTTGATCGAATATTTTTCAG-------AGCAGAAGAGGGGCACCTT 1132

Db        352 CTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAAGCAGCAGCAGAAGAGGGACACATT 411

Qy       1133 ATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGT 1192

Db        412 ATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGT 471

Qy       1193 GATCACTGGGGGCTCTTGTGCACCTT-GAATGTAGTATTGTGAAAAGCTT 1241

Db        472 GATCACTGGGGTCTTCTGTGCAACTTAGGATATAATATTGTAAAATGCTT 521
```

RESULT 8
US-11-266-748A-19660
; Sequence 19660, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19660
; LENGTH: 1179
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-19660

Query Match 18.0%; Score 236.8; DB 9; Length 1179;
Best Local Similarity 81.0%; Pred. No. 3.8e-55;
Matches 290; Conservative 0; Mismatches 62; Indels 6; Gaps 1;

```
Qy        905 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 964

Db          1 CCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACC 60

Qy        965 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 1024

Db         61 AGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAA 120
```

```
Qy       1025 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1084
             ||||||||||||||||||||||||  ||   ||   || ||  | ||  |||||||||| |
Db        121 CATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGT 180

Qy       1085 AAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCT 1138
             || | ||||||||||||||||||||||||      |||||||||||| ||| |||||||
Db        181 AAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 240

Qy       1139 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1198
             | |||||| |||||| |||| || ||||||||||||||||||||||||||| |||||||||
Db        241 CGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCAC 300

Qy       1199 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
             ||||| ||  |||||| |||  || || |||||| ||| ||||  ||  ||  | |||
Db        301 TGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 358
```

RESULT 9
US-11-266-748A-184613
; Sequence 184613, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 184613
; LENGTH: 1205
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (64)..(81)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-184613

Query Match 13.8%; Score 180.8; DB 9; Length 1205;
Best Local Similarity 79.1%; Pred. No. 1.2e-39;
Matches 242; Conservative 0; Mismatches 57; Indels 7; Gaps 2;

```
Qy        957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016
              | |||| || ||||||||||||||  ||||| || | |||| |||||| ||||| ||||
Db        135 AGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCA 194

Qy       1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076
              |||||||||||||||||||||||||||||  ||   ||   || ||  | ||  ||| |
Db        195 AACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTG 254

Qy       1077 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCA------GAGCAGAAGAGGGGCACC 1130
              |||||| ||| | |||||||||||||||||||||||      |||||||||||| |||
Db        255 CTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAAAGCAGCAGCAGAAGAGGGACACA 314
```

```
Qy       1131 TTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGA 1190
              ||||||| | |||||| | |||| |||| ||  ||||||||||||||||||||||||| |
Db        315 TTATTCCCCGAAGTTT-GGCCTTCTTGGATTAAAAAAACTGGACTGTGGTAGATTTCCTA 373

Qy       1191 GTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGC 1250
              |||||||||||| ||  |||||| |||  || || |||||  ||| ||||  ||  ||
Db        374 GTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGT 433

Qy       1251 AGCTTT 1256
              | |||
Db        434 GGGTTT 439
```

RESULT 10
US-11-266-748A-192113
; Sequence 192113, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 192113
; LENGTH: 1205
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (64)..(81)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-192113

Query Match 13.8%; Score 180.8; DB 9; Length 1205;
Best Local Similarity 79.1%; Pred. No. 1.2e-39;
Matches 242; Conservative 0; Mismatches 57; Indels 7; Gaps 2;

```
Qy        957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016
              | |||| || ||||||||||||||  ||||| || | |||| |||||| ||||| ||||
Db        135 AGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCA 194

Qy       1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076
              |||||||||||||||||||||||||||||  ||   ||   || ||  | ||  ||| |
Db        195 AACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTG 254

Qy       1077 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCA------GAGCAGAAGAGGGGCACC 1130
              |||||| ||| | |||||||||||||||||||||||      |||||||||||| |||
Db        255 CTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAAAGCAGCAGCAGAAGAGGGACACA 314

Qy       1131 TTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGA 1190
              ||||||| | |||||| | |||| |||| ||  ||||||||||||||||||||||||| |
Db        315 TTATTCCCCGAAGTTT-GGCCTTCTTGGATTAAAAAAACTGGACTGTGGTAGATTTCCTA 373
```

```
Qy       1191 GTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGC 1250
              ||||||||||||| ||  |||||| |||  || || |||||| ||| ||||  ||  ||
Db        374 GTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGT 433

Qy       1251 AGCTTT 1256
                 | |||
Db        434 GGGTTT 439


RESULT 11
US-11-266-748A-117422
; Sequence 117422, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 117422
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-117422

  Query Match             9.3%;  Score 121.8;  DB 9;  Length 1000;
  Best Local Similarity   80.5%;  Pred. No. 2.3e-23;
  Matches  157;  Conservative    0;  Mismatches   32;  Indels    6;  Gaps    1;

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||| | ||||||||||||||||||||||||      ||||||||
Db          2 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 61

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db         62 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 121

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db        122 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 181

Qy       1242 CCCACTTGCAGCTTT 1256
                ||  ||  | |||
Db        182 TTCAAGTGTGGGTTT 196


RESULT 12
US-11-266-748A-159586/c
; Sequence 159586, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 159586
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-159586

  Query Match             9.3%;  Score 121.8;  DB 9;  Length 1000;
  Best Local Similarity   80.5%;  Pred. No. 2.3e-23;
  Matches  157;  Conservative    0;  Mismatches   32;  Indels    6;  Gaps    1;

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||| | ||||||||||||||||||||||||      ||||||||
Db        999 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 940

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db        939 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 880

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db        879 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 820

Qy       1242 CCCACTTGCAGCTTT 1256
                ||  ||  | |||
Db        819 TTCAAGTGTGGGTTT 805


RESULT 13
US-11-266-748A-289180
; Sequence 289180, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
```

```
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 289180
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-289180

  Query Match             9.3%;  Score 121.8;  DB 9;  Length 1000;
  Best Local Similarity   80.5%;  Pred. No. 2.3e-23;
  Matches  157;  Conservative    0;  Mismatches   32;  Indels    6;  Gaps    1;

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
                | ||  |||||||||| ||| | ||||||||||||||||||||||||      ||||||||
Db          2 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 61

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
                |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db         62 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 121

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
                ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db        122 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 181

Qy       1242 CCCACTTGCAGCTTT 1256
                ||  ||  | |||
Db        182 TTCAAGTGTGGGTTT 196


RESULT 14
US-11-266-748A-340609/c
; Sequence 340609, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 340609
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-340609

  Query Match             9.3%;  Score 121.8;  DB 9;  Length 1000;
  Best Local Similarity   80.5%;  Pred. No. 2.3e-23;
  Matches  157;  Conservative    0;  Mismatches   32;  Indels    6;  Gaps    1;

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
                | ||  |||||||||| ||| | ||||||||||||||||||||||||      ||||||||
Db        999 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 940

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
                |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db        939 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 880

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
                ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db        879 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 820

Qy       1242 CCCACTTGCAGCTTT 1256
                ||  ||  | |||
Db        819 TTCAAGTGTGGGTTT 805


RESULT 15
US-11-266-748A-79844
; Sequence 79844, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79844
;   LENGTH: 1007
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (198)..(198)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (253)..(263)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (926)..(935)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-79844

  Query Match             4.4%;  Score 58.2;  DB 9;  Length 1007;
  Best Local Similarity   75.0%;  Pred. No. 8.8e-06;
  Matches   72;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              |||||| |||| |||||||||||||||||||||||||| |  |||||||| ||| ||| |
```

```
Db        157 AGCTATGAATGCTTCGAATTTATACGACTCACTATAGGAATNTTGGCCCTGGAGCCCAGG 216

Qy         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCG 96
              | | || ||||||| |       |  | | || |||
Db        217 ATTCCGCCACGAGGACCACCTAAATTGAGGAGTGCG 252


RESULT 16
US-11-266-748A-132655/c
; Sequence 132655, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 132655
;   LENGTH: 1007
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (73)..(82)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (745)..(755)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (810)..(810)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-132655

  Query Match             4.4%;  Score 58.2;  DB 9;  Length 1007;
  Best Local Similarity   75.0%;  Pred. No. 8.8e-06;
  Matches   72;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              |||||| |||| |||||||||||||||||||||||||| |  |||||||| ||| ||| |
Db        851 AGCTATGAATGCTTCGAATTTATACGACTCACTATAGGAATNTTGGCCCTGGAGCCCAGG 792

Qy         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCG 96
              | | || ||||||| |       |  | | || |||
Db        791 ATTCCGCCACGAGGACCACCTAAATTGAGGAGTGCG 756


RESULT 17
US-11-266-748A-219847/c
; Sequence 219847, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 219847
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (371)..(371)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-219847

  Query Match             4.4%;  Score 57.8;  DB 9;  Length 617;
  Best Local Similarity   80.0%;  Pred. No. 9.4e-06;
  Matches   68;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 487

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGT 101
                ||   | | || |  |  | |
Db        486 TTTGAGTTTCATAAATAATGCTTCT 462


RESULT 18
US-11-266-748A-239642
; Sequence 239642, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
```

```
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 239642
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (247)..(247)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-239642

  Query Match             4.4%;  Score 57.8;  DB 9;  Length 617;
  Best Local Similarity   80.0%;  Pred. No. 9.4e-06;
  Matches   68;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 131

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGT 101
                ||   | | || |  |  | |
Db        132 TTTGAGTTTCATAAATAATGCTTCT 156


RESULT 19
US-10-515-325-731/c
; Sequence 731, Application US/10515325
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Higgs, Brandon
;  APPLICANT:  Elashoff, Michael
;  TITLE OF INVENTION: Molecular Nephrotoxicology Modeling
;  FILE REFERENCE: 044921-5089-02-US
;  CURRENT APPLICATION NUMBER: US/10/515,325
;  CURRENT FILING DATE:  2004-11-23
;  PRIOR APPLICATION NUMBER: US 10/301,856
;  PRIOR FILING DATE: 2002-11-22
;  PRIOR APPLICATION NUMBER: PCT/US2003/037556
;  PRIOR FILING DATE: 2003-11-24
;  NUMBER OF SEQ ID NOS: 1617
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 731
;   LENGTH: 315
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-10-515-325-731

  Query Match             4.4%;  Score 57.6;  DB 6;  Length 315;
  Best Local Similarity   82.5%;  Pred. No. 8.2e-06;
  Matches   66;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        291 ATTTAATACGACTCCCTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 232

Qy         77 GGGAAGCAGCGTGAAGAGCG 96
              ||||      |  |  ||| |
Db        231 GGGAGTGGTCCCGGCGAGTG 212


RESULT 20
US-11-266-748A-219858/c
; Sequence 219858, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 219858
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-219858

  Query Match             4.3%;  Score 55.8;  DB 9;  Length 614;
  Best Local Similarity   84.0%;  Pred. No. 3.4e-05;
  Matches   63;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGG 476

Qy         77 GGGAAGCAGCGTGAA 91
                |   ||  | ||
Db        475 AATACAAAGTTTAAA 461


RESULT 21
US-11-266-748A-219918/c
; Sequence 219918, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
```

```
; SEQ ID NO 219918
;   LENGTH: 678
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-219918

  Query Match             4.2%;  Score 55.2;  DB 9;  Length 678;
  Best Local Similarity   95.0%;  Pred. No. 5.1e-05;
  Matches   57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        586 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCC 527


RESULT 22
US-11-266-748A-239685
; Sequence 239685, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 239685
;   LENGTH: 678
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-239685

  Query Match             4.2%;  Score 55.2;  DB 9;  Length 678;
  Best Local Similarity   95.0%;  Pred. No. 5.1e-05;
  Matches   57;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         93 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCC 152


RESULT 23
US-10-199-820A-142/c
; Sequence 142, Application US/10199820A
; GENERAL INFORMATION:
;  APPLICANT: Board of Trustees of the University of Illinois
;  APPLICANT:  Primiano, Thomas
;  APPLICANT:  Chang, Bey-dih
;  APPLICANT:  Roninson, Igor
;  TITLE OF INVENTION: Methods and Reagents for Identifying Gene Targets for Treating Ca
;  FILE REFERENCE: 99,216-U
;  CURRENT APPLICATION NUMBER: US/10/199,820A
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: PCT/US02/06254
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/306730
;  PRIOR FILING DATE: 2001-07-20
;  NUMBER OF SEQ ID NOS: 315
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 142
;   LENGTH: 61
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-199-820A-142

  Query Match             4.2%;  Score 54.8;  DB 6;  Length 61;
  Best Local Similarity   96.6%;  Pred. No. 2.6e-05;
  Matches   56;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 1


RESULT 24
US-10-199-820B-142/c
; Sequence 142, Application US/10199820B
; GENERAL INFORMATION:
;  APPLICANT: Board of Trustees of the University of Illinois
;  APPLICANT:  Primiano, Thomas
;  APPLICANT:  Chang, Bey-dih
;  APPLICANT:  Roninson, Igor
;  TITLE OF INVENTION: Methods and Reagents for Identifying Gene Targets for Treating Ca
;  FILE REFERENCE: 99,216-U
;  CURRENT APPLICATION NUMBER: US/10/199,820B
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: PCT/US02/06254
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/306730
;  PRIOR FILING DATE: 2001-07-20
;  NUMBER OF SEQ ID NOS: 322
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 142
;   LENGTH: 61
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-199-820B-142

  Query Match             4.2%;  Score 54.8;  DB 6;  Length 61;
  Best Local Similarity   96.6%;  Pred. No. 2.6e-05;
  Matches   56;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 1


RESULT 25
US-11-266-748A-220017/c
; Sequence 220017, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
```

; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 220017
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-220017

Query Match 4.2%; Score 54.8; DB 9; Length 653;
Best Local Similarity 96.6%; Pred. No. 6.5e-05;
Matches 56; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
Db 562 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 505

RESULT 26
US-11-266-748A-45850/c
; Sequence 45850, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 45850
; LENGTH: 738
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-45850

Query Match 4.2%; Score 54.8; DB 9; Length 738;
Best Local Similarity 96.6%; Pred. No. 6.8e-05;
Matches 56; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
Db 648 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 591

RESULT 27
US-09-831-621A-14
; Sequence 14, Application US/09831621A
; GENERAL INFORMATION:
; APPLICANT: Baum, Peter
; APPLICANT: DuBose, Robert
; APPLICANT: Sims, John E.
; APPLICANT: Youakim, Adel
; APPLICANT: Hasel, Karl W.
; APPLICANT: Hilbush, Brian S.
; TITLE OF INVENTION: Novel DNAs and Polypeptides
; FILE REFERENCE: 98664-A
; CURRENT APPLICATION NUMBER: US/09/831,621A
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/107821
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-831-621A-14

Query Match 4.2%; Score 54.8; DB 5; Length 1044;
Best Local Similarity 96.6%; Pred. No. 7.8e-05;
Matches 56; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
Db 3 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 60

RESULT 28
US-10-461-673-1880
; Sequence 1880, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yunqing
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Ghosh, Malabika
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001

```
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 1880
;    LENGTH: 3289
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-461-673-1880

  Query Match             4.2%;  Score 54.8;  DB 6;  Length 3289;
  Best Local Similarity   96.6%;  Pred. No. 0.00012;
  Matches   56;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 430


RESULT 29
US-11-266-748A-219976/c
; Sequence 219976, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 219976
;    LENGTH: 661
;    TYPE: DNA
;    ORGANISM: Homo Sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (525)..(525)
;    OTHER INFORMATION: n is a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (596)..(596)
;    OTHER INFORMATION: n is a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (627)..(627)
;    OTHER INFORMATION: n is a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (642)..(642)
;    OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-219976

  Query Match             4.1%;  Score 53.8;  DB 9;  Length 661;
  Best Local Similarity   94.8%;  Pred. No. 0.00012;
  Matches   55;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        579 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACNAGG 522


RESULT 30
US-11-266-748A-239716
; Sequence 239716, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 239716
;    LENGTH: 661
;    TYPE: DNA
;    ORGANISM: Homo Sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (20)..(20)
;    OTHER INFORMATION: n is a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (35)..(35)
;    OTHER INFORMATION: n is a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (66)..(66)
;    OTHER INFORMATION: n is a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (137)..(137)
;    OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-239716

  Query Match             4.1%;  Score 53.8;  DB 9;  Length 661;
  Best Local Similarity   94.8%;  Pred. No. 0.00012;
  Matches   55;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db         83 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACNAGG 140
```

```
RESULT 31
US-60-762-056-35503/c
; Sequence 35503, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 35503
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-35503

  Query Match             4.1%;  Score 53.8;  DB 11;  Length 2300;
  Best Local Similarity   58.4%;  Pred. No. 0.0002;
  Matches   94;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |  |  || |   | | ||  |||     |||||| | |||||
Db       1900 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 1841

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db       1840 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 1781

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGC 331
              |||||   |||||||||||| ||| | |   |  | |  ||
Db       1780 TCCTGGCCGAGAACGACTGGGAGATGGAAGTACGCTTTCGC 1740


RESULT 32
US-60-762-056-35504
; Sequence 35504, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 35504
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-35504

  Query Match             4.1%;  Score 53.6;  DB 11;  Length 2300;
  Best Local Similarity   60.1%;  Pred. No. 0.00023;
  Matches   89;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |  |  || |   | | ||  |||     |||||| | |||||
Db       2142 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 2201

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db       2202 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 2261

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCA 318
              |||||   |||||||||||| ||| | |
Db       2262 TCCTGGCCGAGAACGACTGGGAGATGGA 2289


RESULT 33
US-11-266-748A-219957/c
; Sequence 219957, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 219957
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (89)..(89)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (500)..(500)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-219957

  Query Match             4.0%;  Score 52.2;  DB 9;  Length 600;
  Best Local Similarity   93.1%;  Pred. No. 0.00033;
  Matches   54;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | |||||||||||| |||||||||||||||||||||||||||||||  |||||||||
Db        514 ATTTTATACGACTCNCTATAGGGAATTTGGCCCTCGAGGCCAAGAATCGGGCACGAGG 457


RESULT 34
US-60-762-056-20736
; Sequence 20736, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 20736
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-60-762-056-20736

Query Match 4.0%; Score 52.2; DB 11; Length 2300;
Best Local Similarity 68.6%; Pred. No. 0.00055;
Matches 72; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy        141 CCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCAC 200
              | |  ||||||| | | || | | |||| | |  ||||||| ||||| | |  |||||
Db       2109 CGGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCAGCGG 2168

Qy        201 AGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
               |||||    || ||||| | || |||||||||  | | ||||||
Db       2169 CGGCGGCCTGGGCCCGGGCGCAGCGGCGGCGGCGGCGGGGCCTGG 2213
```

RESULT 35
US-11-283-329-179
; Sequence 179, Application US/11283329
; GENERAL INFORMATION:
; APPLICANT: Piu, Fabrice
; TITLE OF INVENTION: ENABLING TOOLS TO IDENTIFY LIGANDS FOR
; TITLE OF INVENTION: HORMONE NUCLEAR RECEPTORS
; FILE REFERENCE: ACADIA.043A
; CURRENT APPLICATION NUMBER: US/11/283,329
; CURRENT FILING DATE: 2005-11-18
; PRIOR APPLICATION NUMBER: 60/629,811
; PRIOR FILING DATE: 2004-11-19
; NUMBER OF SEQ ID NOS: 242
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 179
; LENGTH: 2968
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)...(1953)
; OTHER INFORMATION: CARM1
US-11-283-329-179

Query Match 4.0%; Score 52.2; DB 7; Length 2968;
Best Local Similarity 68.6%; Pred. No. 0.00061;
Matches 72; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy        141 CCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCAC 200
              | |  ||||||| | | || | | |||| | |  ||||||| ||||| | |  |||||
Db          9 CGGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCAGCGG 68

Qy        201 AGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
               |||||    || ||||| | || |||||||||  | | ||||||
Db         69 CGGCGGCCTGGGCCCGGGCGCAGCGGCGGCGGCGGCGGGGCCTGG 113
```

RESULT 36
US-11-266-748A-267412
; Sequence 267412, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 267412
; LENGTH: 971
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-267412

Query Match 3.9%; Score 51.4; DB 9; Length 971;
Best Local Similarity 98.1%; Pred. No. 0.00066;
Matches 52; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         23 TACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db          1 TACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGTG 53
```

RESULT 37
US-11-266-748A-327929/c
; Sequence 327929, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 327929
; LENGTH: 971
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-327929

Query Match 3.9%; Score 51.4; DB 9; Length 971;
Best Local Similarity 98.1%; Pred. No. 0.00066;
Matches 52; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         23 TACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        971 TACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGTG 919
```

RESULT 38
US-11-266-748A-102287/c
; Sequence 102287, Application US/11266748A
; GENERAL INFORMATION:

```
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 102287
;   LENGTH: 890
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (59)..(80)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (214)..(236)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-102287

  Query Match             3.9%;  Score 51.2;  DB 9;  Length 890;
  Best Local Similarity   81.9%;  Pred. No. 0.00072;
  Matches   59;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         13 TTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGA 72
              | |||||||||||||||| |||||||||| |||||||||| |||||||||| | ||
Db        789 TCCGAATTTATACGACTCCCTATAGGGAATCTGGCCCTCGAGTCCAAGAATTCGCCCCGC 730

Qy         73 GGGCGGGAAGCA 84
              ||   | | | |
Db        729 GGAAAGAACGAA 718


RESULT 39
US-11-266-748A-155098
; Sequence 155098, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 155098
;   LENGTH: 890
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (655)..(677)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (811)..(832)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-155098

  Query Match             3.9%;  Score 51.2;  DB 9;  Length 890;
  Best Local Similarity   81.9%;  Pred. No. 0.00072;
  Matches   59;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         13 TTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGA 72
              | |||||||||||||||| |||||||||| |||||||||| |||||||||| | ||
Db        102 TCCGAATTTATACGACTCCCTATAGGGAATCTGGCCCTCGAGTCCAAGAATTCGCCCCGC 161

Qy         73 GGGCGGGAAGCA 84
              ||   | | | |
Db        162 GGAAAGAACGAA 173


RESULT 40
US-10-531-147-349
; Sequence 349, Application US/10531147
; GENERAL INFORMATION:
;  APPLICANT: THOMAS WU
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF TUMOR
;  FILE REFERENCE: P5102R1-US
;  CURRENT APPLICATION NUMBER: US/10/531,147
;  CURRENT FILING DATE:  2005-04-12
;  PRIOR APPLICATION NUMBER: PCT/US03/029126
;  PRIOR FILING DATE: 2003-10-15
;  PRIOR APPLICATION NUMBER: US 60/418,988
;  PRIOR FILING DATE: 2002-10-18
;  NUMBER OF SEQ ID NOS: 4622
; SEQ ID NO 349
;   LENGTH: 2954
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-531-147-349

  Query Match             3.9%;  Score 51;  DB 6;  Length 2954;
  Best Local Similarity   69.7%;  Pred. No. 0.0013;
  Matches   69;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy        147 CGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGG 206
              |||||| | | || | | |||| | |  ||||||| ||||| | |  |||||   |||||
Db          1 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCAGCGGCGGCGG 60

Qy        207 AGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
                  || |||||| | || ||||||||||  | | ||||||
Db         61 CCTGGGCCCGGGCGCAGCGGCGGCGGCGGCGGGGCCTGG 99


RESULT 41
US-10-531-147-1529
```

```
; Sequence 1529, Application US/10531147
; GENERAL INFORMATION:
;  APPLICANT: THOMAS WU
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF TUMOR
;  FILE REFERENCE: P5102R1-US
;  CURRENT APPLICATION NUMBER: US/10/531,147
;  CURRENT FILING DATE:  2005-04-12
;  PRIOR APPLICATION NUMBER: PCT/US03/029126
;  PRIOR FILING DATE: 2003-10-15
;  PRIOR APPLICATION NUMBER: US 60/418,988
;  PRIOR FILING DATE: 2002-10-18
;  NUMBER OF SEQ ID NOS: 4622
; SEQ ID NO 1529
;   LENGTH: 2954
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-531-147-1529

  Query Match             3.9%;  Score 51;  DB 6;  Length 2954;
  Best Local Similarity   69.7%;  Pred. No. 0.0013;
  Matches   69;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy        147 CGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGG 206
              |||||| | | || | | |||| | |  ||||||| ||||| | |  |||||   |||||
Db          1 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCAGCGGCGGCGG 60

Qy        207 AGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
              || ||||| | || |||||||||  | | ||||||
Db         61 CCTGGGCCCGGGCGCAGCGGCGGCGGCGGCGGGGCCTGG 99


RESULT 42
US-11-227-183A-4003/c
; Sequence 4003, Application US/11227183A
; GENERAL INFORMATION:
;  APPLICANT: Fisher, Dane K.
;  APPLICANT:  Lalgudi, Raghunath V.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.336 - 38-21(51230)C/US
;  CURRENT APPLICATION NUMBER: US/11/227,183A
;  CURRENT FILING DATE:  2005-09-16
;  PRIOR APPLICATION NUMBER: US 09/619,643
;  PRIOR FILING DATE: 2000-07-19
;  NUMBER OF SEQ ID NOS: 32236
; SEQ ID NO 4003
;   LENGTH: 602
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3116-021-P1-K1-D8-TCA
US-11-227-183A-4003

  Query Match             3.8%;  Score 49.6;  DB 9;  Length 602;
  Best Local Similarity   52.0%;  Pred. No. 0.0017;
  Matches  106;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy         66 GGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGG 125
              ||   | |||||||  || | | |  | | ||| |   ||  | |     | |||||  ||
Db        568 GGGGGGCGGGCGGGCNGCCGGGCGCCGCGGGGGGGCGGTGGCGTGGGGGTTCGGCGGGGG 509

Qy        126 CGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGG 185
              |   ||| |  |   |  ||||    |  ||  ||| ||  || ||| || || |
Db        508 GGGGGGGCGGGGGGGGCCGGGCGGGTNGGGGCGGGGGGGGGGGGCGGGGGGGGGGTGGT 449

Qy        186 CGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
              |   |  ||      ||||| || ||   ||| |    ||||| ||| | | | | ||
Db        448 GGGGGGGGGCGCGGGGGGCGGCGGGGGGGGGGGGGCCGGGCGGGGGCNNCGGCGGCGGG 389

Qy        246 GCTTTGCGTTGGTGGGGGGATGCG 269
              |   |||  |  ||||||   ||
Db        388 GGCGCGCGCGGCGGGGGGGCCCCG 365


RESULT 43
US-11-329-388-9794
; Sequence 9794, Application US/11329388
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Scott E.
;  APPLICANT:  Fisher, Dane K.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.357 - 38-21(15503)C/US
;  CURRENT APPLICATION NUMBER: US/11/329,388
;  CURRENT FILING DATE:  2006-01-11
;  PRIOR APPLICATION NUMBER: 09/553,094
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS: 18831
; SEQ ID NO 9794
;   LENGTH: 555
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(555)
;   OTHER INFORMATION: unsure at all n locations
;   OTHER INFORMATION: Clone ID: LIB3061-094-Q1-N6-B12
US-11-329-388-9794

  Query Match             3.8%;  Score 49.4;  DB 7;  Length 555;
  Best Local Similarity   53.6%;  Pred. No. 0.0019;
  Matches  125;  Conservative    0;  Mismatches  106;  Indels    2;  Gaps    1;

Qy         37 GGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCG 96
              |||   |||||  | ||    |   ||||  |  || |||  |  | | |  | | |
Db        261 GGGCGGGTGGCCGGGGGGGGGGGGGCGGCGGCCGGTCGGGGGGGGGGTGGGGCGGCGGGGG 320

Qy         97 GGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCC 156
              |  |    | |||    |||||| |  || ||||||  |  | |  |  | || || |
Db        321 GCGGGGCGGGGGGCCGGCTGCGGGGGGGGGGGTCTGGGGGGGGGTGGGGGGGGGGGGGGT 380

Qy        157 CGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCG 216
              |  ||| |||| ||||||| || |  || | | |  |  |    || || || ||   |
Db        381 GG--GGGGCGGCGGGGCGGGGGGGGGGGGGGTCGGGGGGGGGGGGGGGGGCGGGGGGGG 438

Qy        217 GGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCG 269
              || |  | || || |||    | | | |||   || |  || |||||  |||
Db        439 GGGGGGGGGGGGGGGGCGGGGGGGCGGGGGGGTGTGGGGGGGGGGGGGGCG 491


RESULT 44
US-60-762-056-6363/c
; Sequence 6363, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 6363
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-6363

  Query Match             3.7%;  Score 49;  DB 11;  Length 2300;
  Best Local Similarity   57.5%;  Pred. No. 0.0042;
  Matches   88;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;
```

```
Qy          77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
               || | ||   | | ||  ||| ||        |||| | ||||| | |   |   |||
Db        1863 GGAAGGCTCGGCGGAGTCCGGATGGAGGACTCGGACTCGGCGGCAAAGCAGCTGGGCCTG 1804

Qy         137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
                 |    | ||||||| | | || | | | || | |  ||||||| | ||| |||  |||
Db        1803 GCTGAGGCGGCGGCGGTGGCGGCCGCGGCCGCTGTGGCGGCGGCGGCCGCGGCCGCGGCA 1744

Qy         197 GCACAGGCGGAGGAGGACCGGGTGAAGAGGCGG 229
               | |   | || |||||| | ||||  |||  ||
Db        1743 GGAGGCGAGGCGGAGGAGCCGGTGCTGAGCAGG 1711
```

RESULT 45
US-11-266-748A-30544
; Sequence 30544, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30544
; LENGTH: 2737
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-30544

Query Match 3.7%; Score 49; DB 9; Length 2737;
Best Local Similarity 57.5%; Pred. No. 0.0045;
Matches 88; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```
Qy          77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
               || | ||   | | ||  ||| ||        |||| | ||||| | |   |   |||
Db         673 GGAAGGCTCGGCGGAGTCCGGATGGAGGACTCGGACTCGGCGGCAAAGCAGCTGGGCCTG 732

Qy         137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
                 |    | ||||||| | | || | | | || | |  ||||||| | ||| |||  |||
Db         733 GCTGAGGCGGCGGCGGTGGCGGCCGCGGCCGCTGTGGCGGCGGCGGCCGCGGCCGCGGCA 792

Qy         197 GCACAGGCGGAGGAGGACCGGGTGAAGAGGCGG 229
               | |   | || |||||| | ||||  |||  ||
Db         793 GGAGGCGAGGCGGAGGAGCCGGTGCTGAGCAGG 825
```

RESULT 46
US-11-292-078-5998
; Sequence 5998, Application US/11292078
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.346 - 38-21(52274)C
; CURRENT APPLICATION NUMBER: US/11/292,078
; CURRENT FILING DATE: 2005-12-02
; PRIOR APPLICATION NUMBER: US 10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255,619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 5998
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-011-Q1-N6-H6
US-11-292-078-5998

Query Match 3.7%; Score 48.6; DB 9; Length 420;
Best Local Similarity 52.8%; Pred. No. 0.0028;
Matches 105; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```
Qy          66 GGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGG 125
               ||   | |||||||  |  | | |  | | ||| |    | |||| ||| | || |   |
Db         222 GGGGGGGGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGCCCCGGGGGGCCCG 281

Qy         126 CGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGG 185
               |   | | |  |     | |||  ||| |   |   ||  ||  ||| ||| |||   ||
Db         282 GGGGGGCCCGGGGGGGGGCTGCTTCGGGGGAGGGGCGGGGGGGGGGGGGGGGCCGGGG 341

Qy         186 CGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
               | |||  |  |  | || || |  ||   ||| |  | | | | ||     | |||  |
Db         342 GGGCCGGGGGGGGCCCGGGGGGCGGGGGGGGGGGGGGGCCCGTGGTGGGCGGGCCCCG 401

Qy         246 GCTTTGCGTTGGTGGGGGG 264
               |||  |||  || ||||||
Db         402 GCTGGGCGGGGGGGGGGG 420
```

RESULT 47
US-60-762-056-4866/c
; Sequence 4866, Application US/60762056
; GENERAL INFORMATION:
; APPLICANT: Trinklein, Nathan D
; APPLICANT: Aldred, Shelley Force
; APPLICANT: Cooper, Sara J
; APPLICANT: Myers, Richard M
; TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
; TITLE OF INVENTION: EXPRESSION REGULATORY ELEMENTS
; FILE REFERENCE: 33102-701.102
; CURRENT APPLICATION NUMBER: US/60/762,056
; CURRENT FILING DATE: 2006-01-24
; NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 4866
; LENGTH: 2300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-762-056-4866

Query Match 3.7%; Score 48.6; DB 11; Length 2300;
Best Local Similarity 50.2%; Pred. No. 0.0055;
Matches 147; Conservative 0; Mismatches 144; Indels 2; Gaps 1;

```
Qy         105 GAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGC 164
               | ||||  |  ||||||  ||||| |||  | ||  |||      ||  |||    | |
Db        2211 GCGGGGCGCGGGCGGCGGGGGCGTGTGG--GGAGGCCCGCCCGCTCGCCGCCACGGGAGG 2154

Qy         165 CGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGA 224
               ||| |  |||  | || | ||       ||||||   ||||| || ||  |       |
Db        2153 GGGCGGAACGGCGGCCGGCCGCCGGGTCAGCAGCGGCGGCGGCGGCGGGGCCTCGCCATG 2094
```

```
Qy        225 GGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCC 284
              |||||||||  | | | | | |    |||   |  ||      ||  |||| ||    | ||
Db       2093 GGCGGCGGCGGCGGCGGCGGCGACGAGCGGGCGACGGCCCCGGCTCCCCCTCGGCCGGCC 2034

Qy        285 CCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGC 344
                ||    |  ||||  |    | | |     |    | ||   |||||||| ||   |
Db       2033 TATCCGCGGCGGAGGAGCGCTCGGGTCCGCGCGACTGCGGCCGCCGCGCCTACCTCACCC 1974

Qy        345 TGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCC 397
              |||   |      |  | |  |||||||  ||| ||  |  ||  |   |||
Db       1973 CGCCGCGGCCATGGCGGGCGGCGCCGCGCAGGCCGCCTCCCCCCCGCGCTTCC 1921
```

RESULT 48
US-11-329-388-8161
; Sequence 8161, Application US/11329388
; GENERAL INFORMATION:
; APPLICANT: Andersen, Scott E.
; APPLICANT: Fisher, Dane K.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.357 - 38-21(15503)C/US
; CURRENT APPLICATION NUMBER: US/11/329,388
; CURRENT FILING DATE: 2006-01-11
; PRIOR APPLICATION NUMBER: 09/553,094
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18831
; SEQ ID NO 8161
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(627)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3061-085-Q1-N6-E12
US-11-329-388-8161

Query Match 3.7%; Score 48.4; DB 7; Length 627;
Best Local Similarity 52.0%; Pred. No. 0.0038;
Matches 106; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```
Qy         66 GGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGG 125
              |||  | |||||||  |  | |||   | |||  |   | ||||    | ||  |  ||
Db        389 GGCGGGGGGGCGGGTCGGGGGGTGGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 448

Qy        126 CGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGG 185
              |   ||  |  |     |  |||| || |   || |||  ||  ||| |||||| |  ||
Db        449 GGGGGGGGGGGGGGGGGGGGGCGGGGTGGGGGCGGGGGGGGGGGGGNGGGCGGGGGGGG 508

Qy        186 CGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
               | | |  || ||   || || || ||   ||| |  |  | || ||     | |   ||
Db        509 GGGCGGGGGCGGCGGGGGGGGGGNGGGGGGGGGGGGGCGGGGGGGGGGGGGGGGGGGG 568

Qy        246 GCTTTGCGTTGGTGGGGGGATGCG 269
              |    |||  ||  |||||  | |
Db        569 GGGGGGCGGGGGGCGGGGGGGGG 592
```

RESULT 49
US-11-329-388-8142
; Sequence 8142, Application US/11329388
; GENERAL INFORMATION:
; APPLICANT: Andersen, Scott E.
; APPLICANT: Fisher, Dane K.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.357 - 38-21(15503)C/US
; CURRENT APPLICATION NUMBER: US/11/329,388
; CURRENT FILING DATE: 2006-01-11
; PRIOR APPLICATION NUMBER: 09/553,094
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18831
; SEQ ID NO 8142
; LENGTH: 398
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(398)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3061-085-Q1-N6-B12
US-11-329-388-8142

Query Match 3.7%; Score 48.2; DB 7; Length 398;
Best Local Similarity 50.2%; Pred. No. 0.0036;
Matches 119; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

```
Qy         65 CGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATG 124
              |||| || ||| |||  |  | | |  | | ||| | | | ||||      | || | |
Db         18 CGGCGCGGGGGGGGGGGGGGGGTGGGCGGGGGGGGGGGTCGGGGGGGGGGGGGGGGGGGG 77

Qy        125 GCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCG 184
              | | | ||  |  |     |  | || || |   |  ||| ||  ||| ||| || || |
Db         78 GGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGTGG 137

Qy        185 GCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTG 244
              | ||  |  |         ||||| || || | ||| |  |  | || ||      | |   |
Db        138 GGGCGGGGGGTGCGGGCGGCGGGGGGGGGCGGGGGGGGGGCGTGGGGGGGGGGGGGGGCG 197

Qy        245 GGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAG 301
              |||   | |  || ||||||  | |    |  ||    || |   | |||  |
Db        198 GGCGGGGGGGGGGGGGGGGGGGGGGGGGCGGGGGGGGGGGGGAGGGGGGGGGGGGG 254
```

RESULT 50
US-11-311-886-3337
; Sequence 3337, Application US/11311886
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 16517.358 - 38-21(15598)C/US
; CURRENT APPLICATION NUMBER: US/11/311,886
; CURRENT FILING DATE: 2005-12-20
; PRIOR APPLICATION NUMBER: 09/421,106
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 36938
; SEQ ID NO 3337
; LENGTH: 799
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: unsure at all n locations
US-11-311-886-3337

Query Match 3.7%; Score 48.2; DB 9; Length 799;
Best Local Similarity 50.9%; Pred. No. 0.0047;
Matches 112; Conservative 0; Mismatches 107; Indels 1; Gaps 1;

```
Qy         45 GGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTT 104
              |||   | |  ||       |||   | || ||||  |  ||| |  | | ||| |
Db        208 GGCGGGCTACTCCGNNNGCGCGGGGNGCGGNCGGGGGGGGGCGCGCNGGGGGGGGGGGGG 267

Qy        105 GAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGC 164
              | |||| |   |||||      ||   ||      |   |  | ||| |  | ||  |||
Db        268 GGGGGGGCGGGGCGGC-NCCCCGGGGGGGGGNNCGGGGGGGGGGGCCGNCCNCGGGGGGG 326

Qy        165 CGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGA 224
              ||  ||| ||| | | | |||   |  |  ||   || || || || ||||| | |
Db        327 GGGGGGGGGGGGNNGGGGCNGCGGGGGGGGGGGCNNCGGGGGCGGGGGGCCGGGGGGGGC 386

Qy        225 GGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGG 264
              || |||||   | | |   |||    |||   | ||||||
Db        387 GGGGGCGGGGCCCGGGGGGGGGGGGGNGCGGGNGNGGGGGG 426
```

```
RESULT 51
US-11-266-748A-40236
; Sequence 40236, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 40236
;   LENGTH: 1136
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (19)..(19)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (49)..(49)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-40236

  Query Match             3.7%;  Score 48.2;  DB 9;  Length 1136;
  Best Local Similarity   54.6%;  Pred. No. 0.0054;
  Matches   95;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy        132 GCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCG 191
              ||||| |   |   ||||||| ||  |  | | |||| | | | | | ||  |||   |
Db         28 GCAGCCGCAGCCGGGCGGCGGNGCGAGAGGCGGCGGCGGCGGGCGGGCCGCGGGCAGTCA 87

Qy        192 AAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTG 251
                ||   ||||| || ||     || |  || ||||||||  |  ||  |  ||
Db         88 GTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGATGCGGCGGCCCCGCTGAGTCCGCCCGC 147

Qy        252 CGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACG 305
                ||| |  ||||  |  || | ||||||   || | | |  |||| || | |
Db        148 TCCTGGCGCCGGGAGCCAGCCGCGCGATGGCGGCCCGGGCCGGGCGGCAGCATG 201


RESULT 52
US-10-953-349-34914
; Sequence 34914, Application US/10953349
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCONDED THERBY
;  FILE REFERENCE: 2750-1579PUS2
;  CURRENT APPLICATION NUMBER: US/10/953,349
;  CURRENT FILING DATE:  2004-09-30
;  NUMBER OF SEQ ID NOS: 40252
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34914
;   LENGTH: 1109
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
US-10-953-349-34914

  Query Match             3.6%;  Score 47.8;  DB 6;  Length 1109;
  Best Local Similarity   50.2%;  Pred. No. 0.0069;
  Matches  118;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy        124 GGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTC 183
              || |||    | | ||  ||   | || |   |||| | | || | |||| |  | || |
Db        292 GGAGTCGCAAACCGGTAACGTCCCCGCCGCATCCGCGGCGACGCCGGGGCCGAAGCCGGC 351

Qy        184 GGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCT 243
              | || | |  || |   | |  | ||||| || |     |  |||||||| ||||
Db        352 GTCGGCGGGGGCCGGGAAAGGCGCGGAGGGCCAGTCTGTGGTGCGGCGGCTGCAGTCGGA 411

Qy        244 GGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAA 303
              |    | ||| || ||  ||| |||||||      ||    |      |  | |||| ||
Db        412 GCTGATGGCGCTGATGATGGGCGGCGACCCGGGCGTGTCGGCGTTTCCCGAGGGGGACAA 471

Qy        304 CGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGAC 358
              |      ||   |  |    || |   || || |  || |  | |  | | | |
Db        472 CATGCTCCACTGGGTGGGCACCATCGACGGCTCCGCCGGGACGGCCTACGAGGGC 526


RESULT 53
US-11-266-748A-69859
; Sequence 69859, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 69859
;   LENGTH: 1183
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (472)..(498)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (959)..(1050)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1056)..(1056)
```

```
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-69859

  Query Match             3.6%;  Score 47.6;  DB 9;  Length 1183;
  Best Local Similarity   92.4%;  Pred. No. 0.008;
  Matches   61;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          9 ATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 68
              ||||| |||||| |||||||||||||||| ||||| |||||||||||||||| |||||||||
Db        221 ATGATACGAATTAATACGACTCACTATA-GGAATGTGGCCCTCGAGGCCAGGAATTCGGC 279

Qy         69 ACGAGG 74
              ||||||
Db        280 ACGAGG 285


RESULT 54
US-11-266-748A-105667
; Sequence 105667, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 105667
;   LENGTH: 1183
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (472)..(498)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (959)..(1050)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1056)..(1056)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-105667

  Query Match             3.6%;  Score 47.6;  DB 9;  Length 1183;
  Best Local Similarity   92.4%;  Pred. No. 0.008;
  Matches   61;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          9 ATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 68
              ||||| |||||| |||||||||||||||| ||||| |||||||||||||||| |||||||||
Db        221 ATGATACGAATTAATACGACTCACTATA-GGAATGTGGCCCTCGAGGCCAGGAATTCGGC 279

Qy         69 ACGAGG 74
              ||||||
Db        280 ACGAGG 285


RESULT 55
US-11-266-748A-122670/c
; Sequence 122670, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 122670
;   LENGTH: 1183
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (128)..(128)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (134)..(225)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (686)..(712)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-122670

  Query Match             3.6%;  Score 47.6;  DB 9;  Length 1183;
  Best Local Similarity   92.4%;  Pred. No. 0.008;
  Matches   61;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          9 ATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 68
              ||||| |||||| |||||||||||||||| ||||| |||||||||||||||| |||||||||
Db        963 ATGATACGAATTAATACGACTCACTATA-GGAATGTGGCCCTCGAGGCCAGGAATTCGGC 905

Qy         69 ACGAGG 74
              ||||||
Db        904 ACGAGG 899


RESULT 56
US-11-266-748A-69138
; Sequence 69138, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
```

```
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 69138
;   LENGTH: 2080
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (101)..(157)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (264)..(284)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (761)..(769)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1015)..(1031)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-69138

  Query Match             3.6%;  Score 47.6;  DB 9;  Length 2080;
  Best Local Similarity   92.4%;  Pred. No. 0.01;
  Matches   61;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          9 ATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 68
              ||||| |||||| |||||||||| |||| |||||||||||||||| ||||||||||||
Db        160 ATGATACGAATTAATACGACTCA-TATAAGGAATTTGGCCCTCGAGACCAAGAATTCGGC 218

Qy         69 ACGAGG 74
              ||||||
Db        219 ACGAGG 224


RESULT 57
US-11-266-748A-121949/c
; Sequence 121949, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 121949
;   LENGTH: 2080
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1050)..(1066)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1312)..(1320)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1797)..(1817)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1924)..(1980)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-121949

  Query Match             3.6%;  Score 47.6;  DB 9;  Length 2080;
  Best Local Similarity   92.4%;  Pred. No. 0.01;
  Matches   61;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          9 ATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 68
              ||||| |||||| |||||||||| |||| |||||||||||||||| ||||||||||||
Db       1921 ATGATACGAATTAATACGACTCA-TATAAGGAATTTGGCCCTCGAGACCAAGAATTCGGC 1863

Qy         69 ACGAGG 74
              ||||||
Db       1862 ACGAGG 1857


RESULT 58
US-11-353-150-5643
; Sequence 5643, Application US/11353150
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Scott E.
;  APPLICANT:  Conner, Timothy W.
;  APPLICANT:  Lalgudi, Raghunath V.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.365 - 38-21(15459)D/US
;  CURRENT APPLICATION NUMBER: US/11/353,150
;  CURRENT FILING DATE:  2006-02-14
;  PRIOR APPLICATION NUMBER: 09/565,386
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/132,860
;  PRIOR FILING DATE: 1999-05-07
;  NUMBER OF SEQ ID NOS: 83523
; SEQ ID NO 5643
;   LENGTH: 452
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB148-035-Q1-E1-B5
US-11-353-150-5643
```

```
  Query Match             3.6%;  Score 47.2;  DB 7;  Length 452;
  Best Local Similarity   52.0%;  Pred. No. 0.0071;
  Matches  106;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy         66 GGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGG 125
              ||    | ||| |||  || | | |  | | ||| |   || ||||     | ||  |  ||
Db        140 GGGGGGGGGGGGGGGCGCTGGGGGGGGGGGGGGGGGGGGTGGGGGGTGGGGGGGGGGTGGG 199

Qy        126 CGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGG 185
              |     |  |  |  ||| | || || |    |  |||| ||  ||| ||||||  |  ||
Db        200 GGGGATGGGGGGGGGCCGGGGGGGGGGGGGGGGGGGGGCGGGGTGGGGGGGCGGGGGGGG 259

Qy        186 CGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
              |    |   |  ||   ||||| || ||  |||| |  | || || ||     | |   ||
Db        260 GGGTGGGGGGGGCTGGGGCGGGGGGGGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGG 319

Qy        246 GCTTTGCGTTGGTGGGGGGATGCG 269
              |    |  | || | ||||  | |
Db        320 GGGGGGGCTGGGGGTGGGGGGGGG 343


RESULT 59
US-11-266-748A-219917/c
; Sequence 219917, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 219917
;   LENGTH: 748
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-219917

  Query Match             3.6%;  Score 47;  DB 9;  Length 748;
  Best Local Similarity   100.0%;  Pred. No. 0.0098;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 CTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        681 CTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 635


RESULT 60
US-11-266-748A-239684
; Sequence 239684, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 239684
;   LENGTH: 748
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-239684

  Query Match             3.6%;  Score 47;  DB 9;  Length 748;
  Best Local Similarity   100.0%;  Pred. No. 0.0098;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 CTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              |||||||||||||||||||||||||||||||||||||||||||||||
Db         68 CTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 114


RESULT 61
US-11-266-748A-82366
; Sequence 82366, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82366
;   LENGTH: 1077
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (996)..(1014)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-82366

  Query Match             3.6%;  Score 46.8;  DB 9;  Length 1077;
  Best Local Similarity   87.9%;  Pred. No. 0.013;
  Matches   51;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || |||  ||||||| |||||||||||||||||| | ||||||||||||||||||
Db        323 ATTTAATAGCACTCACTGTAGGGAATTTGGCCCTCGGGACCAAGAATTCGGCACGAGG 380


RESULT 62
US-11-266-748A-111289
; Sequence 111289, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 111289
;   LENGTH: 1077
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (996)..(1014)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-111289

  Query Match             3.6%;  Score 46.8;  DB 9;  Length 1077;
  Best Local Similarity   87.9%;  Pred. No. 0.013;
  Matches   51;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || |||  ||||||| |||||||||||||||||| | ||||||||||||||||||
Db        323 ATTTAATAGCACTCACTGTAGGGAATTTGGCCCTCGGGACCAAGAATTCGGCACGAGG 380


RESULT 63
US-11-266-748A-135177/c
; Sequence 135177, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 135177
;   LENGTH: 1077
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (64)..(82)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-135177

  Query Match             3.6%;  Score 46.8;  DB 9;  Length 1077;
  Best Local Similarity   87.9%;  Pred. No. 0.013;
  Matches   51;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
              | || |||  ||||||| |||||||||||||||||| | ||||||||||||||||||
Db        755 ATTTAATAGCACTCACTGTAGGGAATTTGGCCCTCGGGACCAAGAATTCGGCACGAGG 698


RESULT 64
US-60-762-056-19219
; Sequence 19219, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 19219
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-19219

  Query Match             3.6%;  Score 46.8;  DB 11;  Length 2300;
  Best Local Similarity   46.8%;  Pred. No. 0.017;
  Matches  147;  Conservative    0;  Mismatches  167;  Indels    0;  Gaps    0;

Qy         75 GCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCA 134
              | | |  || |   | | |  || |     | | ||  || ||||||  |||| | ||||
Db       1828 GGGCGGGGCCGGCGGCGGCGGCGGCGGGCAGGGCGGGGCCGGCGGCGGCGGCGGCGGGCA 1887

Qy        135 GCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAG 194
              |          ||||||| | | || ||   ||| |||| ||||||| ||||| | |
Db       1888 GGGCGGGGCCGGCGGCGGCGGCGGCGGGCAGGGCGGGGCCGGCGGCGGCGGCGGCGGGCA 1947

Qy        195 CAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGT 254
```

```
                ||   | ||| || ||    || |   |||  | |||    |     ||     |
Db       1948 GGGCGGGGCCGGCGGCGGCGGCGGCGGGCAGGGCGGGGCCGGCGGCGGCGGCGGCGGGCA 2007

Qy        255 TGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGA 314
              || ||||       |   |  ||  || |      | | | || |   |     |   |
Db       2008 GGGCGGGGCCGGCGGCGGCGGCGGCGGGCAGGGCGGGGCCGGCGGCGGCGGCGGCGGGCA 2067

Qy        315 CGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCC 374
              |  |    ||  |  | | |    |||| || | ||  ||    |||  |   || |
Db       2068 GGGCGGGGCCCAGGGCCCCGTGACCCGAGGGGCGACAGCGCGCGGCAGGCGGCGAGCTCG 2127

Qy        375 AGCCTCCCACGTCC 388
              |  | ||  || |
Db       2128 GGGGCCGCAGGTGC 2141
```

RESULT 65
US-11-266-748A-70423/c
; Sequence 70423, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 70423
; LENGTH: 1171
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (38)..(38)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (89)..(89)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (97)..(110)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (942)..(942)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-70423

Query Match 3.6%; Score 46.6; DB 9; Length 1171;
Best Local Similarity 85.1%; Pred. No. 0.015;
Matches 63; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              || |    ||| | |||||| |||||||||||||||| |||||||||||||||||  ||||
Db        944 AGNTTGACATGTTACGAATTAATACGACTCACTATA-GGAATTTGGCCCTCGAGCACAAG 886

Qy         61 AATTCGGCACGAGG 74
              ||||||||||||||
Db        885 AATTCGGCACGAGG 872
```

RESULT 66
US-11-266-748A-106000/c
; Sequence 106000, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 106000
; LENGTH: 1171
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (38)..(38)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (89)..(89)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (97)..(110)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (942)..(942)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-106000

Query Match 3.6%; Score 46.6; DB 9; Length 1171;
Best Local Similarity 85.1%; Pred. No. 0.015;
Matches 63; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
```

```
Db        944 AGNTTGACATGTTACGAATTAATACGACTCACTATA-GGAATTTGGCCCTCGAGCACAAG 886

Qy         61 AATTCGGCACGAGG 74
              ||||||||||||||
Db        885 AATTCGGCACGAGG 872
```

RESULT 67
US-11-266-748A-123234
; Sequence 123234, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 123234
; LENGTH: 1171
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (230)..(230)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1062)..(1075)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1083)..(1083)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1134)..(1134)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1158)..(1158)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-123234

Query Match 3.6%; Score 46.6; DB 9; Length 1171;
Best Local Similarity 85.1%; Pred. No. 0.015;
Matches 63; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              || |    ||| | |||||| ||||||||||||||| ||||||||||||||||  ||||
Db        228 AGNTTGACATGTTACGAATTAATACGACTCACTATA-GGAATTTGGCCCTCGAGCACAAG 286

Qy         61 AATTCGGCACGAGG 74
```

```
              ||||||||||||||
Db        287 AATTCGGCACGAGG 300
```

RESULT 68
US-11-266-748A-81563
; Sequence 81563, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81563
; LENGTH: 1088
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1009)..(1027)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-81563

Query Match 3.5%; Score 46.2; DB 9; Length 1088;
Best Local Similarity 86.4%; Pred. No. 0.019;
Matches 51; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||   |||||||||||||||||| ||||||| |||||||||||||||||| |
Db        240 ATTTAATCGAACTCACTATAGGGAATTTGCCCCTCGAGCCCAAGAATTCGGCACGAGAG 298
```

RESULT 69
US-11-266-748A-134374/c
; Sequence 134374, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03

```
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 134374
;   LENGTH: 1088
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (62)..(80)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-134374

  Query Match             3.5%;  Score 46.2;  DB 9;  Length 1088;
  Best Local Similarity   86.4%;  Pred. No. 0.019;
  Matches   51;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75
              | || ||   |||||||||||||||||||| |||||||| ||||||||||||||||||| |
Db        849 ATTTAATCGAACTCACTATAGGGAATTTGCCCCTCGAGCCCAAGAATTCGGCACGAGAG 791


RESULT 70
US-11-266-748A-47816
; Sequence 47816, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 47816
;   LENGTH: 823
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (34)..(34)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (817)..(817)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-47816

  Query Match             3.5%;  Score 45.8;  DB 9;  Length 823;
  Best Local Similarity   68.1%;  Pred. No. 0.022;
  Matches   62;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy         27 ACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGGGAAGCAGC 86
              |||  |||||||||||||||||||||||||||  |||||||||||| |     ||| |
Db          1 ACTNCCTATAGGGAATTTGGCCCTCGAGGCCAGNAATTCGGCACGAGCGGCACGAGCGGG 60

Qy         87 GTGAAGAGCGGGTGTTTTGAGGGGACCCTGC 117
                |  |||   |   |  |  | || | |||
Db         61 AGGTGGAGATTGCAGTAAGCCGAGATCGTGC 91


RESULT 71
US-11-330-082-8448/c
; Sequence 8448, Application US/11330082
; GENERAL INFORMATION:
;  APPLICANT: Buehler, Robert E.
;  APPLICANT:  Byrum, Joseph R.
;  APPLICANT:  Coombs, Brian E.
;  APPLICANT:  La Rosa, Thomas J.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.356 - 38-21(15760)C/US
;  CURRENT APPLICATION NUMBER: US/11/330,082
;  CURRENT FILING DATE:  2006-01-12
;  PRIOR APPLICATION NUMBER: US 09/552,086
;  PRIOR FILING DATE: 2000-04-19
;  NUMBER OF SEQ ID NOS: 27278
; SEQ ID NO 8448
;   LENGTH: 409
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: jC-gmle01810032d04a2
US-11-330-082-8448

  Query Match             3.5%;  Score 45.6;  DB 7;  Length 409;
  Best Local Similarity   51.5%;  Pred. No. 0.019;
  Matches  105;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy         66 GGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGG 125
              ||   | |||  || ||  | | |  | | | ||| |  |||||| | | | || || ||
Db        318 GGGGGGGGGGGTGGCAGGGGGGGGGGGGGGGGAGTGATGGGAGGGGGCGCGGGGGGGAGGG 259

Qy        126 CGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGG 185
              |   ||  |  |    |  | || || |   |  |||  ||  ||| ||| || |  ||
Db        258 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGAGGGGGGGGGGGGGGGGGGGG 199

Qy        186 CGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGG 245
              |   |  |  |    || || || ||   ||| |  | || || ||     | ||  ||
Db        198 GGGGGGGGGGGGGGGGGGGGGGGGGGGAGGGGGGGGGGGGGGGGGGGGGGGCGGGG 139

Qy        246 GCTTTGCGTTGGTGGGGGGATGCG 269
              |   | |  || ||||||  | |
Db        138 GAGGGGGGCGGGGGGGGGGGGGG 115


RESULT 72
US-11-330-364-11771/c
; Sequence 11771, Application US/11330364
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S.
;  APPLICANT:  Andersen, Scott E.
;  APPLICANT:  Dubois, Patrice
;  APPLICANT:  Mahadeo, Debbie A.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
```

```
;  FILE REFERENCE: 16517.352 - 38-21(51721)C/US
;  CURRENT APPLICATION NUMBER: US/11/330,364
;  CURRENT FILING DATE:  2006-01-12
;  PRIOR APPLICATION NUMBER: US 09/696,664
;  PRIOR FILING DATE: 2000-10-25
;  PRIOR APPLICATION NUMBER: US 60/161,619
;  PRIOR FILING DATE: 1999-10-26
;  NUMBER OF SEQ ID NOS: 17472
; SEQ ID NO 11771
;   LENGTH: 671
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3279-190-Q1-K6-C8
US-11-330-364-11771

  Query Match             3.5%;  Score 45.6;  DB 7;  Length 671;
  Best Local Similarity   63.9%;  Pred. No. 0.023;
  Matches   69;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy        118 GGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGC 177
              | ||  | || | |||    | |  || | || ||   | |||| | |||| | |  |||
Db        364 GCCGCGGACGCCGGGCTTGGGCTTGGAGGTGGTGGCCTCGGCAGCGGCGGCGGCGGCGGC 305

Qy        178 GGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAG 225
              |||| ||||  | | ||||||   |||||||||| | || ||| |||
Db        304 GGCGGCGGCAACAGCAGCAGCGGGGGCGGAGGAGTGCAGGTTGAGGAG 257


RESULT 73
US-60-762-056-22874/c
; Sequence 22874, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 22874
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-22874

  Query Match             3.5%;  Score 45.6;  DB 11;  Length 2300;
  Best Local Similarity   51.5%;  Pred. No. 0.037;
  Matches  105;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy         62 ATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCG 121
              ||  |||| |  ||   | |||          | || |    | | ||  |||| |||
Db       1922 ATCGGGCAGGGCGGGACGCGGCACGTGTTTTGGTCGGAGCGAGGGGCGGGGCCTGAGGCC 1863

Qy        122 ATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCG 181
              |||    |   | |   ||  ||||||      ||| |   ||| | | ||||||||
Db       1862 CTGGAAAGTAAGGGAAAGGTCGCAGCGGCGCCAGGGGCAAGCGGGGCGGAGGGGGCGGCG 1803

Qy        182 TCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGC 241
              ||||   |  || ||  |||||||| ||    || |  | || ||||||      |
Db       1802 GGGGCGGAGGGGGCGGCGGGGGCGGAGGGGGCGGCGGGGCGGCGGGGGCGGCGGGGCGGA 1743

Qy        242 CTGGGCTTTGCGTTGGTGGGGGGA 265
              ||||   |||  ||  |||| |
Db       1742 GGGGGCGGGGCGGCGGCTGGGGAA 1719


RESULT 74
US-11-266-748A-92915
; Sequence 92915, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 92915
;   LENGTH: 611
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-92915

  Query Match             3.5%;  Score 45.4;  DB 9;  Length 611;
  Best Local Similarity   97.9%;  Pred. No. 0.025;
  Matches   46;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         20 TTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCG 66
              |||||||||||||||||||||||||||||||||||||||||||||| |
Db          1 TTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTGG 47


RESULT 75
US-11-266-748A-145726/c
; Sequence 145726, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145726
```

```
;   LENGTH: 611
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-145726

  Query Match             3.5%;  Score 45.4;  DB 9;  Length 611;
  Best Local Similarity   97.9%;  Pred. No. 0.025;
  Matches   46;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         20 TTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCG 66
              |||||||||||||||||||||||||||||||||||||||||||||| |
Db        611 TTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTGG 565


Search completed: March 11, 2006, 16:04:39
Job time : 445 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:19:50 ; Search time 9330.13 Seconds
(without alignments)
1345.232 Million cell updates/sec

Title: US-10-757-745-2_COPY_121_347
Perfect score: 1192
Sequence: 1 IDGLDLNNLSERARGVCSYL.....HIIPRSLDLLGLEKLDCGRF 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170333_4165/app_query.fasta_1
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=rnpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss02h -USER=US10757745_@CGN_1_1_20983_@runat_09032006_170333_4165
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUSB_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUSC_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US075_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US076_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US077_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US078_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US079_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US101_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US104_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US105_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US112_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US117_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US600_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US601_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US603_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US605_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US606_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US607_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1192	100.0	1898	28	US-09-629-469A-13209	Sequence 13209, A
2	1192	100.0	1898	63	US-10-917-503-13209	Sequence 13209, A
3	1192	100.0	1919	42	US-10-170-235-27345	Sequence 27345, A
4	1192	100.0	1920	61	US-10-757-745-1	Sequence 1, Appli
5	1192	100.0	1930	31	US-09-705-256A-7708	Sequence 7708, Ap
6	1192	100.0	1930	75	US-60-164-285-7708	Sequence 7708, Ap
7	1192	100.0	1940	3	PCT-US04-24424-921	Sequence 921, App
8	1192	100.0	1940	61	US-10-783-271-28	Sequence 28, Appl

	9	1192	100.0	1940	80	US-60-490-890-921 Sequence 921, App
	10	1192	100.0	1940	81	US-60-568-073-582 Sequence 582, App
	11	1192	100.0	1940	82	US-60-671-501-156 Sequence 156, App
	12	1192	100.0	1940	82	US-60-685-372-523 Sequence 523, App
	13	1192	100.0	1940	83	US-60-704-740-28 Sequence 28, Appl
	14	1192	100.0	1940	83	US-60-710-726-741 Sequence 741, App
	15	1192	100.0	1948	24	US-09-488-725A-1200 Sequence 1200, Ap
	16	1192	100.0	1948	40	US-10-037-270-889 Sequence 889, App
	17	1192	100.0	1948	42	US-10-105-837-889 Sequence 889, App
	18	1192	100.0	1948	42	US-10-117-722-889 Sequence 889, App
	19	1192	100.0	1948	42	US-10-122-851-889 Sequence 889, App
	20	1192	100.0	1948	43	US-10-258-898A-1200 Sequence 1200, Ap
	21	1192	100.0	1948	43	US-10-286-897-1200 Sequence 1200, Ap
	22	1192	100.0	1958	24	US-09-471-275-1314 Sequence 1314, Ap
	23	1192	100.0	1958	24	US-09-488-725A-4772 Sequence 4772, Ap
	24	1192	100.0	1958	24	US-09-488-725B-8295 Sequence 8295, Ap
	25	1192	100.0	1958	27	US-09-552-317-8295 Sequence 8295, Ap
	26	1192	100.0	1958	43	US-10-258-898A-4772 Sequence 4772, Ap
	27	1192	100.0	1958	43	US-10-286-897-4772 Sequence 4772, Ap
	28	1192	100.0	1958	48	US-10-302-689A-59957 Sequence 59957, A
	29	1192	100.0	1958	48	US-10-302-689A-77389 Sequence 77389, A
	30	1192	100.0	2262	28	US-09-606-680-2868 Sequence 2868, Ap
	31	1192	100.0	2262	28	US-09-606-776-4085 Sequence 4085, Ap
	32	1192	100.0	2262	29	US-09-644-869-9498 Sequence 9498, Ap
	33	1192	100.0	2262	29	US-09-652-121-6794 Sequence 6794, Ap
	34	1192	100.0	2262	29	US-09-652-122-3756 Sequence 3756, Ap
	35	1192	100.0	2262	29	US-09-652-128-9329 Sequence 9329, Ap
	36	1192	100.0	2262	29	US-09-652-816-9572 Sequence 9572, Ap
	37	1192	100.0	2262	31	US-09-700-000-6419 Sequence 6419, Ap
	38	1192	100.0	2262	31	US-09-710-280-2045 Sequence 2045, Ap
	39	1192	100.0	2262	31	US-09-710-282-3926 Sequence 3926, Ap
	40	1192	100.0	2262	31	US-09-726-802-2141 Sequence 2141, Ap
	41	1192	100.0	3080	27	US-09-585-944-2184 Sequence 2184, Ap
c	42	1192	100.0	3082	75	US-60-172-360-956 Sequence 956, App
	43	1192	100.0	3152	38	US-09-981-353-88 Sequence 88, Appl
	44	1192	100.0	3152	77	US-60-239-841-88 Sequence 88, Appl
	45	1184	99.3	1296	1	PCT-US00-05883-170 Sequence 170, App
	46	1184	99.3	1296	36	US-09-925-299-170 Sequence 170, App
	47	1167	97.9	1919	79	US-60-443-566-1795 Sequence 1795, Ap
	48	1167	97.9	1919	79	US-60-452-680-10426 Sequence 10426, A
	49	1167	97.9	1919	79	US-60-455-444-3324 Sequence 3324, Ap
	50	1167	97.9	1919	79	US-60-465-241-3324 Sequence 3324, Ap
	51	1116	93.6	1088	1	PCT-US01-08631-22058 Sequence 22058, A
	52	1116	93.6	1088	48	US-10-302-689A-136240 Sequence 136240,
	53	1116	93.6	1088	51	US-10-450-763-22058 Sequence 22058, A
	54	1102	92.4	674	1	PCT-US02-02870-1067 Sequence 1067, Ap
	55	1102	92.4	674	40	US-10-066-543-1067 Sequence 1067, Ap
	56	1102	92.4	674	43	US-10-214-403-1067 Sequence 1067, Ap
c	57	1090	91.4	1079	22	US-09-248-178-63 Sequence 63, Appl
c	58	1090	91.4	1079	22	US-09-288-950-63 Sequence 63, Appl
c	59	1090	91.4	1079	23	US-09-346-327-63 Sequence 63, Appl
c	60	1090	91.4	1079	29	US-09-687-507-63 Sequence 63, Appl
c	61	1090	91.4	1079	29	US-09-687-507A-63 Sequence 63, Appl
c	62	1090	91.4	1079	31	US-09-745-288-63 Sequence 63, Appl
c	63	1090	91.4	1079	32	US-09-778-381-63 Sequence 63, Appl
c	64	1090	91.4	1079	51	US-10-453-919-63 Sequence 63, Appl
	65	1069	89.7	644	1	PCT-US02-02870-1070 Sequence 1070, Ap
	66	1069	89.7	644	40	US-10-066-543-1070 Sequence 1070, Ap
	67	1069	89.7	644	43	US-10-214-403-1070 Sequence 1070, Ap
	68	1054	88.4	633	1	PCT-US02-02870-1084 Sequence 1084, Ap
	69	1054	88.4	633	40	US-10-066-543-1084 Sequence 1084, Ap
	70	1054	88.4	633	43	US-10-214-403-1084 Sequence 1084, Ap
c	71	1041	87.3	1849	42	US-10-170-235-26834 Sequence 26834, A
	72	1031	86.5	625	1	PCT-US02-02870-548 Sequence 548, App
	73	1031	86.5	625	40	US-10-066-543-548 Sequence 548, App
	74	1031	86.5	625	43	US-10-214-403-548 Sequence 548, App
	75	1020	85.6	625	1	PCT-US02-02870-1031 Sequence 1031, Ap

ALIGNMENTS

RESULT 1
US-09-629-469A-13209

; Sequence 13209, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1086)
US-09-629-469A-13209

Alignment Scores:

Pred. No.:	7.05e-180	Length:	1898
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	28	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-629-469A-13209 (1-1898)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        221 LysLeuAspCysGlyArgPhe 227
                |||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTT 1041
```

RESULT 2
US-10-917-503-13209

```
; Sequence 13209, Application US/10917503
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
;   LENGTH: 1898
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1086)
US-10-917-503-13209
```

```
Alignment Scores:
Pred. No.:                7.05e-180       Length:          1898
Score:                    1192.00         Matches:         227
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       63              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-917-503-13209 (1-1898)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTT 1041
```

RESULT 3
US-10-170-235-27345

```
; Sequence 27345, Application US/10170235
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig
;  TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAI
;  TITLE OF INVENTION:  TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
;  FILE REFERENCE: CL001380
;  CURRENT APPLICATION NUMBER: US/10/170,235
;  CURRENT FILING DATE:  2003-03-17
;  NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 27345
;   LENGTH: 1919
;   TYPE: DNA
;   ORGANISM: HUMAN
```

US-10-170-235-27345

Alignment Scores:

Pred. No.:	7.2e-180	Length:	1919
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	42	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-170-235-27345 (1-1919)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 617

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 677

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 737

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 797

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        798 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 857

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 917

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 977

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1037

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1038 AAACTGGACTGTGGTAGATTT 1058
```

RESULT 4
US-10-757-745-1
; Sequence 1, Application US/10757745
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/10/757,745
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/697,863A
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1627)..(1627)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1108)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

Alignment Scores:

Pred. No.:	7.21e-180	Length:	1920
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	61	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-757-745-1 (1-1920)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
```

```
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTT 1060

RESULT 5
US-09-705-256A-7708
; Sequence 7708, Application US/09705256A
; GENERAL INFORMATION:
;  APPLICANT: Ma, Xiao-Jun; Dotson, Stanton B.:  Monsanto Company
;  TITLE OF INVENTION: Tumor Associated Molecules (TAMs): Targets for diagnosis, treatme
;  TITLE OF INVENTION:  and prevention of cancer
;  FILE REFERENCE: 3214
;  CURRENT APPLICATION NUMBER: US/09/705,256A
;  CURRENT FILING DATE:  1999-05-11
;  PRIOR APPLICATION NUMBER: US 60/164,285
;  PRIOR FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 8259
; SEQ ID NO 7708
;   LENGTH: 1930
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-705-256A-7708

Alignment Scores:
Pred. No.:              7.28e-180      Length:          1930
Score:                  1192.00        Matches:         227
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            100.0%         Indels:          0
DB:                     31             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-705-256A-7708 (1-1930)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 446

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 506

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 566

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 626

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 686

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 746

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 806

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 866

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 926

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 986

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1046

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||
Db       1047 AAACTGGACTGTGGTAGATTT 1067

RESULT 6
US-60-164-285-7708
; Sequence 7708, Application US/60164285
; GENERAL INFORMATION:
;  APPLICANT: Ma, Xiao-Jun
;  TITLE OF INVENTION: Tumor Associated Molecules (TAMs): Targets for diagnosis, treatm
;  FILE REFERENCE: 3214
;  CURRENT APPLICATION NUMBER: US/60/164,285
;  CURRENT FILING DATE: 1999-11-05
;  NUMBER OF SEQ ID NOS: 8259
; SEQ ID NO 7708
;   LENGTH: 1930
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-164-285-7708

Alignment Scores:
Pred. No.:              7.28e-180      Length:          1930
Score:                  1192.00        Matches:         227
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            100.0%         Indels:          0
DB:                     75             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-164-285-7708 (1-1930)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 446

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 506

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 566

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 626

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 686

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 746

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 806

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 866

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 926
```

```
Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 986

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1046

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1047 AAACTGGACTGTGGTAGATTT 1067
```

RESULT 7
PCT-US04-24424-921
; Sequence 921, Application PC/TUS0424424
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PCT
; CURRENT APPLICATION NUMBER: PCT/US04/24424
; CURRENT FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US 60/490,890
; PRIOR FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2786
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-24424-921

Alignment Scores:

Pred. No.:	7.35e-180	Length:	1940
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x PCT-US04-24424-921 (1-1940)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066
```

RESULT 8
US-10-783-271-28
; Sequence 28, Application US/10783271
; GENERAL INFORMATION:
; APPLICANT: Veridex, LLC
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
; FILE REFERENCE: VDX-5003 USNP
; CURRENT APPLICATION NUMBER: US/10/783,271
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: human
US-10-783-271-28

Alignment Scores:

Pred. No.:	7.35e-180	Length:	1940
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	61	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-783-271-28 (1-1940)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
```

```
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180

Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200

Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220

Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227

Db       1046 AAACTGGACTGTGGTAGATTT 1066
```

RESULT 9
US-60-490-890-921
; Sequence 921, Application US/60490890
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PSP
; CURRENT APPLICATION NUMBER: US/60/490,890
; CURRENT FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2779
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-490-890-921

Alignment Scores:

Pred. No.:	7.35e-180	Length:	1940
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	80	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-490-890-921 (1-1940)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20

Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40

Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60

Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80

Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100

Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120

Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140

Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180

Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200

Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220

Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227

Db       1046 AAACTGGACTGTGGTAGATTT 1066
```

RESULT 10
US-60-568-073-582
; Sequence 582, Application US/60568073
; GENERAL INFORMATION:
; APPLICANT: Huang, Fei
; APPLICANT: Han, Xia
; APPLICANT: Shaw, Peter
; APPLICANT: Clark, Edwin
; TITLE OF INVENTION: COLON CANCER BIOMARKERS
; FILE REFERENCE: 10205 PSP
; CURRENT APPLICATION NUMBER: US/60/568,073
; CURRENT FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 1305
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 582
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-568-073-582

Alignment Scores:

Pred. No.:	7.35e-180	Length:	1940
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	81	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-568-073-582 (1-1940)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20

Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40

Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60

Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
```

```
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066

RESULT 11
US-60-671-501-156
; Sequence 156, Application US/60671501
; GENERAL INFORMATION:
;  APPLICANT: Wim Van Criekinge
;  APPLICANT:  Josef Straub
;  TITLE OF INVENTION: METHYLATION MARKERS FOR DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF CANCERS
;  FILE REFERENCE: 006791.00001
;  CURRENT APPLICATION NUMBER: US/60/671,501
;  CURRENT FILING DATE:  2005-04-15
;  NUMBER OF SEQ ID NOS: 420
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 156
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-671-501-156

Alignment Scores:
Pred. No.:              7.35e-180       Length:         1940
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     82              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-671-501-156 (1-1940)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066

RESULT 12
US-60-685-372-523
; Sequence 523, Application US/60685372
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John V
;  APPLICANT:  Bradley, Walter E
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall D
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul V
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: GeneMap of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 059908-5005-PR
;  CURRENT APPLICATION NUMBER: US/60/685,372
;  CURRENT FILING DATE:  2005-05-31
;  NUMBER OF SEQ ID NOS: 2738
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 523
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-685-372-523

Alignment Scores:
Pred. No.:              7.35e-180       Length:         1940
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     82              Gaps:           0
```

US-10-757-745-2_COPY_121_347 (1-227) x US-60-685-372-523 (1-1940)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066
```

RESULT 13
US-60-704-740-28
; Sequence 28, Application US/60704740
; GENERAL INFORMATION:
; APPLICANT: Wang, Yixin
; APPLICANT: Zhang, Yi
; APPLICANT: Atkins, David
; APPLICANT: Sieuwerts, Anieta M.
; APPLICANT: Smid, Marcel
; APPLICANT: Klijn, Jan G.M.
; APPLICANT: Martens, John W.M.
; APPLICANT: Foekens, John A.
; TITLE OF INVENTION: Predicting Bone Relapse of Breast Cancer
; FILE REFERENCE: VDX 5026 USPSP
; CURRENT APPLICATION NUMBER: US/60/704,740
; CURRENT FILING DATE: 2005-08-02
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: human
US-60-704-740-28

Alignment Scores:

Pred. No.:	7.35e-180	Length:	1940
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	83	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-704-740-28 (1-1940)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066
```

RESULT 14
US-60-710-726-741
; Sequence 741, Application US/60710726
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John V
; APPLICANT: Bradley, Walter E
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Serre, Valerie
; APPLICANT: Eerdewegh, Paul V

```
;  APPLICANT:  Little, Randall D
;  APPLICANT:  Keith , Tim
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: GeneMap of the human genes associated with Crohn's disease
;  FILE REFERENCE: 059908-5009-PR
;  CURRENT APPLICATION NUMBER: US/60/710,726
;  CURRENT FILING DATE:  2005-08-24
;  NUMBER OF SEQ ID NOS: 20166
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 741
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homosapiens
US-60-710-726-741

Alignment Scores:
Pred. No.:              7.35e-180       Length:         1940
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     83              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-710-726-741 (1-1940)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066
```

```
RESULT 15
US-09-488-725A-1200
; Sequence 1200, Application US/09488725A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784PLPCT
;  CURRENT APPLICATION NUMBER: US/09/488,725A
;  CURRENT FILING DATE:  2000-12-22
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7144
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-09-488-725A-1200

Alignment Scores:
Pred. No.:              7.41e-180       Length:         1948
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     24              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-488-725A-1200 (1-1948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818
```

```
Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079

RESULT 16
US-10-037-270-889
; Sequence 889, Application US/10037270
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/10/037,270
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-037-270-889

Alignment Scores:
Pred. No.:              7.41e-180      Length:         1948
Score:                  1192.00        Matches:        227
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     40             Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-037-270-889 (1-1948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079

RESULT 17
US-10-105-837-889
; Sequence 889, Application US/10105837
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BDIVA
;  CURRENT APPLICATION NUMBER: US/10/105,837
;  CURRENT FILING DATE:  2002-03-21
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-105-837-889

Alignment Scores:
Pred. No.:              7.41e-180       Length:         1948
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     42              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-105-837-889 (1-1948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079

RESULT 18
US-10-117-722-889
; Sequence 889, Application US/10117722
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BCIP
;  CURRENT APPLICATION NUMBER: US/10/117,722
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-117-722-889

Alignment Scores:
Pred. No.:              7.41e-180       Length:         1948
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     42              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-117-722-889 (1-1948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079
```

RESULT 19
US-10-122-851-889
; Sequence 889, Application US/10122851
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BDV3
; CURRENT APPLICATION NUMBER: US/10/122,851
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-122-851-889

Alignment Scores:

Pred. No.:	7.41e-180	Length:	1948
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	42	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-122-851-889 (1-1948)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079
```

RESULT 20
US-10-258-898A-1200
; Sequence 1200, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/258,898A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-258-898A-1200

Alignment Scores:

Pred. No.:	7.41e-180	Length:	1948
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	43	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-258-898A-1200 (1-1948)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy       101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy       121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy       141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy       161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy       181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy       201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy       221 LysLeuAspCysGlyArgPhe 227
             |||||||||||||||||||||
Db      1059 AAACTGGACTGTGGTAGATTT 1079
```

RESULT 21
US-10-286-897-1200
; Sequence 1200, Application US/10286897
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-286-897-1200

Alignment Scores:			
Pred. No.:	7.41e-180	Length:	1948
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	43	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-286-897-1200 (1-1948)

```
Qy         1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy        21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy        41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy        61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy        81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy       101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy       121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy       141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy       161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy       181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy       201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy       221 LysLeuAspCysGlyArgPhe 227
             |||||||||||||||||||||
Db      1059 AAACTGGACTGTGGTAGATTT 1079
```

RESULT 22
US-09-471-275-1314
; Sequence 1314, Application US/09471275
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Contigs Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 782
; CURRENT APPLICATION NUMBER: US/09/471,275

```
;  CURRENT FILING DATE: 1999-12-23
;  EARLIER APPLICATION NUMBER: US 09/235,076
;  EARLIER FILING DATE: 1999-01-20
;  EARLIER APPLICATION NUMBER: US 09/234,611
;  EARLIER FILING DATE: 1999-01-22
;  EARLIER APPLICATION NUMBER: US 09/240,371
;  EARLIER FILING DATE: 1999-01-29
;  EARLIER APPLICATION NUMBER: US 09/277,227
;  EARLIER FILING DATE: 1999-03-25
;  EARLIER APPLICATION NUMBER: US 09/271,490
;  EARLIER FILING DATE: 1999-03-18
;  EARLIER APPLICATION NUMBER: US 09/293,972
;  EARLIER FILING DATE: 1999-04-15
;  EARLIER APPLICATION NUMBER: US 09/274,861
;  EARLIER FILING DATE: 1999-03-23
;  EARLIER APPLICATION NUMBER: US 60/125,453
;  EARLIER FILING DATE: 1999-03-19
;  EARLIER APPLICATION NUMBER: US 60/126,605
;  EARLIER FILING DATE: 1999-03-26
;  EARLIER APPLICATION NUMBER: US 09/306,350
;  EARLIER FILING DATE: 1999-05-07
;  EARLIER APPLICATION NUMBER: US 09/399,720
;  EARLIER FILING DATE: 1999-09-21
;  EARLIER APPLICATION NUMBER: US 09/404,284
;  EARLIER FILING DATE: 1999-09-21
;  EARLIER APPLICATION NUMBER:
;  EARLIER FILING DATE: 1999-12-16
;  NUMBER OF SEQ ID NOS: 10451
;  SOFTWARE: pt_CT_genes Version 1.0
; SEQ ID NO 1314
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-471-275-1314

Alignment Scores:
Pred. No.:              7.48e-180       Length:         1958
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     24              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-471-275-1314 (1-1958)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTT 1091

RESULT 23
US-09-488-725A-4772
; Sequence 4772, Application US/09488725A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/09/488,725A
;  CURRENT FILING DATE:  2000-12-22
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7144
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-488-725A-4772

Alignment Scores:
Pred. No.:              7.48e-180       Length:         1958
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     24              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-488-725A-4772 (1-1958)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy          41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy          81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy         101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy         121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy         141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy         161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy         181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy         201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy         221 LysLeuAspCysGlyArgPhe 227
               |||||||||||||||||||||
Db        1071 AAACTGGACTGTGGTAGATTT 1091

RESULT 24
US-09-488-725B-8295
; Sequence 8295, Application US/09488725B
; GENERAL INFORMATION:
;  APPLICANT: Yuanhua T. Tang
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Ankura Sinku
;  APPLICANT:  Chenghua Liu
;  APPLICANT:  Radoje T. Drmanac
;  TITLE OF INVENTION: Novel Contigs Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 784
;  CURRENT APPLICATION NUMBER: US/09/488,725B
;  CURRENT FILING DATE:  2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/004,182
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: US 09/034,341
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 09/045,400
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: US 09/321,214
;  PRIOR FILING DATE: 1999-05-26
;  PRIOR APPLICATION NUMBER: US 09/131,598
;  PRIOR FILING DATE: 1998-08-10
;  PRIOR APPLICATION NUMBER: US 09/170,294
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: US 09/179,473
;  PRIOR FILING DATE: 1998-10-27
;  PRIOR APPLICATION NUMBER: US 09/181,430
;  PRIOR FILING DATE: 1998-10-28
;  PRIOR APPLICATION NUMBER: US 09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  PRIOR APPLICATION NUMBER: US 09/234,611
;  PRIOR FILING DATE: 1999-01-22
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10289
;  SOFTWARE: pt_CT_genes Version 1.01
; SEQ ID NO 8295
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-488-725B-8295

Alignment Scores:
Pred. No.:              7.48e-180       Length:         1958
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     24              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-488-725B-8295 (1-1958)

Qy           1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy          21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy          41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy          81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy         101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy         121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy         141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy         161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy         181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy         201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy         221 LysLeuAspCysGlyArgPhe 227
```

```
                    ||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTT 1091

RESULT 25
US-09-552-317-8295
; Sequence 8295, Application US/09552317
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Sinku, Ankura
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids
;  TITLE OF INVENTION:  and Polypeptides
;  FILE REFERENCE: 784CIP
;  CURRENT APPLICATION NUMBER: US/09/552,317
;  CURRENT FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 10289
;  SOFTWARE: pt_CT_genes Version 1.01
; SEQ ID NO 8295
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-552-317-8295

Alignment Scores:
Pred. No.:                7.48e-180       Length:         1958
Score:                    1192.00         Matches:        227
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       27              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-552-317-8295 (1-1958)

Qy           1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy          21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy          41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy          81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy         101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy         121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy         141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy         161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy         181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy         201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuGluLeuGluGlu 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy         221 LysLeuAspCysGlyArgPhe 227
                ||||||||||||||||||||
Db        1071 AAACTGGACTGTGGTAGATTT 1091

RESULT 26
US-10-258-898A-4772
; Sequence 4772, Application US/10258898A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/258,898A
;  CURRENT FILING DATE:  2002-10-29
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-258-898A-4772

Alignment Scores:
Pred. No.:                7.48e-180       Length:         1958
Score:                    1192.00         Matches:        227
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       43              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-258-898A-4772 (1-1958)

Qy           1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy          21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy          41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        221 LysLeuAspCysGlyArgPhe 227
                |||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTT 1091

RESULT 27
US-10-286-897-4772
; Sequence 4772, Application US/10286897
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/286,897
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US/09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US/09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US/09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US/09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US/09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US/09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US/09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-286-897-4772

Alignment Scores:
Pred. No.:              7.48e-180       Length:         1958
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0

Query Match:            100.0%          Indels:         0
DB:                     43              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-286-897-4772 (1-1958)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        221 LysLeuAspCysGlyArgPhe 227
                |||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTT 1091

RESULT 28
US-10-302-689A-59957
; Sequence 59957, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ballinger, Dennis
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Liu, Jin
;  APPLICANT:  Loeb, Deborah
;  APPLICANT:  Montgomery, Julia, R.
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
```

```
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 59957
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-302-689A-59957

Alignment Scores:
Pred. No.:              7.48e-180      Length:          1958
Score:                  1192.00        Matches:         227
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            100.0%         Indels:          0
DB:                     48             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-302-689A-59957 (1-1958)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTT 1091

RESULT 29
US-10-302-689A-77389
; Sequence 77389, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ballinger, Dennis
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Liu, Jin
;  APPLICANT:  Loeb, Deborah
;  APPLICANT:  Montgomery, Julia, R.
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 77389
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-302-689A-77389

Alignment Scores:
Pred. No.:              7.48e-180      Length:          1958
Score:                  1192.00        Matches:         227
Percent Similarity:     100.0%         Conservative:    0
Best Local Similarity:  100.0%         Mismatches:      0
Query Match:            100.0%         Indels:          0
DB:                     48             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-302-689A-77389 (1-1958)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTT 1091
```

RESULT 30

```
US-09-606-680-2868
; Sequence 2868, Application US/09606680
; GENERAL INFORMATION:
;  APPLICANT: Lloyd, Clare M.
;  APPLICANT:  Williamson, Mark
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 1600.1131-001
;  CURRENT APPLICATION NUMBER: US/09/606,680
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER: 60/141,227
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 60/141,226
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 4394
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2868
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-606-680-2868
```

```
Alignment Scores:
Pred. No.:              9.86e-180       Length:         2262
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     28              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-606-680-2868 (1-2262)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProProPro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTT 1355
```

RESULT 31

```
US-09-606-776-4085
; Sequence 4085, Application US/09606776
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  APPLICANT:  Myers, Paul
;  APPLICANT:  Gearing, David P.
;  APPLICANT:  Pan, Yang
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 1600.1129-001
;  CURRENT APPLICATION NUMBER: US/09/606,776
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER: 60/141,578
```

```
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 60/141,379
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: 60/141,138
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: 60/141,581
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 5415
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4085
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-606-776-4085

Alignment Scores:
Pred. No.:              9.86e-180       Length:         2262
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     28              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-606-776-4085 (1-2262)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTT 1355

RESULT 32
US-09-644-869-9498
; Sequence 9498, Application US/09644869
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  APPLICANT:  McCarthy, Sean A.
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Monahan, John
;  APPLICANT:  Richardson, Jennifer
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1182-001
;  CURRENT APPLICATION NUMBER: US/09/644,869
;  CURRENT FILING DATE:  2000-08-28
;  PRIOR APPLICATION NUMBER:  60/151,062
;  PRIOR FILING DATE:  1999-08-27
;  NUMBER OF SEQ ID NOS:  9708
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9498
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-644-869-9498

Alignment Scores:
Pred. No.:              9.86e-180       Length:         2262
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     29              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-644-869-9498 (1-2262)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154
```

```
Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTT 1355

RESULT 33
US-09-652-121-6794
; Sequence 6794, Application US/09652121
; GENERAL INFORMATION:
;  APPLICANT:  DiStefano, Peter
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1188-001
;  CURRENT APPLICATION NUMBER: US/09/652,121
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,129
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  7615
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6794
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-121-6794

Alignment Scores:
Pred. No.:                9.86e-180       Length:          2262
Score:                    1192.00         Matches:         227
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       29              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-652-121-6794 (1-2262)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTT 1355

RESULT 34
US-09-652-122-3756
; Sequence 3756, Application US/09652122
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1194-001
;  CURRENT APPLICATION NUMBER: US/09/652,122
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,421
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  5020
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3756
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-122-3756

Alignment Scores:
Pred. No.:                9.86e-180       Length:          2262
Score:                    1192.00         Matches:         227
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       29              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-652-122-3756 (1-2262)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
```

```
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100

Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120

Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140

Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180

Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200

Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220

Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227

Db       1335 AAACTGGACTGTGGTAGATTT 1355
```

RESULT 35
US-09-652-128-9329
; Sequence 9329, Application US/09652128
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1171-001
; CURRENT APPLICATION NUMBER: US/09/652,128
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,133
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10265
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9329
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2262)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-128-9329

Alignment Scores:

Pred. No.:	9.86e-180	Length:	2262
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	29	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-652-128-9329 (1-2262)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20

Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40

Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60

Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80

Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100

Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120

Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140

Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180

Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200

Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220

Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227

Db       1335 AAACTGGACTGTGGTAGATTT 1355
```

RESULT 36
US-09-652-816-9572
; Sequence 9572, Application US/09652816
; GENERAL INFORMATION:
; APPLICANT: Gutierrez-Ramos, Jose-Carlos
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1177-001
; CURRENT APPLICATION NUMBER: US/09/652,816
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,111
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 9647
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9572
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2262)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-816-9572

Alignment Scores:

Pred. No.:	9.86e-180	Length:	2262
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0

```
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     29              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-652-816-9572 (1-2262)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTT 1355

RESULT 37
US-09-700-000-6419
; Sequence 6419, Application US/09700000
; GENERAL INFORMATION:
;  APPLICANT:  Richardson, Jennifer
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2022-001
;  CURRENT APPLICATION NUMBER: US/09/700,000
;  CURRENT FILING DATE:  2000-10-30
;  PRIOR APPLICATION NUMBER:  60/162,619
;  PRIOR FILING DATE:  1999-10-29
;  NUMBER OF SEQ ID NOS:  7171
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6419
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-700-000-6419

Alignment Scores:
Pred. No.:              9.86e-180       Length:         2262
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-700-000-6419 (1-2262)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTT 1355

RESULT 38
US-09-710-280-2045
; Sequence 2045, Application US/09710280
; GENERAL INFORMATION:
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Gearing, David P.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2007-001
```

```
;  CURRENT APPLICATION NUMBER: US/09/710,280
;  CURRENT FILING DATE:  2000-11-10
;  PRIOR APPLICATION NUMBER:  60/164,537
;  PRIOR FILING DATE:  1999-11-10
;  NUMBER OF SEQ ID NOS:  2763
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2045
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-710-280-2045

Alignment Scores:
Pred. No.:              9.86e-180       Length:         2262
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-710-280-2045 (1-2262)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTT 1355
```

```
RESULT 39
US-09-710-282-3926
; Sequence 3926, Application US/09710282
; GENERAL INFORMATION:
;  APPLICANT:  Gearing, David P.
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Kingsbury, Gillian A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2004-001
;  CURRENT APPLICATION NUMBER: US/09/710,282
;  CURRENT FILING DATE:  2000-11-10
;  PRIOR APPLICATION NUMBER:  60/164,256
;  PRIOR FILING DATE:  1999-11-09
;  NUMBER OF SEQ ID NOS:  4662
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3926
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-710-282-3926

Alignment Scores:
Pred. No.:              9.86e-180       Length:         2262
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-710-282-3926 (1-2262)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214
```

```
Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTT 1355

RESULT 40
US-09-726-802-2141
; Sequence 2141, Application US/09726802
; GENERAL INFORMATION:
;  APPLICANT:  Acton, Susan L.
;  APPLICANT:  Falb, Dean A.
;  APPLICANT:
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2025-001
;  CURRENT APPLICATION NUMBER: US/09/726,802
;  CURRENT FILING DATE:  2000-11-30
;  PRIOR APPLICATION NUMBER:  60/168,012
;  PRIOR FILING DATE:  1999-11-30
;  NUMBER OF SEQ ID NOS:  2872
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2141
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-726-802-2141

Alignment Scores:
Pred. No.:              9.86e-180       Length:         2262
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-726-802-2141 (1-2262)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 734

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 794

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 854

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 914

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 974

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1034

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1094

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1154

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1214

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1274

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1275 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1334

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1335 AAACTGGACTGTGGTAGATTT 1355

RESULT 41
US-09-585-944-2184
; Sequence 2184, Application US/09585944
; GENERAL INFORMATION:
;  APPLICANT: Hodgson, David M.
;  APPLICANT: Lincoln, Stephen E.
;  APPLICANT: Russo, Frank D.
;  APPLICANT: Spiro, Peter A.
;  APPLICANT: Banville, Steve C.
;  APPLICANT: Bratcher, Shawn R.
;  APPLICANT: Dufour, Gerard E.
;  APPLICANT: Cohen, Howard J.
;  APPLICANT: Rosen, Bruce
;  APPLICANT: Chalup, Michael S.
;  APPLICANT: Hillman, Jennifer L.
;  APPLICANT: Jones, Anissa L.
;  APPLICANT: Yu, Jimmy Y.
;  APPLICANT: Greenawalt, Lila B.
;  APPLICANT: Panzer, Scott R.
;  APPLICANT: Roseberry, Ann M.
;  APPLICANT: Wright, Rachel J.
;  APPLICANT: Daniels, Susan E.
;  TITLE OF INVENTION: SECRETORY MOLECULES
;  FILE REFERENCE: PT-0064 US
;  CURRENT APPLICATION NUMBER: US/09/585,944
;  CURRENT FILING DATE: 2000-06-02
;  EARLIER APPLICATION NUMBER: US 60/137,412
;  EARLIER FILING DATE: 1999-06-03
;  EARLIER APPLICATION NUMBER: US 60/147,542
;  EARLIER FILING DATE: 1999-08-05
;  EARLIER APPLICATION NUMBER: US 60/147,501
;  EARLIER FILING DATE: 1999-08-05
;  EARLIER APPLICATION NUMBER: US 60/147,500
;  EARLIER FILING DATE: 1999-08-05
;  NUMBER OF SEQ ID NOS: 2515
;  SOFTWARE: PERL Program
; SEQ ID NO 2184
;   LENGTH: 3080
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 232992.7
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 125, 1098
;   OTHER INFORMATION: a, t, c, g, or other
US-09-585-944-2184

Alignment Scores:
```

```
Pred. No.:              1.78e-179       Length:         3080
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     27              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-585-944-2184 (1-3080)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1540 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 1599

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1600 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 1659

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1660 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1719

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1720 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1779

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1780 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1839

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1840 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1899

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1900 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1959

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1960 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 2019

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2020 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 2079

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2080 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 2139

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2140 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 2199

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       2200 AAACTGGACTGTGGTAGATTT 2220

RESULT 42
US-60-172-360-956/c
; Sequence 956, Application US/60172360
; GENERAL INFORMATION:
;  APPLICANT: Morris, MacDonald
;  APPLICANT: Lal, Preeti
;  APPLICANT: Diep, Dinh
;  TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
;  TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymorp
;  FILE REFERENCE: GX-0007 P
;  CURRENT APPLICATION NUMBER: US/60/172,360
;  CURRENT FILING DATE: 1999-12-16
;  NUMBER OF SEQ ID NOS: 29838
;  SOFTWARE: PERL Program
; SEQ ID NO 956
;   LENGTH: 3082
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 232992.7c
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 1983, 2958
;   OTHER INFORMATION: a, t, c, g, or other
US-60-172-360-956

Alignment Scores:
Pred. No.:              1.78e-179       Length:         3082
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     75              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-172-360-956 (1-3082)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1541 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 1482

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1481 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 1422

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1421 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1362

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1361 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1302

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1301 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1242

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1241 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1182

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1181 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1122

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1121 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1062

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1061 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1002

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1001 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 942

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 882

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db        881 AAACTGGACTGTGGTAGATTT 861

RESULT 43
US-09-981-353-88
; Sequence 88, Application US/09981353
; GENERAL INFORMATION:
```

```
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE: PA-0038 US
;  CURRENT APPLICATION NUMBER: US/09/981,353
;  CURRENT FILING DATE:  2001-10-11
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 232992.1
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

Alignment Scores:
Pred. No.:                1.86e-179       Length:         3152
Score:                    1192.00         Matches:        227
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       38              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-981-353-88 (1-3152)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 1671

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1672 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 1731

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1732 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1791

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1792 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1851

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1852 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1911

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1912 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1971

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1972 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 2031

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2032 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 2091

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2092 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 2151

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2152 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 2211

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2212 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 2271

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       2272 AAACTGGACTGTGGTAGATTT 2292

RESULT 44
US-60-239-841-88
; Sequence 88, Application US/60239841
; GENERAL INFORMATION:
;  APPLICANT:  Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION:  GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE:  PA-0038 P
;  CURRENT APPLICATION NUMBER: US/60/239,841
;  CURRENT FILING DATE:  2000-10-11
;  NUMBER OF SEQ ID NOS:  194
;  SOFTWARE:  PERL Program
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 232992.1
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-60-239-841-88

Alignment Scores:
Pred. No.:                1.86e-179       Length:         3152
Score:                    1192.00         Matches:        227
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       77              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-239-841-88 (1-3152)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 1671

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1672 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 1731

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1732 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1791

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1792 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1851

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1852 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1911

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1912 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1971

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1972 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 2031

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2032 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 2091

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2092 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 2151

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2152 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 2211

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2212 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 2271

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       2272 AAACTGGACTGTGGTAGATTT 2292

RESULT 45
PCT-US00-05883-170
; Sequence 170, Application PC/TUS0005883
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT: Steve Ruben
;  TITLE OF INVENTION: Human Colon Cancer Associated Gene Sequences and Polypeptides
;  FILE REFERENCE: PA102PCT
;  CURRENT APPLICATION NUMBER: PCT/US00/05883
;  CURRENT FILING DATE: 2000-03-08
;  EARLIER APPLICATION NUMBER: 60/124,270
;  EARLIER FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
PCT-US00-05883-170

Alignment Scores:
Pred. No.:              6.75e-179      Length:         1296
Score:                  1184.00        Matches:        226
Percent Similarity:     99.6%          Conservative:   0
Best Local Similarity:  99.6%          Mismatches:     1
Query Match:            99.3%          Indels:         0
DB:                     1              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x PCT-US00-05883-170 (1-1296)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 443

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 503

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACA 563

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 623

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTT 683

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 743

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 803

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 863

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 GTGGATGTCTGGGAGTTTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAA 923

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 983

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        984 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1043

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1044 AAACTGGACTGTGGTAGATTT 1064

RESULT 46
US-09-925-299-170
; Sequence 170, Application US/09925299
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:              6.75e-179      Length:         1296
Score:                  1184.00        Matches:        226
Percent Similarity:     99.6%          Conservative:   0
Best Local Similarity:  99.6%          Mismatches:     1
Query Match:            99.3%          Indels:         0
DB:                     36             Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-299-170 (1-1296)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 443

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 503
```

```
Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACA 563

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 623

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTT 683

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 743

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 803

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 863

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 GTGGATGTCTGGGAGTTTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAA 923

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 983

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        984 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1043

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1044 AAACTGGACTGTGGTAGATTT 1064
```

RESULT 47
US-60-443-566-1795
; Sequence 1795, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1795
; LENGTH: 1919
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-443-566-1795

Alignment Scores:

Pred. No.:	8.19e-176	Length:	1919
Score:	1167.00	Matches:	223
Percent Similarity:	98.2%	Conservative:	0
Best Local Similarity:	98.2%	Mismatches:	4
Query Match:	97.9%	Indels:	0
DB:	79	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-443-566-1795 (1-1919)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 617

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        618 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 677

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 737

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        738 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 797

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        798 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 857

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 917

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 977

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1037

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1038 AAACTGGACTGTGGTAGATTT 1058
```

RESULT 48
US-60-452-680-10426
; Sequence 10426, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10426
; LENGTH: 1919
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-452-680-10426

Alignment Scores:

Pred. No.:	8.19e-176	Length:	1919
Score:	1167.00	Matches:	223
Percent Similarity:	98.2%	Conservative:	0
Best Local Similarity:	98.2%	Mismatches:	4
Query Match:	97.9%	Indels:	0
DB:	79	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-452-680-10426 (1-1919)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 617

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        618 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 677

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 737

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        738 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 797

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        798 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 857

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 917

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 977

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1037

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1038 AAACTGGACTGTGGTAGATTT 1058
```

RESULT 49
US-60-455-444-3324
; Sequence 3324, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3324
; LENGTH: 1919
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-455-444-3324

Alignment Scores:

Pred. No.:	8.19e-176	Length:	1919
Score:	1167.00	Matches:	223
Percent Similarity:	98.2%	Conservative:	0
Best Local Similarity:	98.2%	Mismatches:	4
Query Match:	97.9%	Indels:	0

```
DB:                  79            Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-455-444-3324 (1-1919)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 617

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        618 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 677

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 737

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        738 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 797

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        798 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 857

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 917

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 977

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1037

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1038 AAACTGGACTGTGGTAGATTT 1058
```

RESULT 50
US-60-465-241-3324
; Sequence 3324, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3324
; LENGTH: 1919
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-465-241-3324

Alignment Scores:

Pred. No.:	8.19e-176	Length:	1919

```
Score:                 1167.00        Matches:       223
Percent Similarity:    98.2%          Conservative:  0
Best Local Similarity: 98.2%          Mismatches:    4
Query Match:           97.9%          Indels:        0
DB:                    79             Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x US-60-465-241-3324 (1-1919)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 437

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 497

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        498 CTAAAGAAGAGATCARGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 557

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 617

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db        618 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCAKGTGAAYGTGTCAGGAAATGAGCTT 677

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 737

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        738 TTAAAAATGGTTTTRAAGAAAATGSAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 797

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        798 GGAGATACAAATCTAAGGGATCRAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 857

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 917

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 977

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1037

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1038 AAACTGGACTGTGGTAGATTT 1058

RESULT 51
PCT-US01-08631-22058
; Sequence 22058, Application PC/TUS0108631
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-049
;  CURRENT APPLICATION NUMBER: PCT/US01/08631
;  CURRENT FILING DATE:  2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22058
;   LENGTH: 1088
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1)..(687)
;   OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=1199.
PCT-US01-08631-22058

Alignment Scores:
Pred. No.:             5.21e-168      Length:        1088
Score:                 1116.00        Matches:       212
Percent Similarity:    99.1%          Conservative:  1
Best Local Similarity: 98.6%          Mismatches:    2
Query Match:           93.6%          Indels:        0
DB:                    1              Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x PCT-US01-08631-22058 (1-1088)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy        213 SerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 645

RESULT 52
US-10-302-689A-136240
; Sequence 136240, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ballinger, Dennis
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
```

```
;  APPLICANT:  Liu, Jin
;  APPLICANT:  Loeb, Deborah
;  APPLICANT:  Montgomery, Julia, R.
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 136240
;   LENGTH: 1088
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-302-689A-136240

Alignment Scores:
Pred. No.:              5.21e-168      Length:          1088
Score:                  1116.00        Matches:         212
Percent Similarity:     99.1%          Conservative:    1
Best Local Similarity:  98.6%          Mismatches:      2
Query Match:            93.6%          Indels:          0
DB:                     48             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-302-689A-136240 (1-1088)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy        213 SerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 645

RESULT 53
US-10-450-763-22058
; Sequence 22058, Application US/10450763
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22058
;   LENGTH: 1088
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1)..(687)
;   OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=1199.
US-10-450-763-22058

Alignment Scores:
Pred. No.:              5.21e-168      Length:          1088
Score:                  1116.00        Matches:         212
Percent Similarity:     99.1%          Conservative:    1
Best Local Similarity:  98.6%          Mismatches:      2
Query Match:            93.6%          Indels:          0
DB:                     51             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-450-763-22058 (1-1088)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112

Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132

Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152

Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172

Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192

Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212

Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy        213 SerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227

Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 645

RESULT 54
PCT-US02-02870-1067
; Sequence 1067, Application PC/TUS0202870
; GENERAL INFORMATION:
;  APPLICANT: Corixa Coporation
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563PC
;  CURRENT APPLICATION NUMBER: PCT/US02/02870
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1067
;   LENGTH: 674
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11, 377, 640, 654
;   OTHER INFORMATION: n = A,T,C or G
PCT-US02-02870-1067

Alignment Scores:
Pred. No.:              3.9e-166      Length:         674
Score:                  1102.00       Matches:        213
Percent Similarity:     98.6%         Conservative:   0
Best Local Similarity:  98.6%         Mismatches:     2
Query Match:            92.4%         Indels:         1
DB:                     1             Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x PCT-US02-02870-1067 (1-674)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32

Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 77

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52

Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72

Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92

Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112

Db        258 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132

Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        133 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152

Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172

Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192

Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProAr 212

Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy        212 gSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227

Db        618 AAGTTTGGACCTTCTTGGATTANAAAAACTGGACTGNGGTAGATTT 663

RESULT 55
US-10-066-543-1067
; Sequence 1067, Application US/10066543
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1067
;   LENGTH: 674
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11, 377, 640, 654
```

```
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1067

Alignment Scores:
Pred. No.:                3.9e-166        Length:         674
Score:                    1102.00         Matches:        213
Percent Similarity:       98.6%           Conservative:   0
Best Local Similarity:    98.6%           Mismatches:     2
Query Match:              92.4%           Indels:         1
DB:                       40              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-1067 (1-674)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 77

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        133 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProAr 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy        212 gSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||||   |||||||||   |||||||||
Db        618 AAGTTTGGACCTTCTTGGATTANAAAAACTGGACTGNGGTAGATTT 663

RESULT 56
US-10-214-403-1067
; Sequence 1067, Application US/10214403
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Chenault, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563C1
;  CURRENT APPLICATION NUMBER: US/10/214,403
;  CURRENT FILING DATE:  2002-08-06
;  NUMBER OF SEQ ID NOS: 3420
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1067
;   LENGTH: 674
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11, 377, 640, 654
;   OTHER INFORMATION: n = A,T,C or G
US-10-214-403-1067

Alignment Scores:
Pred. No.:                3.9e-166        Length:         674
Score:                    1102.00         Matches:        213
Percent Similarity:       98.6%           Conservative:   0
Best Local Similarity:    98.6%           Mismatches:     2
Query Match:              92.4%           Indels:         1
DB:                       43              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-214-403-1067 (1-674)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 77

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        133 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProAr 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy        212 gSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||||   |||||||||   |||||||||
Db        618 AAGTTTGGACCTTCTTGGATTANAAAAACTGGACTGNGGTAGATTT 663

RESULT 57
US-09-248-178-63/c
; Sequence 63, Application US/09248178
```

```
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  TITLE OF INVENTION:  COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS OF BREAST CANCER AND
;  TITLE OF INVENTION:  METHODS FOR THEIR USE
;  FILE REFERENCE:  210121.446C2
;  CURRENT APPLICATION NUMBER: US/09/248,178
;  CURRENT FILING DATE:  1999-02-09
;  NUMBER OF SEQ ID NOS:  94
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-248-178-63

Alignment Scores:
Pred. No.:              8.48e-164       Length:         1079
Score:                  1090.00         Matches:        211
Percent Similarity:     99.5%           Conservative:   0
Best Local Similarity:  99.5%           Mismatches:     1
Query Match:            91.4%           Indels:         1
DB:                     22              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-248-178-63 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 58
US-09-288-950-63/c
```

```
; Sequence 63, Application US/09288950
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Xu, Jiangchun
;  APPLICANT: Dillon, Davin C.
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
;  TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.446C3
;  CURRENT APPLICATION NUMBER: US/09/288,950
;  CURRENT FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-288-950-63

Alignment Scores:
Pred. No.:              8.48e-164       Length:         1079
Score:                  1090.00         Matches:        211
Percent Similarity:     99.5%           Conservative:   0
Best Local Similarity:  99.5%           Mismatches:     1
Query Match:            91.4%           Indels:         1
DB:                     22              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-288-950-63 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 59
```

US-09-346-327-63/c
; Sequence 63, Application US/09346327
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C4
; CURRENT APPLICATION NUMBER: US/09/346,327
; CURRENT FILING DATE: 1999-07-02
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-346-327-63

Alignment Scores:

Pred. No.:	8.48e-164	Length:	1079
Score:	1090.00	Matches:	211
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	91.4%	Indels:	1
DB:	23	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-346-327-63 (1-1079)

```
Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443
```

RESULT 60
US-09-687-507-63/c
; Sequence 63, Application US/09687507
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C6
; CURRENT APPLICATION NUMBER: US/09/687,507
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-687-507-63

Alignment Scores:

Pred. No.:	8.48e-164	Length:	1079
Score:	1090.00	Matches:	211
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	91.4%	Indels:	1
DB:	29	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-687-507-63 (1-1079)

```
Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||
```

```
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 61
US-09-687-507A-63/c
; Sequence 63, Application US/09687507A
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Retter, Marc W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.446C6
;  CURRENT APPLICATION NUMBER: US/09/687,507A
;  CURRENT FILING DATE:  2000-10-12
;  NUMBER OF SEQ ID NOS: 116
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-687-507A-63

Alignment Scores:
Pred. No.:              8.48e-164      Length:         1079
Score:                  1090.00        Matches:        211
Percent Similarity:     99.5%          Conservative:   0
Best Local Similarity:  99.5%          Mismatches:     1
Query Match:            91.4%          Indels:         1
DB:                     29             Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-687-507A-63 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479
```

```
Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 62
US-09-745-288-63/c
; Sequence 63, Application US/09745288
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  TITLE OF INVENTION:  COMPOUNDS FOR IMMUNOTHERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE:  210121.446D1
;  CURRENT APPLICATION NUMBER: US/09/745,288
;  CURRENT FILING DATE:  2000-12-19
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-745-288-63

Alignment Scores:
Pred. No.:              8.48e-164      Length:         1079
Score:                  1090.00        Matches:        211
Percent Similarity:     99.5%          Conservative:   0
Best Local Similarity:  99.5%          Mismatches:     1
Query Match:            91.4%          Indels:         1
DB:                     31             Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-745-288-63 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479
```

```
Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443
```

RESULT 63
US-09-778-381-63/c
; Sequence 63, Application US/09778381
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C7
; CURRENT APPLICATION NUMBER: US/09/778,381
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-778-381-63

Alignment Scores:

Pred. No.:	8.48e-164	Length:	1079
Score:	1090.00	Matches:	211
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	91.4%	Indels:	1
DB:	32	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-778-381-63 (1-1079)

```
Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

```
Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443
```

RESULT 64
US-10-453-919-63/c
; Sequence 63, Application US/10453919
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C7
; CURRENT APPLICATION NUMBER: US/10/453,919
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-453-919-63

Alignment Scores:

Pred. No.:	8.48e-164	Length:	1079
Score:	1090.00	Matches:	211
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	91.4%	Indels:	1
DB:	51	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-453-919-63 (1-1079)

```
Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
```
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
                |||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 65
PCT-US02-02870-1070
; Sequence 1070, Application PC/TUS0202870
; GENERAL INFORMATION:
;  APPLICANT: Corixa Coporation
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563PC
;  CURRENT APPLICATION NUMBER: PCT/US02/02870
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1070
;   LENGTH: 644
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US02-02870-1070

Alignment Scores:
Pred. No.:              8.07e-161       Length:         644
Score:                  1069.00         Matches:        203
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  99.5%           Mismatches:     0
Query Match:            89.7%           Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x PCT-US02-02870-1070 (1-644)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
                ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||
Db        332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy        213 SerLeuAspLeu 216
                ||||||||||||
Db        632 AGTTTGGACCTT 643

RESULT 66
US-10-066-543-1070
; Sequence 1070, Application US/10066543
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1070
;   LENGTH: 644
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-066-543-1070

Alignment Scores:
Pred. No.:              8.07e-161       Length:         644
Score:                  1069.00         Matches:        203
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  99.5%           Mismatches:     0
Query Match:            89.7%           Indels:         0
DB:                     40              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-1070 (1-644)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy        213 SerLeuAspLeu 216
              ||||||||||||
Db        632 AGTTTGGACCTT 643

RESULT 67
US-10-214-403-1070
; Sequence 1070, Application US/10214403
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Chenault, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563C1
;  CURRENT APPLICATION NUMBER: US/10/214,403
;  CURRENT FILING DATE:  2002-08-06
;  NUMBER OF SEQ ID NOS: 3420
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1070
;   LENGTH: 644
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-214-403-1070

Alignment Scores:
Pred. No.:              8.07e-161       Length:         644
Score:                  1069.00         Matches:        203
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  99.5%           Mismatches:     0
Query Match:            89.7%           Indels:         0
DB:                     43              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-214-403-1070 (1-644)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy        213 SerLeuAspLeu 216
              ||||||||||||
Db        632 AGTTTGGACCTT 643

RESULT 68
PCT-US02-02870-1084
; Sequence 1084, Application PC/TUS0202870
; GENERAL INFORMATION:
;  APPLICANT: Corixa Coporation
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563PC
;  CURRENT APPLICATION NUMBER: PCT/US02/02870
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1084
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3
;   OTHER INFORMATION: n = A,T,C or G
PCT-US02-02870-1084

Alignment Scores:
```

```
Pred. No.:              2.12e-158       Length:         633
Score:                  1054.00         Matches:        200
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            88.4%           Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x PCT-US02-02870-1084 (1-633)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 633

RESULT 69
US-10-066-543-1084
; Sequence 1084, Application US/10066543
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1084
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1084

Alignment Scores:
Pred. No.:              2.12e-158       Length:         633
Score:                  1054.00         Matches:        200
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            88.4%           Indels:         0
DB:                     40              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-1084 (1-633)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 633

RESULT 70
US-10-214-403-1084
; Sequence 1084, Application US/10214403
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Chenault, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
```

```
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563C1
;  CURRENT APPLICATION NUMBER: US/10/214,403
;  CURRENT FILING DATE:  2002-08-06
;  NUMBER OF SEQ ID NOS: 3420
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1084
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3
;   OTHER INFORMATION: n = A,T,C or G
US-10-214-403-1084

Alignment Scores:
Pred. No.:              2.12e-158       Length:         633
Score:                  1054.00         Matches:        200
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            88.4%           Indels:         0
DB:                     43              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-214-403-1084 (1-633)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 633

RESULT 71
US-10-170-235-26834/c
; Sequence 26834, Application US/10170235
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig
;  TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
;  TITLE OF INVENTION:  TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
;  FILE REFERENCE: CL001380
;  CURRENT APPLICATION NUMBER: US/10/170,235
;  CURRENT FILING DATE:  2003-03-17
;  NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 26834
;   LENGTH: 1849
;   TYPE: DNA
;   ORGANISM: HUMAN
US-10-170-235-26834

Alignment Scores:
Pred. No.:              2.12e-155       Length:         1849
Score:                  1041.00         Matches:        198
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            87.3%           Indels:         0
DB:                     42              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-170-235-26834 (1-1849)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 535

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 475

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 415

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 355

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 295

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 235

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 175

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 115

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 55

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 1

RESULT 72
PCT-US02-02870-548
; Sequence 548, Application PC/TUS0202870
; GENERAL INFORMATION:
;  APPLICANT: Corixa Coporation
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
```

```
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563PC
;  CURRENT APPLICATION NUMBER: PCT/US02/02870
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 19, 71
;   OTHER INFORMATION: n = A,T,C or G
PCT-US02-02870-548

Alignment Scores:
Pred. No.:                1.12e-154       Length:         625
Score:                    1031.00         Matches:        198
Percent Similarity:       99.0%           Conservative:   0
Best Local Similarity:    99.0%           Mismatches:     1
Query Match:              86.5%           Indels:         1
DB:                       1               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x PCT-US02-02870-548 (1-625)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy         53 -GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 72
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy         72 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy         92 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        112 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy        132 oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 623

RESULT 73
US-10-066-543-548

; Sequence 548, Application US/10066543
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 19, 71
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-548

Alignment Scores:
Pred. No.:                1.12e-154       Length:         625
Score:                    1031.00         Matches:        198
Percent Similarity:       99.0%           Conservative:   0
Best Local Similarity:    99.0%           Mismatches:     1
Query Match:              86.5%           Indels:         1
DB:                       40              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-548 (1-625)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy         53 -GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 72
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy         72 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy         92 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        112 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy        132 oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 211
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 623

RESULT 74
US-10-214-403-548
; Sequence 548, Application US/10214403
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Chenault, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563C1
;  CURRENT APPLICATION NUMBER: US/10/214,403
;  CURRENT FILING DATE:  2002-08-06
;  NUMBER OF SEQ ID NOS: 3420
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 19, 71
;   OTHER INFORMATION: n = A,T,C or G
US-10-214-403-548

Alignment Scores:
Pred. No.:              1.12e-154       Length:         625
Score:                  1031.00         Matches:        198
Percent Similarity:     99.0%           Conservative:   0
Best Local Similarity:  99.0%           Mismatches:     1
Query Match:            86.5%           Indels:         1
DB:                     43              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-214-403-548 (1-625)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
                 ||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy         53 -GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 72
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy         72 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 92
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy         92 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 112
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        112 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 132
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy        132 oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 211
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 623

RESULT 75
PCT-US02-02870-1031
; Sequence 1031, Application PC/TUS0202870
; GENERAL INFORMATION:
;  APPLICANT: Corixa Coporation
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563PC
;  CURRENT APPLICATION NUMBER: PCT/US02/02870
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1031
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 10
;   OTHER INFORMATION: n = A,T,C or G
PCT-US02-02870-1031

Alignment Scores:
Pred. No.:              6.79e-153       Length:         625
Score:                  1020.00         Matches:        194
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            85.6%           Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x PCT-US02-02870-1031 (1-625)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 581

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 206
              |||||||||||||||||||||||||||||||||||||||
Db        582 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAG 623

Search completed: March 11, 2006, 12:06:25
Job time : 9370.96 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:22:00 ; Search time 16.3777 Seconds
(without alignments)
1013.401 Million cell updates/sec

Title: US-10-757-745-2_COPY_115_121
Perfect score: 38
Sequence: 1 SLITWNI 7

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5007642 seqs, 1185514400 residues

Total number of hits satisfying chosen parameters: 10015284

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170335_4211/app_query.fasta_1
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_240_@runat_09032006_170335_4211 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq1:*
9: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	752	6	US-10-917-503B-4908	Sequence 4908, Ap
2	38	100.0	1288	9	US-11-266-748A-226184	Sequence 226184,
3	38	100.0	1898	6	US-10-917-503B-13209	Sequence 13209, A
4	38	100.0	1940	6	US-10-567-867-921	Sequence 921, App
5	38	100.0	1940	11	US-60-742-219-2071	Sequence 2071, Ap
6	36	94.7	324	9	US-11-266-748A-412604	Sequence 412604,
7	36	94.7	438	7	US-11-330-364-16964	Sequence 16964, A
c 8	36	94.7	659	7	US-11-353-150-23111	Sequence 23111, A
c 9	36	94.7	816	9	US-11-266-748A-74302	Sequence 74302, A
10	36	94.7	816	9	US-11-266-748A-127113	Sequence 127113,
c 11	36	94.7	972	6	US-10-461-673-984	Sequence 984, App
c 12	36	94.7	993	9	US-11-266-748A-56667	Sequence 56667, A
c 13	36	94.7	993	9	US-11-266-748A-394321	Sequence 394321,
14	36	94.7	993	9	US-11-266-748A-465367	Sequence 465367,
15	36	94.7	1000	9	US-11-266-748A-197416	Sequence 197416,
c 16	36	94.7	1000	9	US-11-266-748A-198059	Sequence 198059,
17	36	94.7	1000	9	US-11-266-748A-202645	Sequence 202645,
c 18	36	94.7	1000	9	US-11-266-748A-224243	Sequence 224243,
c 19	36	94.7	1000	9	US-11-266-748A-292898	Sequence 292898,
20	36	94.7	1000	9	US-11-266-748A-344327	Sequence 344327,
c 21	36	94.7	1000	9	US-11-266-748A-404749	Sequence 404749,
22	36	94.7	1000	9	US-11-266-748A-475795	Sequence 475795,
c 23	36	94.7	1054	9	US-11-266-748A-18075	Sequence 18075, A
c 24	36	94.7	1061	9	US-11-266-748A-251826	Sequence 251826,
25	36	94.7	1061	9	US-11-266-748A-312343	Sequence 312343,
c 26	36	94.7	1219	9	US-11-266-748A-182011	Sequence 182011,
27	36	94.7	1219	9	US-11-266-748A-242559	Sequence 242559,
28	36	94.7	2300	11	US-60-762-056-22013	Sequence 22013, A
c 29	36	94.7	2300	11	US-60-762-056-24854	Sequence 24854, A
30	36	94.7	2300	11	US-60-762-056-24855	Sequence 24855, A
c 31	36	94.7	2300	11	US-60-762-056-35417	Sequence 35417, A
c 32	36	94.7	2560	6	US-10-461-673-1106	Sequence 1106, Ap
c 33	36	94.7	2700	6	US-10-276-817B-4457	Sequence 4457, Ap
c 34	36	94.7	3466	6	US-10-515-325-1100	Sequence 1100, Ap
c 35	36	94.7	3466	6	US-10-515-325-1336	Sequence 1336, Ap
c 36	36	94.7	3758	11	US-60-732-162-261	Sequence 261, App
c 37	36	94.7	3758	11	US-60-751-455-2043	Sequence 2043, Ap
38	36	94.7	3962	6	US-10-461-673-75	Sequence 75, Appl
39	36	94.7	684973	9	US-11-266-748A-32013	Sequence 32013, A
c 40	35	92.1	62	9	US-11-130-645B-322741	Sequence 322741,
41	35	92.1	63	9	US-11-130-645B-71304	Sequence 71304, A
42	35	92.1	64	9	US-11-130-645B-52383	Sequence 52383, A
43	35	92.1	64	9	US-11-130-645B-332515	Sequence 332515,
c 44	35	92.1	286	7	US-11-330-082-1351	Sequence 1351, Ap
45	35	92.1	354	9	US-11-266-748A-430212	Sequence 430212,
46	35	92.1	536	9	US-11-266-748A-219748	Sequence 219748,
47	35	92.1	1620	9	US-11-266-748A-228613	Sequence 228613,
48	35	92.1	1831	11	US-60-751-455-2154	Sequence 2154, Ap
49	35	92.1	1865	9	US-11-266-748A-162764	Sequence 162764,
50	35	92.1	1894	9	US-11-266-748A-19824	Sequence 19824, A
51	35	92.1	1907	9	US-11-266-748A-253327	Sequence 253327,
c 52	35	92.1	1907	9	US-11-266-748A-313844	Sequence 313844,
c 53	35	92.1	2300	11	US-60-762-056-16200	Sequence 16200, A
c 54	35	92.1	2300	11	US-60-762-056-23584	Sequence 23584, A
55	35	92.1	3378	6	US-10-276-817B-1699	Sequence 1699, Ap
c 56	35	92.1	6051	9	US-11-266-748A-56674	Sequence 56674, A
c 57	35	92.1	6051	11	US-60-742-219-1999	Sequence 1999, Ap
c 58	35	92.1	246282	9	US-11-266-748A-32092	Sequence 32092, A
59	35	92.1	440684	6	US-10-540-898-114	Sequence 114, App
60	34	89.5	19	8	US-11-313-452-984899	Sequence 984899,
c 61	34	89.5	340	9	US-11-227-183A-10563	Sequence 10563, A
c 62	34	89.5	354	9	US-11-313-816-11722	Sequence 11722, A
63	34	89.5	390	7	US-11-329-388-2343	Sequence 2343, Ap
64	34	89.5	424	9	US-11-313-816-11719	Sequence 11719, A
65	34	89.5	428	7	US-11-353-150-66475	Sequence 66475, A
66	34	89.5	429	7	US-11-329-388-2953	Sequence 2953, Ap
c 67	34	89.5	451	9	US-11-266-748A-266797	Sequence 266797,
c 68	34	89.5	451	9	US-11-266-748A-280499	Sequence 280499,
69	34	89.5	451	9	US-11-266-748A-327314	Sequence 327314,
70	34	89.5	459	9	US-11-311-886-3178	Sequence 3178, Ap
c 71	34	89.5	468	9	US-11-313-816-7526	Sequence 7526, Ap
72	34	89.5	581	9	US-11-266-748A-480077	Sequence 480077,
73	34	89.5	660	9	US-11-266-748A-232881	Sequence 232881,
74	34	89.5	841	9	US-11-266-748A-381213	Sequence 381213,
c 75	34	89.5	841	9	US-11-266-748A-430969	Sequence 430969,

ALIGNMENTS

RESULT 1

```
US-10-917-503B-4908
; Sequence 4908, Application US/10917503B
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4908
;   LENGTH: 752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (685)..(685)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (694)..(694)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (740)..(740)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-10-917-503B-4908

Alignment Scores:
Pred. No.:                578             Length:         752
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-917-503B-4908 (1-752)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363

RESULT 2
US-11-266-748A-226184
; Sequence 226184, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 226184
;   LENGTH: 1288
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-226184

Alignment Scores:
Pred. No.:                1.01e+03        Length:         1288
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-226184 (1-1288)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        195 TCTCTCATTACCTGGAATATT 215

RESULT 3
US-10-917-503B-13209
; Sequence 13209, Application US/10917503B
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
```

; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1086)
US-10-917-503B-13209

Alignment Scores:

Pred. No.:	1.5e+03	Length:	1898
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-917-503B-13209 (1-1898)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363
```

RESULT 4
US-10-567-867-921
; Sequence 921, Application US/10567867
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PCT
; CURRENT APPLICATION NUMBER: US/10/567,867
; CURRENT FILING DATE: 2006-01-27
; PRIOR APPLICATION NUMBER: US 60/490,890
; PRIOR FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2786
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-567-867-921

Alignment Scores:

Pred. No.:	1.53e+03	Length:	1940
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-567-867-921 (1-1940)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        368 TCTCTCATTACCTGGAATATT 388
```

RESULT 5
US-60-742-219-2071
; Sequence 2071, Application US/60742219
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John Verner
; APPLICANT: Bradley, Walter Edward
; APPLICANT: Paquin, Bruno
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Little, Randall David
; APPLICANT: Keith, Tim
; APPLICANT: Cousineau, Johanne
; APPLICANT: Eerdewegh, Paul Van
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
; FILE REFERENCE: 306522-2000
; CURRENT APPLICATION NUMBER: US/60/742,219
; CURRENT FILING DATE: 2005-12-05
; NUMBER OF SEQ ID NOS: 7303
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2071
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-742-219-2071

Alignment Scores:

Pred. No.:	1.53e+03	Length:	1940
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-742-219-2071 (1-1940)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        368 TCTCTCATTACCTGGAATATT 388
```

RESULT 6
US-11-266-748A-412604
; Sequence 412604, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 412604
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:

```
;   NAME/KEY: misc_feature
;   LOCATION: (162)..(162)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (170)..(170)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (178)..(178)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (184)..(184)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-412604

Alignment Scores:
Pred. No.:                592             Length:         324
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-412604 (1-324)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db         72 TCCCTCATAACTTGGAATCTA 92

RESULT 7
US-11-330-364-16964
; Sequence 16964, Application US/11330364
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S.
;  APPLICANT:  Andersen, Scott E.
;  APPLICANT:  Dubois, Patrice
;  APPLICANT:  Mahadeo, Debbie A.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.352 - 38-21(51721)C/US
;  CURRENT APPLICATION NUMBER: US/11/330,364
;  CURRENT FILING DATE:  2006-01-12
;  PRIOR APPLICATION NUMBER: US 09/696,664
;  PRIOR FILING DATE: 2000-10-25
;  PRIOR APPLICATION NUMBER: US 60/161,619
;  PRIOR FILING DATE: 1999-10-26
;  NUMBER OF SEQ ID NOS: 17472
; SEQ ID NO 16964
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3279-200-Q1-N6-F3
US-11-330-364-16964

Alignment Scores:
Pred. No.:                808             Length:         438
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-330-364-16964 (1-438)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db          2 TCGATAATAACATGGAACATT 22

RESULT 8
US-11-353-150-23111/c
; Sequence 23111, Application US/11353150
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Scott E.
;  APPLICANT:  Conner, Timothy W.
;  APPLICANT:  Lalgudi, Raghunath V.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.365 - 38-21(15459)D/US
;  CURRENT APPLICATION NUMBER: US/11/353,150
;  CURRENT FILING DATE:  2006-02-14
;  PRIOR APPLICATION NUMBER: 09/565,386
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/132,860
;  PRIOR FILING DATE: 1999-05-07
;  NUMBER OF SEQ ID NOS: 83523
; SEQ ID NO 23111
;   LENGTH: 659
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: cC-zmle189044h10d1
US-11-353-150-23111

Alignment Scores:
Pred. No.:                1.23e+03        Length:         659
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-353-150-23111 (1-659)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db         30 TCGATAATAACATGGAACATT 10

RESULT 9
US-11-266-748A-74302/c
; Sequence 74302, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 74302
;   LENGTH: 816
;   TYPE: DNA
```

; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (22)..(22)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (670)..(745)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-74302

Alignment Scores:

Pred. No.:	1.53e+03	Length:	816
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-74302 (1-816)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db         73 TCATTACTAACTTGGAATATT 53
```

RESULT 10
US-11-266-748A-127113
; Sequence 127113, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 127113
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (72)..(147)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (795)..(795)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-127113

Alignment Scores:

Pred. No.:	1.53e+03	Length:	816
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-127113 (1-816)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        744 TCATTACTAACTTGGAATATT 764
```

RESULT 11
US-10-461-673-984/c
; Sequence 984, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yunqing
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Ghosh, Malabika
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 984
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (227)..(952)
US-10-461-673-984

Alignment Scores:

```
Pred. No.:              1.84e+03        Length:         972
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-461-673-984 (1-972)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        199 AGCATTATTACGTGGAACATC 179

RESULT 12
US-11-266-748A-56667/c
; Sequence 56667, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56667
;   LENGTH: 993
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-56667

Alignment Scores:
Pred. No.:              1.88e+03        Length:         993
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-56667 (1-993)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db         32 AGCATTATTACGTGGAACATC 12

RESULT 13
US-11-266-748A-394321/c
; Sequence 394321, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 394321
;   LENGTH: 993
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-394321

Alignment Scores:
Pred. No.:              1.88e+03        Length:         993
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-394321 (1-993)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db         32 AGCATTATTACGTGGAACATC 12

RESULT 14
US-11-266-748A-465367
; Sequence 465367, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 465367
;   LENGTH: 993
;   TYPE: DNA
```

```
;   ORGANISM: Homo Sapiens
US-11-266-748A-465367

Alignment Scores:
Pred. No.:                1.88e+03        Length:         993
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-465367 (1-993)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        962 AGCATTATTACGTGGAACATC 982

RESULT 15
US-11-266-748A-197416
; Sequence 197416, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 197416
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-197416

Alignment Scores:
Pred. No.:                1.89e+03        Length:         1000
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-197416 (1-1000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        925 TCAATAATAACATGGAATATA 945

RESULT 16
US-11-266-748A-198059/c
; Sequence 198059, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 198059
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-198059

Alignment Scores:
Pred. No.:                1.89e+03        Length:         1000
Score:                    36.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-198059 (1-1000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        744 TCCTTAGTGACATGGAATGTG 724

RESULT 17
US-11-266-748A-202645
; Sequence 202645, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
```

```
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 202645
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-202645

Alignment Scores:
Pred. No.:                1.89e+03        Length:         1000
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-202645 (1-1000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||||||||||||||:::
Db        350 AGCTTAATTACGTGGAACTTA 370

RESULT 18
US-11-266-748A-224243/c
; Sequence 224243, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 224243
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-224243

Alignment Scores:
Pred. No.:                1.89e+03        Length:         1000
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-224243 (1-1000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        257 TCATTACTAACTTGGAATATT 237

RESULT 19
US-11-266-748A-292898/c
; Sequence 292898, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 292898
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-292898

Alignment Scores:
Pred. No.:                1.89e+03        Length:         1000
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              94.7%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-292898 (1-1000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        257 TCATTACTAACTTGGAATATT 237

RESULT 20
US-11-266-748A-344327
; Sequence 344327, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
```

```
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 344327
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-344327

Alignment Scores:
Pred. No.:              1.89e+03        Length:         1000
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-344327 (1-1000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        744 TCATTACTAACTTGGAATATT 764

RESULT 21
US-11-266-748A-404749/c
; Sequence 404749, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 404749
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-404749

Alignment Scores:
Pred. No.:              1.89e+03        Length:         1000
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-404749 (1-1000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        257 TCATTACTAACTTGGAATATT 237

RESULT 22
US-11-266-748A-475795
; Sequence 475795, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 475795
;   LENGTH: 1000
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-475795

Alignment Scores:
Pred. No.:              1.89e+03        Length:         1000
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-475795 (1-1000)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db        744 TCATTACTAACTTGGAATATT 764

RESULT 23
US-11-266-748A-18075/c
; Sequence 18075, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
```

```
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18075
;   LENGTH: 1054
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-18075

Alignment Scores:
Pred. No.:              2e+03       Length:         1054
Score:                  36.00       Matches:        6
Percent Similarity:     100.0%      Conservative:   1
Best Local Similarity:  85.7%       Mismatches:     0
Query Match:            94.7%       Indels:         0
DB:                     9           Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-18075 (1-1054)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        414 TCATTACTAACTTGGAATATT 394

RESULT 24
US-11-266-748A-251826/c
; Sequence 251826, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 251826
;   LENGTH: 1061
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (27)..(27)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (943)..(943)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-251826

Alignment Scores:
Pred. No.:              2.01e+03    Length:         1061
Score:                  36.00       Matches:        6
Percent Similarity:     100.0%      Conservative:   1
Best Local Similarity:  85.7%       Mismatches:     0
Query Match:            94.7%       Indels:         0
DB:                     9           Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-251826 (1-1061)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db         82 AGCATTATTACGTGGAACATC 62

RESULT 25
US-11-266-748A-312343
; Sequence 312343, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 312343
;   LENGTH: 1061
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (119)..(119)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1035)..(1035)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-312343

Alignment Scores:
Pred. No.:              2.01e+03    Length:         1061
Score:                  36.00       Matches:        6
Percent Similarity:     100.0%      Conservative:   1
Best Local Similarity:  85.7%       Mismatches:     0
Query Match:            94.7%       Indels:         0
DB:                     9           Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-312343 (1-1061)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        980 AGCATTATTACGTGGAACATC 1000

RESULT 26
```

US-11-266-748A-182011/c
; Sequence 182011, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 182011
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-182011

Alignment Scores:

Pred. No.:	2.32e+03	Length:	1219
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-182011 (1-1219)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        188 AGCATTATTACGTGGAACATC 168
```

RESULT 27
US-11-266-748A-242559
; Sequence 242559, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 242559
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-242559

Alignment Scores:

Pred. No.:	2.32e+03	Length:	1219
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-242559 (1-1219)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db       1032 AGCATTATTACGTGGAACATC 1052
```

RESULT 28
US-60-762-056-22013
; Sequence 22013, Application US/60762056
; GENERAL INFORMATION:
; APPLICANT: Trinklein, Nathan D
; APPLICANT: Aldred, Shelley Force
; APPLICANT: Cooper, Sara J
; APPLICANT: Myers, Richard M
; TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
; TITLE OF INVENTION: EXPRESSION REGULATORY ELEMENTS
; FILE REFERENCE: 33102-701.102
; CURRENT APPLICATION NUMBER: US/60/762,056
; CURRENT FILING DATE: 2006-01-24
; NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 22013
; LENGTH: 2300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-762-056-22013

Alignment Scores:

Pred. No.:	4.46e+03	Length:	2300
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	94.7%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-762-056-22013 (1-2300)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||||
Db        235 TCTATCATTACTTGGAATATA 255
```

RESULT 29
US-60-762-056-24854/c
; Sequence 24854, Application US/60762056
; GENERAL INFORMATION:
; APPLICANT: Trinklein, Nathan D
; APPLICANT: Aldred, Shelley Force
; APPLICANT: Cooper, Sara J
; APPLICANT: Myers, Richard M
; TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
; TITLE OF INVENTION: EXPRESSION REGULATORY ELEMENTS
; FILE REFERENCE: 33102-701.102
; CURRENT APPLICATION NUMBER: US/60/762,056
; CURRENT FILING DATE: 2006-01-24

```
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 24854
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-24854

Alignment Scores:
Pred. No.:              4.46e+03        Length:         2300
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-762-056-24854 (1-2300)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::||||||||||||||||
Db       2132 AGCATTATTACGTGGAACATC 2112

RESULT 30
US-60-762-056-24855
; Sequence 24855, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 24855
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-24855

Alignment Scores:
Pred. No.:              4.46e+03        Length:         2300
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-762-056-24855 (1-2300)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::||||||||||||||||
Db       1205 AGCATTATTACGTGGAACATC 1225

RESULT 31
US-60-762-056-35417/c
; Sequence 35417, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 35417
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-35417

Alignment Scores:
Pred. No.:              4.46e+03        Length:         2300
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-762-056-35417 (1-2300)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2109 TCCCTCATAACTTGGAATCTA 2089

RESULT 32
US-10-461-673-1106/c
; Sequence 1106, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 1106
;   LENGTH: 2560
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (320)..(2545)
US-10-461-673-1106

Alignment Scores:
Pred. No.:              4.98e+03        Length:         2560
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-461-673-1106 (1-2560)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       2336 TCATTACTAACTTGGAATATT 2316

RESULT 33
US-10-276-817B-4457/c
; Sequence 4457, Application US/10276817B
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-104
;  CURRENT APPLICATION NUMBER: US/10/276,817B
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: PCT/US01/14827
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 16102
;  SOFTWARE: Custom
; SEQ ID NO 4457
;   LENGTH: 2700
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (153)..(1517)
;   OTHER INFORMATION: 95% homologous to Homo sapiens Human SEC10 protein sequence
;   OTHER INFORMATION: (clone ID 1795045.0.77),accession number AAB48377,Smith-Waterman
;   OTHER INFORMATION: Score=2431.
US-10-276-817B-4457

Alignment Scores:
Pred. No.:              5.26e+03        Length:         2700
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-276-817B-4457 (1-2700)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       2434 TCATTACTAACTTGGAATATT 2414

RESULT 34
US-10-515-325-1100/c
; Sequence 1100, Application US/10515325
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Higgs, Brandon
;  APPLICANT:  Elashoff, Michael
;  TITLE OF INVENTION: Molecular Nephrotoxicology Modeling
;  FILE REFERENCE: 044921-5089-02-US
;  CURRENT APPLICATION NUMBER: US/10/515,325
;  CURRENT FILING DATE:  2004-11-23
;  PRIOR APPLICATION NUMBER: US 10/301,856
;  PRIOR FILING DATE: 2002-11-22
;  PRIOR APPLICATION NUMBER: PCT/US2003/037556
;  PRIOR FILING DATE: 2003-11-24
;  NUMBER OF SEQ ID NOS: 1617
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1100
;   LENGTH: 3466
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-10-515-325-1100

Alignment Scores:
Pred. No.:              6.81e+03        Length:         3466
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-515-325-1100 (1-3466)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2097 TCTCTCATCACATGGAATCTG 2077

RESULT 35
US-10-515-325-1336/c
; Sequence 1336, Application US/10515325
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Higgs, Brandon
;  APPLICANT:  Elashoff, Michael
;  TITLE OF INVENTION: Molecular Nephrotoxicology Modeling
;  FILE REFERENCE: 044921-5089-02-US
;  CURRENT APPLICATION NUMBER: US/10/515,325
;  CURRENT FILING DATE:  2004-11-23
;  PRIOR APPLICATION NUMBER: US 10/301,856
;  PRIOR FILING DATE: 2002-11-22
;  PRIOR APPLICATION NUMBER: PCT/US2003/037556
;  PRIOR FILING DATE: 2003-11-24
;  NUMBER OF SEQ ID NOS: 1617
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1336
;   LENGTH: 3466
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-10-515-325-1336

Alignment Scores:
Pred. No.:              6.81e+03        Length:         3466
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-515-325-1336 (1-3466)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2097 TCTCTCATCACATGGAATCTG 2077

RESULT 36
US-60-732-162-261/c
; Sequence 261, Application US/60732162
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John V
```

```
;  APPLICANT:  Bradley, Walter E
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Fournier, Helene
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Debrus, Sophie
;  APPLICANT:  Eerdewegh, Paul V
;  APPLICANT:  Little, Randall D
;  APPLICANT:  Keith , Tim
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Genemap of the Human Genes Associated With Asthma Disease
;  FILE REFERENCE: 059908-5010-PR
;  CURRENT APPLICATION NUMBER: US/60/732,162
;  CURRENT FILING DATE:  2005-11-02
;  NUMBER OF SEQ ID NOS: 4417
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 261
;   LENGTH: 3758
;   TYPE: DNA
;   ORGANISM: Homosapiens
US-60-732-162-261

Alignment Scores:
Pred. No.:                7.4e+03        Length:          3758
Score:                    36.00          Matches:         6
Percent Similarity:       100.0%         Conservative:    1
Best Local Similarity:    85.7%          Mismatches:      0
Query Match:              94.7%          Indels:          0
DB:                       11             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-732-162-261 (1-3758)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2315 TCTCTCATCACATGGAATCTG 2295

RESULT 37
US-60-751-455-2043/c
; Sequence 2043, Application US/60751455
; GENERAL INFORMATION:
;  APPLICANT: Radich, Jerald P.
;  APPLICANT:  Dai, Hongyue
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Schelter, Janell
;  APPLICANT:  Linsley, Peter S.
;  TITLE OF INVENTION: GENES ASSOCIATED WITH PROGRESSION AND RESPONSE IN
;  TITLE OF INVENTION:  CHRONIC MYELOID LEUKEMIA AND USES THEREOF
;  FILE REFERENCE: 9301-253-888
;  CURRENT APPLICATION NUMBER: US/60/751,455
;  CURRENT FILING DATE:  2005-12-15
;  NUMBER OF SEQ ID NOS: 7936
;  SOFTWARE: FastSEQ 4.01
; SEQ ID NO 2043
;   LENGTH: 3758
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-751-455-2043

Alignment Scores:
Pred. No.:                7.4e+03        Length:          3758
Score:                    36.00          Matches:         6
Percent Similarity:       100.0%         Conservative:    1
Best Local Similarity:    85.7%          Mismatches:      0
Query Match:              94.7%          Indels:          0
DB:                       11             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-751-455-2043 (1-3758)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||:::
Db       2315 TCTCTCATCACATGGAATCTG 2295

RESULT 38
US-10-461-673-75
; Sequence 75, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 75
;   LENGTH: 3962
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (652)..(3900)
US-10-461-673-75

Alignment Scores:
Pred. No.:                7.81e+03       Length:          3962
Score:                    36.00          Matches:         5
Percent Similarity:       100.0%         Conservative:    2
Best Local Similarity:    71.4%          Mismatches:      0
Query Match:              94.7%          Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-461-673-75 (1-3962)

Qy          1 SerLeuIleThrTrpAsnIle 7
```

```
                  ||||||:::|||||||||:::
Db       2351 AGCTTAGTAACTTGGAATGTT 2371

RESULT 39
US-11-266-748A-32013
; Sequence 32013, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32013
;   LENGTH: 684973
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-32013

Alignment Scores:
Pred. No.:              1.46e+06        Length:         684973
Score:                  36.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            94.7%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-32013 (1-684973)

Qy          1 SerLeuIleThrTrpAsnIle 7
                ||||||||||||||||||:::
Db     398299 TCTTTAATTACCTGGAATTTG 398319

RESULT 40
US-11-130-645B-322741/c
; Sequence 322741, Application US/11130645B
; GENERAL INFORMATION:
;  APPLICANT: ROSETTA GENOMICS LTD
;  TITLE OF INVENTION: MICRORNAS AND USES THEREOF
;  FILE REFERENCE: 06087.0202.CPUS13
;  CURRENT APPLICATION NUMBER: US/11/130,645B
;  CURRENT FILING DATE:  2005-05-16
;  NUMBER OF SEQ ID NOS: 760616
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 322741
;   LENGTH: 62
;   TYPE: RNA
;   ORGANISM: hsa
US-11-130-645B-322741

Alignment Scores:
Pred. No.:              168             Length:         62
Score:                  35.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-130-645B-322741 (1-62)

Qy          1 SerLeuIleThrTrpAsnIle 7
                ||||||:::|||||||||||||
Db         59 TCTCTGATGACATGGAACATC 39

RESULT 41
US-11-130-645B-71304
; Sequence 71304, Application US/11130645B
; GENERAL INFORMATION:
;  APPLICANT: ROSETTA GENOMICS LTD
;  TITLE OF INVENTION: MICRORNAS AND USES THEREOF
;  FILE REFERENCE: 06087.0202.CPUS13
;  CURRENT APPLICATION NUMBER: US/11/130,645B
;  CURRENT FILING DATE:  2005-05-16
;  NUMBER OF SEQ ID NOS: 760616
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 71304
;   LENGTH: 63
;   TYPE: RNA
;   ORGANISM: hsa
US-11-130-645B-71304

Alignment Scores:
Pred. No.:              171             Length:         63
Score:                  35.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-130-645B-71304 (1-63)

Qy          1 SerLeuIleThrTrpAsnIle 7
                ||||||:::|||||||||:::
Db         26 UCUCUUGUCACAUGGAAUCUA 46

RESULT 42
US-11-130-645B-52383
; Sequence 52383, Application US/11130645B
; GENERAL INFORMATION:
;  APPLICANT: ROSETTA GENOMICS LTD
;  TITLE OF INVENTION: MICRORNAS AND USES THEREOF
;  FILE REFERENCE: 06087.0202.CPUS13
;  CURRENT APPLICATION NUMBER: US/11/130,645B
;  CURRENT FILING DATE:  2005-05-16
;  NUMBER OF SEQ ID NOS: 760616
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 52383
;   LENGTH: 64
;   TYPE: RNA
;   ORGANISM: hsa
US-11-130-645B-52383

Alignment Scores:
Pred. No.:              174             Length:         64
Score:                  35.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-130-645B-52383 (1-64)

Qy          1 SerLeuIleThrTrpAsnIle 7
                ||||||:::|||||||||:::
Db         27 UCUCUUGUCACAUGGAAUCUA 47
```

RESULT 43
US-11-130-645B-332515
; Sequence 332515, Application US/11130645B
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: MICRORNAS AND USES THEREOF
; FILE REFERENCE: 06087.0202.CPUS13
; CURRENT APPLICATION NUMBER: US/11/130,645B
; CURRENT FILING DATE: 2005-05-16
; NUMBER OF SEQ ID NOS: 760616
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 332515
; LENGTH: 64
; TYPE: RNA
; ORGANISM: hsa
US-11-130-645B-332515

Alignment Scores:

Pred. No.:	174	Length:	64
Score:	35.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-130-645B-332515 (1-64)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db         27 UCUCUUGUCACAUGGAAUCUA 47
```

RESULT 44
US-11-330-082-1351/c
; Sequence 1351, Application US/11330082
; GENERAL INFORMATION:
; APPLICANT: Buehler, Robert E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Coombs, Brian E.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.356 - 38-21(15760)C/US
; CURRENT APPLICATION NUMBER: US/11/330,082
; CURRENT FILING DATE: 2006-01-12
; PRIOR APPLICATION NUMBER: US 09/552,086
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 27278
; SEQ ID NO 1351
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: 700996434H1
US-11-330-082-1351

Alignment Scores:

Pred. No.:	813	Length:	286
Score:	35.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-330-082-1351 (1-286)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||::::::|||||||||||||
Db         70 TCAATTGTAACCTGGAACATC 50
```

RESULT 45
US-11-266-748A-430212
; Sequence 430212, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 430212
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-430212

Alignment Scores:

Pred. No.:	1.01e+03	Length:	354
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-430212 (1-354)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db        179 GCCTTAATCACTTGGAATATT 199
```

RESULT 46
US-11-266-748A-219748
; Sequence 219748, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276

```
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 219748
;   LENGTH: 536
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-219748

Alignment Scores:
Pred. No.:                1.55e+03        Length:         536
Score:                    35.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-219748 (1-536)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||
Db        270 TCTCTGATGACATGGAACATC 290

RESULT 47
US-11-266-748A-228613
; Sequence 228613, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 228613
;   LENGTH: 1620
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-228613

Alignment Scores:
Pred. No.:                4.86e+03        Length:         1620
Score:                    35.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-228613 (1-1620)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db        778 TCTCTTGTCACATGGAATCTA 798

RESULT 48
US-60-751-455-2154
; Sequence 2154, Application US/60751455
; GENERAL INFORMATION:
;  APPLICANT: Radich, Jerald P.
;  APPLICANT:  Dai, Hongyue
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Schelter, Janell
;  APPLICANT:  Linsley, Peter S.
;  TITLE OF INVENTION: GENES ASSOCIATED WITH PROGRESSION AND RESPONSE IN
;  TITLE OF INVENTION:  CHRONIC MYELOID LEUKEMIA AND USES THEREOF
;  FILE REFERENCE: 9301-253-888
;  CURRENT APPLICATION NUMBER: US/60/751,455
;  CURRENT FILING DATE:  2005-12-15
;  NUMBER OF SEQ ID NOS: 7936
;  SOFTWARE: FastSEQ 4.01
; SEQ ID NO 2154
;   LENGTH: 1831
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-751-455-2154

Alignment Scores:
Pred. No.:                5.51e+03        Length:         1831
Score:                    35.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       11              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-751-455-2154 (1-1831)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db        692 TCTCTTGTCACATGGAATCTA 712

RESULT 49
US-11-266-748A-162764
; Sequence 162764, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 162764
;   LENGTH: 1865
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-162764
```

```
Alignment Scores:
Pred. No.:              5.61e+03        Length:         1865
Score:                  35.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-162764 (1-1865)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db        791 TCTCTTGTCACATGGAATCTA 811

RESULT 50
US-11-266-748A-19824
; Sequence 19824, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19824
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-19824

Alignment Scores:
Pred. No.:              5.7e+03         Length:         1894
Score:                  35.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-19824 (1-1894)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db        755 TCTCTTGTCACATGGAATCTA 775

RESULT 51
US-11-266-748A-253327
; Sequence 253327, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 253327
;   LENGTH: 1907
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-253327

Alignment Scores:
Pred. No.:              5.74e+03        Length:         1907
Score:                  35.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-253327 (1-1907)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db        763 TCTCTTGTCACATGGAATCTA 783

RESULT 52
US-11-266-748A-313844/c
; Sequence 313844, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 313844
```

```
;   LENGTH: 1907
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-313844

Alignment Scores:
Pred. No.:              5.74e+03        Length:         1907
Score:                  35.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-313844 (1-1907)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||:::
Db       1145 TCTCTTGTCACATGGAATCTA 1125

RESULT 53
US-60-762-056-16200/c
; Sequence 16200, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 16200
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-16200

Alignment Scores:
Pred. No.:              6.97e+03        Length:         2300
Score:                  35.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  85.7%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-762-056-16200 (1-2300)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||||||
Db        103 TCTCTGATGACATGGAATATA 83

RESULT 54
US-60-762-056-23584/c
; Sequence 23584, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 23584
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-23584

Alignment Scores:
Pred. No.:              6.97e+03        Length:         2300
Score:                  35.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-762-056-23584 (1-2300)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db        822 TCCTTGGTCACCTGGAATTTA 802

RESULT 55
US-10-276-817B-1699
; Sequence 1699, Application US/10276817B
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 21272-104
;  CURRENT APPLICATION NUMBER: US/10/276,817B
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: PCT/US01/14827
;  PRIOR FILING DATE: 2001-05-16
;  PRIOR APPLICATION NUMBER: 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 16102
;  SOFTWARE: Custom
; SEQ ID NO 1699
;   LENGTH: 3378
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1533)..(1886)
;   OTHER INFORMATION: 47% homologous to Homo sapiens Similar to dCMP
;   OTHER INFORMATION: deaminase,accession number BC001286,Smith-Waterman Score=176.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(3378)
;   OTHER INFORMATION: n = a,t,c or g
US-10-276-817B-1699

Alignment Scores:
Pred. No.:              1.04e+04        Length:         3378
Score:                  35.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  71.4%           Mismatches:     0
Query Match:            92.1%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-276-817B-1699 (1-3378)

Qy          1 SerLeuIleThrTrpAsnIle 7
              ||||||:::|||||||||:::
Db       2047 TCTCTTGTCACATGGAATCTA 2067

RESULT 56
US-11-266-748A-56674/c
; Sequence 56674, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
```

```
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56674
;   LENGTH: 6051
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-56674

Alignment Scores:
Pred. No.:                1.89e+04        Length:         6051
Score:                    35.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-56674 (1-6051)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       2481 TCTCTGATGACATGGAACATC 2461

RESULT 57
US-60-742-219-1999/c
; Sequence 1999, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1999
;   LENGTH: 6051
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-1999

Alignment Scores:
Pred. No.:                1.89e+04        Length:         6051
Score:                    35.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       11              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-742-219-1999 (1-6051)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||:::|||||||||||||
Db       2481 TCTCTGATGACATGGAACATC 2461

RESULT 58
US-11-266-748A-32092/c
; Sequence 32092, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32092
;   LENGTH: 246282
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-32092

Alignment Scores:
Pred. No.:                8.24e+05        Length:         246282
Score:                    35.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              92.1%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-32092 (1-246282)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||||||||||||||||||
Db     127929 GCACTCATAACATGGAATATT 127909

RESULT 59
US-10-540-898-114
; Sequence 114, Application US/10540898
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: CHIR0056-101 (PP023367.0003)
;  CURRENT APPLICATION NUMBER: US/10/540,898
;  CURRENT FILING DATE:  2005-06-27
;  PRIOR APPLICATION NUMBER: US 10/330,773
;  PRIOR FILING DATE: 2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114
;   LENGTH: 440684
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
```

; NAME/KEY: misc_feature
; LOCATION: (1)...(440684)
; OTHER INFORMATION: n = A,T,C or G
US-10-540-898-114

Alignment Scores:

Pred. No.:	1.45e+06	Length:	440684
Score:	35.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	92.1%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-540-898-114 (1-440684)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::|||||||||||||:::
Db     395255 TCAATAATTACATGGAATGTT 395275
```

RESULT 60
US-11-313-452-984899
; Sequence 984899, Application US/11313452
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/313,452
; CURRENT FILING DATE: 2005-12-21
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 984899
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-313-452-984899

Alignment Scores:

Pred. No.:	77.8	Length:	19
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-313-452-984899 (1-19)

```
Qy          2 LeuIleThrTrpAsnIle 7
              |||||||||||||||||
Db          1 CUCAUUACCUGGAAUAUU 18
```

RESULT 61
US-11-227-183A-10563/c
; Sequence 10563, Application US/11227183A
; GENERAL INFORMATION:
; APPLICANT: Fisher, Dane K.
; APPLICANT: Lalgudi, Raghunath V.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.336 - 38-21(51230)C/US
; CURRENT APPLICATION NUMBER: US/11/227,183A
; CURRENT FILING DATE: 2005-09-16
; PRIOR APPLICATION NUMBER: US 09/619,643
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 32236
; SEQ ID NO 10563
; LENGTH: 340
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3150-056-Q1-N1-G8
US-11-227-183A-10563

Alignment Scores:

Pred. No.:	1.52e+03	Length:	340
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	71.4%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-227-183A-10563 (1-340)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||||
Db        157 GCCCTTGTGACTTGGAACATT 137
```

RESULT 62
US-11-313-816-11722/c
; Sequence 11722, Application US/11313816
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Ruan, Yijun G.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.353 - 38-21(15404)C/US
; CURRENT APPLICATION NUMBER: US/11/313,816
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: 09/333,534
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 18718
; SEQ ID NO 11722
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB25-101-Q1-E1-F3
US-11-313-816-11722

Alignment Scores:

Pred. No.:	1.58e+03	Length:	354
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-313-816-11722 (1-354)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||:::|||||||||
Db        268 TCTCTCATCTCGTGGAACATA 248
```

RESULT 63
US-11-329-388-2343
; Sequence 2343, Application US/11329388
; GENERAL INFORMATION:
; APPLICANT: Andersen, Scott E.
; APPLICANT: Fisher, Dane K.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.357 - 38-21(15503)C/US
; CURRENT APPLICATION NUMBER: US/11/329,388
; CURRENT FILING DATE: 2006-01-11
; PRIOR APPLICATION NUMBER: 09/553,094
; PRIOR FILING DATE: 2000-04-18

```
;  NUMBER OF SEQ ID NOS: 18831
; SEQ ID NO 2343
;   LENGTH: 390
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3061-037-Q1-K1-G1
US-11-329-388-2343

Alignment Scores:
Pred. No.:                1.75e+03        Length:         390
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-329-388-2343 (1-390)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||||
Db        358 GCACTTGTGACCTGGAACATT 378

RESULT 64
US-11-313-816-11719
; Sequence 11719, Application US/11313816
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Ruan, Yijun G.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.353 - 38-21(15404)C/US
;  CURRENT APPLICATION NUMBER: US/11/313,816
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: 09/333,534
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 18718
; SEQ ID NO 11719
;   LENGTH: 424
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB25-101-Q1-E1-F1
US-11-313-816-11719

Alignment Scores:
Pred. No.:                1.91e+03        Length:         424
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       9               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-313-816-11719 (1-424)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||:::::|||||||||||||
Db         96 TCAGTGGTGACGTGGAACATC 116

RESULT 65
US-11-353-150-66475
; Sequence 66475, Application US/11353150
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Scott E.
;  APPLICANT:  Conner, Timothy W.
;  APPLICANT:  Lalgudi, Raghunath V.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.365 - 38-21(15459)D/US
;  CURRENT APPLICATION NUMBER: US/11/353,150
;  CURRENT FILING DATE:  2006-02-14
;  PRIOR APPLICATION NUMBER: 09/565,386
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/132,860
;  PRIOR FILING DATE: 1999-05-07
;  NUMBER OF SEQ ID NOS: 83523
; SEQ ID NO 66475
;   LENGTH: 428
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3076-032-Q1-K1-A1
US-11-353-150-66475

Alignment Scores:
Pred. No.:                1.93e+03        Length:         428
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-353-150-66475 (1-428)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||||
Db        377 GCACTTGTGACCTGGAACATC 397

RESULT 66
US-11-329-388-2953
; Sequence 2953, Application US/11329388
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Scott E.
;  APPLICANT:  Fisher, Dane K.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.357 - 38-21(15503)C/US
;  CURRENT APPLICATION NUMBER: US/11/329,388
;  CURRENT FILING DATE:  2006-01-11
;  PRIOR APPLICATION NUMBER: 09/553,094
;  PRIOR FILING DATE: 2000-04-18
;  NUMBER OF SEQ ID NOS: 18831
; SEQ ID NO 2953
;   LENGTH: 429
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(429)
;   OTHER INFORMATION: unsure at all n locations
;   OTHER INFORMATION: Clone ID: LIB3061-015-Q1-K1-D8
US-11-329-388-2953

Alignment Scores:
Pred. No.:                1.93e+03        Length:         429
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    71.4%           Mismatches:     0
Query Match:              89.5%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-329-388-2953 (1-429)

Qy          1 SerLeuIleThrTrpAsnIle 7
              :::|||:::|||||||||||||
Db         31 GCACTTGTGACCTGGAACATC 51

RESULT 67
US-11-266-748A-266797/c
; Sequence 266797, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
```

; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 266797
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (83)..(186)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-266797

Alignment Scores:

Pred. No.:	2.03e+03	Length:	451
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-266797 (1-451)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||:::|||||||||
Db        367 AGCCTCATCTCCTGGAATATT 347
```

RESULT 68
US-11-266-748A-280499/c
; Sequence 280499, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 280499
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (83)..(186)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-280499

Alignment Scores:

Pred. No.:	2.03e+03	Length:	451
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-280499 (1-451)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||:::|||||||||
Db        367 AGCCTCATCTCCTGGAATATT 347
```

RESULT 69
US-11-266-748A-327314
; Sequence 327314, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 327314
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (266)..(369)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-327314

Alignment Scores:

Pred. No.:	2.03e+03	Length:	451
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0

```
Query Match:            89.5%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-327314 (1-451)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||:::|||||||||
Db         85 AGCCTCATCTCCTGGAATATT 105

RESULT 70
US-11-311-886-3178
; Sequence 3178, Application US/11311886
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
;  TITLE OF INVENTION:  PLANTS
;  FILE REFERENCE: 16517.358 - 38-21(15598)C/US
;  CURRENT APPLICATION NUMBER: US/11/311,886
;  CURRENT FILING DATE:  2005-12-20
;  PRIOR APPLICATION NUMBER: 09/421,106
;  PRIOR FILING DATE: 1999-10-15
;  NUMBER OF SEQ ID NOS: 36938
; SEQ ID NO 3178
;   LENGTH: 459
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
US-11-311-886-3178

Alignment Scores:
Pred. No.:              2.07e+03       Length:        459
Score:                  34.00          Matches:       6
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            89.5%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-311-886-3178 (1-459)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db        194 AGTCTAATAACATGGAAC 211

RESULT 71
US-11-313-816-7526/c
; Sequence 7526, Application US/11313816
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Ruan, Yijun G.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.353 - 38-21(15404)C/US
;  CURRENT APPLICATION NUMBER: US/11/313,816
;  CURRENT FILING DATE:  2005-12-22
;  PRIOR APPLICATION NUMBER: 09/333,534
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 18718
; SEQ ID NO 7526
;   LENGTH: 468
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB24-074-Q1-E1-G5
US-11-313-816-7526

Alignment Scores:
Pred. No.:              2.11e+03       Length:        468
Score:                  34.00          Matches:       6
Percent Similarity:     100.0%         Conservative:  1
Best Local Similarity:  85.7%          Mismatches:    0
Query Match:            89.5%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-313-816-7526 (1-468)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||:::|||||||||
Db        241 TCTCTCATCTCGTGGAACATA 221

RESULT 72
US-11-266-748A-480077
; Sequence 480077, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 480077
;   LENGTH: 581
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44)..(44)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (46)..(46)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-480077

Alignment Scores:
Pred. No.:              2.64e+03       Length:        581
Score:                  34.00          Matches:       6
Percent Similarity:     100.0%         Conservative:  0
Best Local Similarity:  100.0%         Mismatches:    0
Query Match:            89.5%          Indels:        0
DB:                     9              Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-480077 (1-581)

Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db        506 TCACTAATAACTTGGAAT 523

RESULT 73
US-11-266-748A-232881
; Sequence 232881, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
```

; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 232881
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (169)..(211)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-232881

Alignment Scores:

Pred. No.:	3.01e+03	Length:	660
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-232881 (1-660)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db        600 TCTCTCATCACTTGGAAT 617
```

RESULT 74
US-11-266-748A-381213
; Sequence 381213, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 381213
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50)..(69)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-381213

Alignment Scores:

Pred. No.:	3.86e+03	Length:	841
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	89.5%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-381213 (1-841)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db        563 TCACTAATAACTTGGAAT 580
```

RESULT 75
US-11-266-748A-430969/c
; Sequence 430969, Application US/11266748A
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 430969
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (773)..(792)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-430969

Alignment Scores:

Pred. No.:	3.86e+03	Length:	841
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	0

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 89.5% Indels: 0
DB: 9 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x US-11-266-748A-430969 (1-841)

```
Qy          1 SerLeuIleThrTrpAsn 6
              ||||||||||||||||||
Db        279 TCACTAATAACTTGGAAT 262
```

Search completed: March 10, 2006, 23:59:20
Job time : 32.3777 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:19:50 ; Search time 287.713 Seconds
(without alignments)
1345.232 Million cell updates/sec

Title: US-10-757-745-2_COPY_115_121
Perfect score: 38
Sequence: 1 SLITWNI 7

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170333_4165/app_query.fasta_1
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=rnpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss02h -USER=US10757745_@CGN_1_1_20983_@runat_09032006_170333_4165
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUSB_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUSC_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US075_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US076_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US077_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US078_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US079_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US101_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US104_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US105_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US112_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US117_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US600_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US601_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US603_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US605_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US606_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US607_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	164	25	US-09-504-576A-614	Sequence 614, App
2	38	100.0	164	35	US-09-912-293-47580	Sequence 47580, A
3	38	100.0	247	21	US-09-105-427-3620	Sequence 3620, Ap
4	38	100.0	248	20	US-09-015-080-208	Sequence 208, App
5	38	100.0	248	26	US-09-540-212A-31516	Sequence 31516, A
6	38	100.0	264	19	US-08-918-182-1407	Sequence 1407, Ap
7	38	100.0	264	27	US-09-540-766-17346	Sequence 17346, A
8	38	100.0	264	74	US-60-024-832-1407	Sequence 1407, Ap

	9	38	100.0	290	20	US-09-036-591-1392 Sequence 1392, Ap
	10	38	100.0	290	27	US-09-540-766-40150 Sequence 40150, A
	11	38	100.0	290	74	US-60-036-549-1392 Sequence 1392, Ap
	12	38	100.0	306	26	US-09-540-764-53395 Sequence 53395, A
	13	38	100.0	306	49	US-10-349-781-53395 Sequence 53395, A
	14	38	100.0	326	23	US-09-321-214-23541 Sequence 23541, A
	15	38	100.0	326	25	US-09-516-335-23541 Sequence 23541, A
	16	38	100.0	326	31	US-09-733-811-23541 Sequence 23541, A
	17	38	100.0	326	31	US-09-733-811A-23541 Sequence 23541, A
	18	38	100.0	326	38	US-09-975-640-23541 Sequence 23541, A
	19	38	100.0	326	38	US-09-975-640A-23541 Sequence 23541, A
	20	38	100.0	351	25	US-09-515-128-10314 Sequence 10314, A
	21	38	100.0	351	31	US-09-721-544-10314 Sequence 10314, A
	22	38	100.0	356	23	US-09-321-214-1021 Sequence 1021, Ap
	23	38	100.0	356	25	US-09-516-335-1021 Sequence 1021, Ap
	24	38	100.0	356	31	US-09-733-811-1021 Sequence 1021, Ap
	25	38	100.0	356	31	US-09-733-811A-1021 Sequence 1021, Ap
	26	38	100.0	356	38	US-09-975-640-1021 Sequence 1021, Ap
	27	38	100.0	356	38	US-09-975-640A-1021 Sequence 1021, Ap
	28	38	100.0	372	25	US-09-515-128-17577 Sequence 17577, A
	29	38	100.0	372	31	US-09-721-544-17577 Sequence 17577, A
	30	38	100.0	384	29	US-09-644-868-2256 Sequence 2256, Ap
C	31	38	100.0	391	1	PCT-US02-02870-826 Sequence 826, App
C	32	38	100.0	391	40	US-10-066-543-826 Sequence 826, App
C	33	38	100.0	391	43	US-10-214-403-826 Sequence 826, App
	34	38	100.0	394	23	US-09-362-510-41529 Sequence 41529, A
	35	38	100.0	394	23	US-09-362-510A-41529 Sequence 41529, A
	36	38	100.0	394	35	US-09-904-013-41529 Sequence 41529, A
	37	38	100.0	396	25	US-09-528-409-24161 Sequence 24161, A
	38	38	100.0	396	36	US-09-933-524-24161 Sequence 24161, A
	39	38	100.0	396	36	US-09-933-524A-24161 Sequence 24161, A
	40	38	100.0	397	25	US-09-528-409-26306 Sequence 26306, A
	41	38	100.0	397	36	US-09-933-524-26306 Sequence 26306, A
	42	38	100.0	397	36	US-09-933-524A-26306 Sequence 26306, A
	43	38	100.0	398	24	US-09-489-036-16373 Sequence 16373, A
	44	38	100.0	398	37	US-09-943-143-16373 Sequence 16373, A
	45	38	100.0	403	24	US-09-489-036-16721 Sequence 16721, A
	46	38	100.0	403	37	US-09-943-143-16721 Sequence 16721, A
	47	38	100.0	409	24	US-09-489-036-16392 Sequence 16392, A
	48	38	100.0	409	37	US-09-943-143-16392 Sequence 16392, A
	49	38	100.0	414	24	US-09-404-549-2518 Sequence 2518, Ap
	50	38	100.0	414	24	US-09-404-549A-2518 Sequence 2518, Ap
	51	38	100.0	420	25	US-09-528-409-24160 Sequence 24160, A
	52	38	100.0	420	36	US-09-933-524-24160 Sequence 24160, A
	53	38	100.0	420	36	US-09-933-524A-24160 Sequence 24160, A
	54	38	100.0	430	24	US-09-489-036-26234 Sequence 26234, A
	55	38	100.0	430	37	US-09-943-143-26234 Sequence 26234, A
	56	38	100.0	433	25	US-09-528-409-74096 Sequence 74096, A
	57	38	100.0	433	36	US-09-933-524-74096 Sequence 74096, A
	58	38	100.0	433	36	US-09-933-524A-74096 Sequence 74096, A
	59	38	100.0	438	23	US-09-353-690-10783 Sequence 10783, A
	60	38	100.0	438	25	US-09-528-409-3652 Sequence 3652, Ap
	61	38	100.0	438	35	US-09-922-340-10783 Sequence 10783, A
	62	38	100.0	438	35	US-09-922-340A-10783 Sequence 10783, A
	63	38	100.0	438	36	US-09-933-524-3652 Sequence 3652, Ap
	64	38	100.0	438	36	US-09-933-524A-3652 Sequence 3652, Ap
	65	38	100.0	439	23	US-09-321-214-11519 Sequence 11519, A
	66	38	100.0	439	25	US-09-516-335-11519 Sequence 11519, A
	67	38	100.0	439	31	US-09-733-811-11519 Sequence 11519, A
	68	38	100.0	439	31	US-09-733-811A-11519 Sequence 11519, A
	69	38	100.0	439	38	US-09-975-640-11519 Sequence 11519, A
	70	38	100.0	439	38	US-09-975-640A-11519 Sequence 11519, A
	71	38	100.0	440	25	US-09-528-409-14287 Sequence 14287, A
	72	38	100.0	440	36	US-09-933-524-14287 Sequence 14287, A
	73	38	100.0	440	36	US-09-933-524A-14287 Sequence 14287, A
	74	38	100.0	449	23	US-09-304-649-421 Sequence 421, App
	75	38	100.0	449	33	US-09-840-424-421 Sequence 421, App

ALIGNMENTS

RESULT 1
US-09-504-576A-614

; Sequence 614, Application US/09504576A
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 48
; FILE REFERENCE: PO48
; CURRENT APPLICATION NUMBER: US/09/504,576A
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/120,430
; PRIOR FILING DATE: 1999-02-16
; NUMBER OF SEQ ID NOS: 14367
; SEQ ID NO 614
; LENGTH: 164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (35)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (40)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (47)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (93)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (113)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (123)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (150)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (161)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-504-576A-614

Alignment Scores:

Pred. No.:	3.81e+03	Length:	164
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	25	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-504-576A-614 (1-164)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         48 TCTCTGATAACTTGGAACATC 68
```

RESULT 2
US-09-912-293-47580
; Sequence 47580, Application US/09912293
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 100
; FILE REFERENCE: PO-100
; CURRENT APPLICATION NUMBER: US/09/912,293
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 08/103,744
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 09/249,651
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 08/104,507
; PRIOR FILING DATE: 1993-08-09

```
;  PRIOR APPLICATION NUMBER: 08/196,363
;  PRIOR FILING DATE: 1994-02-15
;  PRIOR APPLICATION NUMBER: 09/859,490
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: 08/196,362
;  PRIOR FILING DATE: 1994-02-15
;  PRIOR APPLICATION NUMBER: 08/221,623
;  PRIOR FILING DATE: 1994-03-31
;  PRIOR APPLICATION NUMBER: 08/220,691
;  PRIOR FILING DATE: 1994-03-31
;  PRIOR APPLICATION NUMBER: 09/741,830
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 09/813,155
;  PRIOR FILING DATE: 2001-03-21
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 244538
; SEQ ID NO 47580
;   LENGTH: 164
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (35)..(35)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (40)..(40)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (47)..(47)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (93)..(93)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (113)..(113)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (123)..(123)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (150)..(150)
;   OTHER INFORMATION: n is equal to a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(161)
;   OTHER INFORMATION: n is equal to a,t,g, or c
US-09-912-293-47580

Alignment Scores:
Pred. No.:                3.81e+03        Length:          164
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       35              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-912-293-47580 (1-164)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         48 TCTCTGATAACTTGGAACATC 68

RESULT 3
US-09-105-427-3620
; Sequence 3620, Application US/09105427
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Naughton, Rebecca E.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  HUMAN PROSTATE WITH BENIGN PROSTATIC HYPERPLASIA
;    NUMBER OF SEQUENCES:  3842
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/105,427
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PD-0390P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 845-4166
;  INFORMATION FOR SEQ ID NO:  3620:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  247 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  3320501H1
US-09-105-427-3620

Alignment Scores:
Pred. No.:                5.99e+03        Length:          247
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       21              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-105-427-3620 (1-247)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        142 TCTCTCATTACCTGGAATATT 162

RESULT 4
US-09-015-080-208
; Sequence 208, Application US/09015080
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuve, Laura L.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Naughton, Rebecca E.
;    APPLICANT:  Klingler, Tod M.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  LUNG WITH SQUAMOUS CELL CARCINOMA
;    NUMBER OF SEQUENCES:  4133
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/015,080
;      FILING DATE:  Herewith
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,039
;      FILING DATE:  January 28, 1997
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PD-0310 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  208:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  248 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  2603934H1
US-09-015-080-208

Alignment Scores:
Pred. No.:                6.02e+03        Length:         248
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       20              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-015-080-208 (1-248)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         11 TCTCTCATTACCTGGAATATT 31

RESULT 5
US-09-540-212A-31516
; Sequence 31516, Application US/09540212A
; GENERAL INFORMATION:
;  APPLICANT: Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION: POLYNUCLEOTIDES OF AIRWAY AND LUNG SYSTEM TISSUE
;  FILE REFERENCE: PD-1034 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,212A
;  CURRENT FILING DATE:  2000-03-31
;  NUMBER OF SEQ ID NOS: 67551
;  SOFTWARE: PERL Program
; SEQ ID NO 31516
;   LENGTH: 248
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00625939
;   NAME/KEY: unsure
;   LOCATION: 61
;   OTHER INFORMATION: a, t, c, g, or other
US-09-540-212A-31516

Alignment Scores:
Pred. No.:                6.02e+03        Length:         248
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       26              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-540-212A-31516 (1-248)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         11 TCTCTCATTACCTGGAATATT 31

RESULT 6
US-08-918-182-1407
; Sequence 1407, Application US/08918182
;  GENERAL INFORMATION:
;    APPLICANT:  Gooding, Douglas H.
;    APPLICANT:  Stuve, Laura L.
;    APPLICANT:  Stuart, Susan G.
;    APPLICANT:  Ito, Laura Y.
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Delegeane, Angelo M.
;    APPLICANT:  Naughton, Rebecca E.
;    APPLICANT:  Klingler, Tod M.
;    TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;    TITLE OF INVENTION:  HUMAN SMALL INTESTINE
;    NUMBER OF SEQUENCES:  1686
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/918,182
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/024,832
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C., Ph.D.
;      REGISTRATION NUMBER:  39132
;      REFERENCE/DOCKET NUMBER:  PD-0234P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 845-4166
;  INFORMATION FOR SEQ ID NO:  1407:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  264 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  1807687
US-08-918-182-1407
```

```
Alignment Scores:
Pred. No.:                6.45e+03        Length:         264
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       19              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-08-918-182-1407 (1-264)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        197 TCTCTCATTACCTGGAATATT 217

RESULT 7
US-09-540-766-17346
; Sequence 17346, Application US/09540766
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE
;  FILE REFERENCE:  PD-1024 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,766
;  CURRENT FILING DATE:  2000-03-30
;  PRIOR APPLICATION DATA removed - refer tp PALM or File Wrapper
;  NUMBER OF SEQ ID NOS:  77960
;  SOFTWARE:  PERL Program
; SEQ ID NO 17346
;   LENGTH: 264
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00289414
;   NAME/KEY: unsure
;   LOCATION: 17
;   OTHER INFORMATION: a, t, c, g, or other
US-09-540-766-17346

Alignment Scores:
Pred. No.:                6.45e+03        Length:         264
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       27              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-540-766-17346 (1-264)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        197 TCTCTCATTACCTGGAATATT 217

RESULT 8
US-60-024-832-1407
; Sequence 1407, Application US/60024832
;  GENERAL INFORMATION:
;     APPLICANT:  Gooding, Douglas H.
;     APPLICANT:  Stuve, Laura L.
;     APPLICANT:  Stuart, Susan G.
;     APPLICANT:  Ito, Laura Y.
;     APPLICANT:  Akerblom, Ingrid E.
;     APPLICANT:  Delegeane, Angelo M.
;     APPLICANT:  Naughton, Rebecca E.
;     APPLICANT:  Klingler, Tod M.
;     TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;     TITLE OF INVENTION:  HUMAN SMALL INTESTINE
;     NUMBER OF SEQUENCES:  1686
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;       STREET:  3174 PORTER DRIVE
;       CITY:  PALO ALTO
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/60/024,832
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CERRONE, MICHAEL C., Ph.D.
;       REGISTRATION NUMBER:  39132
;       REFERENCE/DOCKET NUMBER:  PD-0234P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 855-0555
;       TELEFAX:  (415) 845-4166
;   INFORMATION FOR SEQ ID NO:  1407:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  264 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     IMMEDIATE SOURCE:
;       CLONE:  1807687
US-60-024-832-1407

Alignment Scores:
Pred. No.:                6.45e+03        Length:         264
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       74              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-024-832-1407 (1-264)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        197 TCTCTCATTACCTGGAATATT 217

RESULT 9
US-09-036-591-1392
; Sequence 1392, Application US/09036591
;  GENERAL INFORMATION:
;     APPLICANT:  Gooding, Douglas H.
;     APPLICANT:  Stuve, Laura L.
;     APPLICANT:  Stuart, Susan G.
;     APPLICANT:  Ito, Laura Y.
;     APPLICANT:  Akerblom, Ingrid E.
;     APPLICANT:  Delegeane, Angelo M.
;     APPLICANT:  Naughton, Rebecca E.
;     APPLICANT:  Klingler, Tod M.
;     TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;     TITLE OF INVENTION:  COLON TUMOR
;     NUMBER OF SEQUENCES:  2385
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;       STREET:  3174 PORTER DRIVE
;       CITY:  PALO ALTO
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/036,591
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/036,549
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CERRONE, MICHAEL C.
;       REGISTRATION NUMBER:  39,132
;       REFERENCE/DOCKET NUMBER:  PD-0322P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 855-0555
;       TELEFAX:  (415) 845-4166
;   INFORMATION FOR SEQ ID NO:  1392:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  290 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     IMMEDIATE SOURCE:
;       CLONE:  2792292H1
US-09-036-591-1392

Alignment Scores:
Pred. No.:                7.16e+03        Length:         290
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       20              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-036-591-1392 (1-290)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        255 TCTCTCATTACCTGGAATATT 275

RESULT 10
US-09-540-766-40150
; Sequence 40150, Application US/09540766
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF GASTROINTESTINAL SYSTEM TISSUE
;  FILE REFERENCE:  PD-1024 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,766
;  CURRENT FILING DATE:  2000-03-30
;  PRIOR APPLICATION DATA removed - refer tp PALM or File Wrapper
;  NUMBER OF SEQ ID NOS:  77960
;  SOFTWARE:  PERL Program
; SEQ ID NO 40150
;   LENGTH: 290
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu00856020
;   NAME/KEY: unsure
;   LOCATION: 23
;   OTHER INFORMATION: a, t, c, g, or other
US-09-540-766-40150

Alignment Scores:
Pred. No.:                7.16e+03        Length:         290
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       27              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-540-766-40150 (1-290)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        255 TCTCTCATTACCTGGAATATT 275

RESULT 11
US-60-036-549-1392
; Sequence 1392, Application US/60036549
;  GENERAL INFORMATION:
;     APPLICANT:  Gooding, Douglas H.
;     APPLICANT:  Stuve, Laura L.
;     APPLICANT:  Stuart, Susan G.
;     APPLICANT:  Ito, Laura Y.
;     APPLICANT:  Akerblom, Ingrid E.
;     APPLICANT:  Delegeane, Angelo M.
;     APPLICANT:  Naughton, Rebecca E.
;     APPLICANT:  Klingler, Tod M.
;     TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
;     TITLE OF INVENTION:  COLON TUMOR
;     NUMBER OF SEQUENCES:  2385
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;       STREET:  3174 PORTER DRIVE
;       CITY:  PALO ALTO
;       STATE:  CALIFORNIA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/60/036,549
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CERRONE, MICHAEL C.
;       REGISTRATION NUMBER:  39,132
;       REFERENCE/DOCKET NUMBER:  PD-0322P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 855-0555
;       TELEFAX:  (415) 845-4166
;   INFORMATION FOR SEQ ID NO:  1392:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  290 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     IMMEDIATE SOURCE:
;       CLONE:  2792292H1
US-60-036-549-1392

Alignment Scores:
Pred. No.:                7.16e+03        Length:         290
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       74              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-60-036-549-1392 (1-290)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        255 TCTCTCATTACCTGGAATATT 275
```

```
RESULT 12
US-09-540-764-53395
; Sequence 53395, Application US/09540764
; GENERAL INFORMATION:
;  APPLICANT:  Seilhamer, Jeffrey J.
;  APPLICANT:  Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.
;  APPLICANT:  Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.
;  APPLICANT:  Naughton, Rebecca E.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES OF CELLS AND TISSUES OF THE FEMALE REPRODUCTIVE
;  FILE REFERENCE:  PD-1028 CIP
;  CURRENT APPLICATION NUMBER: US/09/540,764
;  CURRENT FILING DATE:  2000-03-30
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS:  61458
;  SOFTWARE:  PERL Program
; SEQ ID NO 53395
;   LENGTH: 306
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu01021970
US-09-540-764-53395

Alignment Scores:
Pred. No.:                7.59e+03        Length:         306
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       26              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-540-764-53395 (1-306)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        213 TCTCTCATTACCTGGAATATT 233

RESULT 13
US-10-349-781-53395
; Sequence 53395, Application US/10349781
; GENERAL INFORMATION:
;  APPLICANT: Seilhamer, Jeffrey J.; Delegeane, Angelo M.
;  APPLICANT:  Stuart, Susan G.; Stuve, Laura L.
;  APPLICANT:  Mullahy, Sara J.; Naughton, Rebecca E.
;  TITLE OF INVENTION: POLYNUCLEOTIDES OF CELLS AND TISSUES OF THE FEMALE REPRODUCTIVE S
;  FILE REFERENCE: PD-1028-2 CON
;  CURRENT APPLICATION NUMBER: US/10/349,781
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: 09/540,764
;  PRIOR FILING DATE: 2000-03-30
;  PRIOR APPLICATION NUMBER: 08/992,868
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/032,838
;  PRIOR FILING DATE: 1996-12-13
;  PRIOR APPLICATION NUMBER: 08/734,050
;  PRIOR FILING DATE: 1996-10-18
;  PRIOR APPLICATION NUMBER: 60/006,111
;  PRIOR FILING DATE: 1995-10-24
;  PRIOR APPLICATION NUMBER: 08/733,814
;  PRIOR FILING DATE: 1996-10-18
;  PRIOR APPLICATION NUMBER: 60/005,864
;  PRIOR FILING DATE: 1995-10-26
;  PRIOR APPLICATION NUMBER: 08/763,920
;  PRIOR FILING DATE: 1996-12-11
;  PRIOR APPLICATION NUMBER: 60/008,794
;  PRIOR FILING DATE: 1995-12-14
;  PRIOR APPLICATION NUMBER: 08/772,783
;  PRIOR FILING DATE: 1996-12-23
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 61458
;  SOFTWARE: PERL Program
; SEQ ID NO 53395
;   LENGTH: 306
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: hu01021970
US-10-349-781-53395

Alignment Scores:
Pred. No.:                7.59e+03        Length:         306
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       49              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-349-781-53395 (1-306)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        213 TCTCTCATTACCTGGAATATT 233

RESULT 14
US-09-321-214-23541
; Sequence 23541, Application US/09321214
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  APPLICANT: Drmanac, Radoje
;  APPLICANT: Crkvenjakov, Radomir
;  APPLICANT: Dickson, Mark
;  APPLICANT: Drmanac, Snezana
;  APPLICANT: Labat, Ivan
;  APPLICANT: Leshkowitz, Dena
;  APPLICANT: Kita, David
;  APPLICANT: Ford, John
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-740
;  CURRENT APPLICATION NUMBER: US/09/321,214
;  CURRENT FILING DATE: 1999-05-26
;  EARLIER APPLICATION NUMBER: US 60/088,041
;  EARLIER FILING DATE: 1998-06-02
;  NUMBER OF SEQ ID NOS: 31906
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23541
;   LENGTH: 326
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-321-214-23541

Alignment Scores:
Pred. No.:                8.14e+03        Length:         326
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       23              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-321-214-23541 (1-326)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        255 TCTCTCATTACCTGGAATATT 275

RESULT 15
US-09-516-335-23541
; Sequence 23541, Application US/09516335
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT: Asghari, Vida
;  APPLICANT: Damavandi, Simin
```

; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/516,335
; CURRENT FILING DATE: 2000-03-01
; EARLIER APPLICATION NUMBER: 09/321,214
; EARLIER FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: 60/088,041
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23541
; LENGTH: 326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-516-335-23541

Alignment Scores:

Pred. No.:	8.14e+03	Length:	326
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	25	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-516-335-23541 (1-326)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        255 TCTCTCATTACCTGGAATATT 275
```

RESULT 16
US-09-733-811-23541
; Sequence 23541, Application US/09733811
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/733,811
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/321,214
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23541
; LENGTH: 326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-733-811-23541

Alignment Scores:

Pred. No.:	8.14e+03	Length:	326
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	31	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-733-811-23541 (1-326)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        255 TCTCTCATTACCTGGAATATT 275
```

RESULT 17
US-09-733-811A-23541
; Sequence 23541, Application US/09733811A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Ford, John
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-740
; CURRENT APPLICATION NUMBER: US/09/733,811A
; CURRENT FILING DATE: 2000-12-08

; PRIOR APPLICATION NUMBER: 09/321,214
; PRIOR FILING DATE: 1999-05-26
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23541
; LENGTH: 326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-733-811A-23541

Alignment Scores:

Pred. No.:	8.14e+03	Length:	326
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	31	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-733-811A-23541 (1-326)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        255 TCTCTCATTACCTGGAATATT 275
```

RESULT 18
US-09-975-640-23541
; Sequence 23541, Application US/09975640
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/975,640
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/733,811
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23541
; LENGTH: 326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-975-640-23541

Alignment Scores:

Pred. No.:	8.14e+03	Length:	326
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	38	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-975-640-23541 (1-326)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        255 TCTCTCATTACCTGGAATATT 275
```

RESULT 19
US-09-975-640A-23541
; Sequence 23541, Application US/09975640A
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/975,640A
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/733,811
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23541
; LENGTH: 326

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-975-640A-23541

Alignment Scores:
Pred. No.:                8.14e+03        Length:         326
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       38              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-975-640A-23541 (1-326)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        255 TCTCTCATTACCTGGAATATT 275

RESULT 20
US-09-515-128-10314
; Sequence 10314, Application US/09515128
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT: Asghari, Vida
;  APPLICANT: Damavandi, Simin
;  APPLICANT: Dickson, Mark
;  APPLICANT: Drake, Jim
;  APPLICANT: Drmanac, Radoje
;  APPLICANT: Engleman, Carrie
;  APPLICANT: Faulkner, Brandy
;  APPLICANT: Garcia, Veronica
;  APPLICANT: Giedt, Gretchen
;  APPLICANT: Hunter, Kelly
;  APPLICANT: Jessen, Aaron
;  APPLICANT: Jones, Lee
;  APPLICANT: Kita, David
;  APPLICANT: Labat, Ivan
;  APPLICANT: Laroya, Mimi
;  APPLICANT: Lomelli, Michelle
;  APPLICANT: Nguyen, Phuong
;  APPLICANT: Nogra, Margie
;  APPLICANT: Palencia, Servando
;  APPLICANT: Raisi, Fariba
;  APPLICANT: Smith, Benjamin
;  APPLICANT: Tkach, Joe
;  APPLICANT: Tran, Lien
;  APPLICANT: Verna, Ron
;  APPLICANT: Yang, Fei
;  APPLICANT: Yim, Kenneth
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
;  FILE REFERENCE: 728CIP
;  CURRENT APPLICATION NUMBER: US/09/515,128
;  CURRENT FILING DATE: 2000-02-28
;  EARLIER APPLICATION NUMBER: 09/034,341
;  EARLIER FILING DATE: 1998-02-13
;  NUMBER OF SEQ ID NOS: 24489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10314
;   LENGTH: 351
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-515-128-10314

Alignment Scores:
Pred. No.:                8.83e+03        Length:         351
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       25              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-515-128-10314 (1-351)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        248 TCTCTCATTACCTGGAATATT 268

RESULT 21
US-09-721-544-10314
; Sequence 10314, Application US/09721544
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT: Asghari, Vida
;  APPLICANT: Damavandi, Simin
;  APPLICANT: Dickson, Mark
;  APPLICANT: Drake, Jim
;  APPLICANT: Drmanac, Radoje
;  APPLICANT: Engleman, Carrie
;  APPLICANT: Faulkner, Brandy
;  APPLICANT: Garcia, Veronica
;  APPLICANT: Giedt, Gretchen
;  APPLICANT: Hunter, Kelly
;  APPLICANT: Jessen, Aaron
;  APPLICANT: Jones, Lee
;  APPLICANT: Kita, David
;  APPLICANT: Labat, Ivan
;  APPLICANT: Laroya, Mimi
;  APPLICANT: Lomelli, Michelle
;  APPLICANT: Nguyen, Phuong
;  APPLICANT: Nogra, Margie
;  APPLICANT: Palencia, Servando
;  APPLICANT: Raisi, Fariba
;  APPLICANT: Smith, Benjamin
;  APPLICANT: Tkach, Joe
;  APPLICANT: Tran, Lien
;  APPLICANT: Verna, Ron
;  APPLICANT: Yang, Fei
;  APPLICANT: Yim, Kenneth
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
;  FILE REFERENCE: 728CIP
;  CURRENT APPLICATION NUMBER: US/09/721,544
;  CURRENT FILING DATE: 2000-11-21
;  PRIOR APPLICATION NUMBER: 09/515,128
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: 09/034,341
;  PRIOR FILING DATE: 1998-02-13
;  NUMBER OF SEQ ID NOS: 24489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10314
;   LENGTH: 351
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-721-544-10314

Alignment Scores:
Pred. No.:                8.83e+03        Length:         351
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       31              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-721-544-10314 (1-351)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        248 TCTCTCATTACCTGGAATATT 268

RESULT 22
US-09-321-214-1021
; Sequence 1021, Application US/09321214
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  APPLICANT: Drmanac, Radoje
```

; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Ford, John
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-740
; CURRENT APPLICATION NUMBER: US/09/321,214
; CURRENT FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: US 60/088,041
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1021
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-321-214-1021

Alignment Scores:

Pred. No.:	8.97e+03	Length:	356
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	23	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-321-214-1021 (1-356)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        287 TCTCTCATTACCTGGAATATT 307
```

RESULT 23
US-09-516-335-1021
; Sequence 1021, Application US/09516335
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/516,335
; CURRENT FILING DATE: 2000-03-01
; EARLIER APPLICATION NUMBER: 09/321,214
; EARLIER FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: 60/088,041
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1021
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-516-335-1021

Alignment Scores:

Pred. No.:	8.97e+03	Length:	356
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	25	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-516-335-1021 (1-356)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        287 TCTCTCATTACCTGGAATATT 307
```

RESULT 24
US-09-733-811-1021
; Sequence 1021, Application US/09733811
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/733,811
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/321,214

```
;  PRIOR FILING DATE:  1999-05-26
;  PRIOR APPLICATION NUMBER:  60/088,041
;  PRIOR FILING DATE:  1998-06-02
;  NUMBER OF SEQ ID NOS:  31906
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1021
;   LENGTH: 356
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-733-811-1021

Alignment Scores:
Pred. No.:              8.97e+03        Length:         356
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-733-811-1021 (1-356)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        287 TCTCTCATTACCTGGAATATT 307

RESULT 25
US-09-733-811A-1021
; Sequence 1021, Application US/09733811A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Crkvenjakov, Radomir
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drmanac, Snezana
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Kita, David
;  APPLICANT:  Ford, John
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-740
;  CURRENT APPLICATION NUMBER: US/09/733,811A
;  CURRENT FILING DATE:  2000-12-08
;  PRIOR APPLICATION NUMBER: 09/321,214
;  PRIOR FILING DATE: 1999-05-26
;  NUMBER OF SEQ ID NOS: 31906
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1021
;   LENGTH: 356
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-733-811A-1021

Alignment Scores:
Pred. No.:              8.97e+03        Length:         356
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-733-811A-1021 (1-356)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        287 TCTCTCATTACCTGGAATATT 307

RESULT 26
US-09-975-640-1021
; Sequence 1021, Application US/09975640
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT:  Asghari, Vida
;  APPLICANT:  Damavandi, Simin
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drake, Jim
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Engleman, Carrie
;  APPLICANT:  Faulkner, Brandy
;  APPLICANT:  Fox, Melvin
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Giedt, Gretchen
;  APPLICANT:  Jessen, Aaron
;  APPLICANT:  Jomek, Leni
;  APPLICANT:  Jones, Lee
;  APPLICANT:  Kita, David
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Laroya, Mimi
;  APPLICANT:  Lomelli, Michelle
;  APPLICANT:  Nelken, Sarah
;  APPLICANT:  Nguyen, Kody
;  APPLICANT:  Nguyen, Lynne
;  APPLICANT:  Nguyen, Phuong
;  APPLICANT:  Nogra, Margie
;  APPLICANT:  Palencia, Servando
;  APPLICANT:  Raisi, Fariba
;  APPLICANT:  Randhwa, Gurpreet
;  APPLICANT:  Sidhu, Navjiwan
;  APPLICANT:  Smith, Benjamin
;  APPLICANT:  Smythe, Ashleigh
;  APPLICANT:  Tkach, Joe
;  APPLICANT:  Tran, Lien
;  APPLICANT:  Verna, Ron
;  APPLICANT:  Wachter, Adam
;  APPLICANT:  Wu, James
;  APPLICANT:  Yim, Kenneth
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 740CIP
;  CURRENT APPLICATION NUMBER: US/09/975,640
;  CURRENT FILING DATE:  2001-10-11
;  PRIOR APPLICATION NUMBER: 09/733,811
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/088,041
;  PRIOR FILING DATE: 1998-06-02
;  NUMBER OF SEQ ID NOS: 31906
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1021
;   LENGTH: 356
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-975-640-1021

Alignment Scores:
Pred. No.:              8.97e+03        Length:         356
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     38              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-975-640-1021 (1-356)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        287 TCTCTCATTACCTGGAATATT 307

RESULT 27
US-09-975-640A-1021
; Sequence 1021, Application US/09975640A
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT:  Asghari, Vida
;  APPLICANT:  Damavandi, Simin
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drake, Jim
```

```
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Engleman, Carrie
;  APPLICANT:  Faulkner, Brandy
;  APPLICANT:  Fox, Melvin
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Giedt, Gretchen
;  APPLICANT:  Jessen, Aaron
;  APPLICANT:  Jomek, Leni
;  APPLICANT:  Jones, Lee
;  APPLICANT:  Kita, David
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Laroya, Mimi
;  APPLICANT:  Lomelli, Michelle
;  APPLICANT:  Nelken, Sarah
;  APPLICANT:  Nguyen, Kody
;  APPLICANT:  Nguyen, Lynne
;  APPLICANT:  Nguyen, Phuong
;  APPLICANT:  Nogra, Margie
;  APPLICANT:  Palencia, Servando
;  APPLICANT:  Raisi, Fariba
;  APPLICANT:  Randhwa, Gurpreet
;  APPLICANT:  Sidhu, Navjiwan
;  APPLICANT:  Smith, Benjamin
;  APPLICANT:  Smythe, Ashleigh
;  APPLICANT:  Tkach, Joe
;  APPLICANT:  Tran, Lien
;  APPLICANT:  Verna, Ron
;  APPLICANT:  Wachter, Adam
;  APPLICANT:  Wu, James
;  APPLICANT:  Yim, Kenneth
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 740CIP
;  CURRENT APPLICATION NUMBER: US/09/975,640A
;  CURRENT FILING DATE:  2001-10-11
;  PRIOR APPLICATION NUMBER: 09/733,811
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/088,041
;  PRIOR FILING DATE: 1998-06-02
;  NUMBER OF SEQ ID NOS: 31906
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1021
;   LENGTH: 356
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-975-640A-1021

Alignment Scores:
Pred. No.:                8.97e+03        Length:         356
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       38              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-975-640A-1021 (1-356)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        287 TCTCTCATTACCTGGAATATT 307

RESULT 28
US-09-515-128-17577
; Sequence 17577, Application US/09515128
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT: Asghari, Vida
;  APPLICANT: Damavandi, Simin
;  APPLICANT: Dickson, Mark
;  APPLICANT: Drake, Jim
;  APPLICANT: Drmanac, Radoje
;  APPLICANT: Engleman, Carrie
;  APPLICANT: Faulkner, Brandy
;  APPLICANT: Garcia, Veronica
;  APPLICANT: Giedt, Gretchen
;  APPLICANT: Hunter, Kelly
;  APPLICANT: Jessen, Aaron
;  APPLICANT: Jones, Lee
;  APPLICANT: Kita, David
;  APPLICANT: Labat, Ivan
;  APPLICANT: Laroya, Mimi
;  APPLICANT: Lomelli, Michelle
;  APPLICANT: Nguyen, Phuong
;  APPLICANT: Nogra, Margie
;  APPLICANT: Palencia, Servando
;  APPLICANT: Raisi, Fariba
;  APPLICANT: Smith, Benjamin
;  APPLICANT: Tkach, Joe
;  APPLICANT: Tran, Lien
;  APPLICANT: Verna, Ron
;  APPLICANT: Yang, Fei
;  APPLICANT: Yim, Kenneth
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
;  FILE REFERENCE: 728CIP
;  CURRENT APPLICATION NUMBER: US/09/515,128
;  CURRENT FILING DATE: 2000-02-28
;  EARLIER APPLICATION NUMBER: 09/034,341
;  EARLIER FILING DATE: 1998-02-13
;  NUMBER OF SEQ ID NOS: 24489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17577
;   LENGTH: 372
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-515-128-17577

Alignment Scores:
Pred. No.:                9.42e+03        Length:         372
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       25              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-515-128-17577 (1-372)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        349 TCTCTCATTACCTGGAATATT 369

RESULT 29
US-09-721-544-17577
; Sequence 17577, Application US/09721544
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT:  Asghari, Vida
;  APPLICANT:  Damavandi, Simin
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drake, Jim
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Engleman, Carrie
;  APPLICANT:  Faulkner, Brandy
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Giedt, Gretchen
;  APPLICANT:  Hunter, Kelly
;  APPLICANT:  Jessen, Aaron
;  APPLICANT:  Jones, Lee
;  APPLICANT:  Kita, David
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Laroya, Mimi
;  APPLICANT:  Lomelli, Michelle
;  APPLICANT:  Nguyen, Phuong
;  APPLICANT:  Nogra, Margie
;  APPLICANT:  Palencia, Servando
;  APPLICANT:  Raisi, Fariba
```

```
;  APPLICANT:  Smith, Benjamin
;  APPLICANT:  Tkach, Joe
;  APPLICANT:  Tran, Lien
;  APPLICANT:  Verna, Ron
;  APPLICANT:  Yang, Fei
;  APPLICANT:  Yim, Kenneth
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From a cDNA Library of Fetal Liver-Spleen
;  FILE REFERENCE: 728CIP
;  CURRENT APPLICATION NUMBER: US/09/721,544
;  CURRENT FILING DATE:  2000-11-21
;  PRIOR APPLICATION NUMBER: 09/515,128
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: 09/034,341
;  PRIOR FILING DATE: 1998-02-13
;  NUMBER OF SEQ ID NOS: 24489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17577
;   LENGTH: 372
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-721-544-17577

Alignment Scores:
Pred. No.:                9.42e+03        Length:          372
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       31              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-721-544-17577 (1-372)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        349 TCTCTCATTACCTGGAATATT 369

RESULT 30
US-09-644-868-2256
; Sequence 2256, Application US/09644868
; GENERAL INFORMATION:
;  APPLICANT:  White, David
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Pan, Yang
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1168-001
;  CURRENT APPLICATION NUMBER: US/09/644,868
;  CURRENT FILING DATE:  2000-08-28
;  PRIOR APPLICATION NUMBER:  60/151,063
;  PRIOR FILING DATE:  1999-08-27
;  NUMBER OF SEQ ID NOS:  10075
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2256
;   LENGTH: 384
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(384)
;   OTHER INFORMATION: n = A,T,C or G
US-09-644-868-2256

Alignment Scores:
Pred. No.:                9.76e+03        Length:          384
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       29              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-644-868-2256 (1-384)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        115 TCTCTCATTACCTGGAATATT 135

RESULT 31
PCT-US02-02870-826/c
; Sequence 826, Application PC/TUS0202870
; GENERAL INFORMATION:
;  APPLICANT: Corixa Coporation
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563PC
;  CURRENT APPLICATION NUMBER: PCT/US02/02870
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 826
;   LENGTH: 391
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 373
;   OTHER INFORMATION: n = A,T,C or G
PCT-US02-02870-826

Alignment Scores:
Pred. No.:                9.95e+03        Length:          391
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       1               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x PCT-US02-02870-826 (1-391)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        103 TCTCTCATTACCTGGAATATT 83

RESULT 32
US-10-066-543-826/c
; Sequence 826, Application US/10066543
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

; SEQ ID NO 826
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 373
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-826

Alignment Scores:

Pred. No.:	9.95e+03	Length:	391
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	40	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-066-543-826 (1-391)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        103 TCTCTCATTACCTGGAATATT 83
```

RESULT 33
US-10-214-403-826/c
; Sequence 826, Application US/10214403
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563C1
; CURRENT APPLICATION NUMBER: US/10/214,403
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 3420
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 826
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 373
; OTHER INFORMATION: n = A,T,C or G
US-10-214-403-826

Alignment Scores:

Pred. No.:	9.95e+03	Length:	391
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	43	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-10-214-403-826 (1-391)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        103 TCTCTCATTACCTGGAATATT 83
```

RESULT 34
US-09-362-510-41529
; Sequence 41529, Application US/09362510
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-759CON1
; CURRENT APPLICATION NUMBER: US/09/362,510
; CURRENT FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 09/221,820
; EARLIER FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 62165
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 41529
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-362-510-41529

Alignment Scores:

Pred. No.:	1e+04	Length:	394
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	23	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-362-510-41529 (1-394)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        344 TCTCTCATTACCTGGAATATT 364
```

RESULT 35
US-09-362-510A-41529
; Sequence 41529, Application US/09362510A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-759CON1
; CURRENT APPLICATION NUMBER: US/09/362,510A
; CURRENT FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 09/221,820
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 62165
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 41529
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-362-510A-41529

Alignment Scores:

Pred. No.:	1e+04	Length:	394
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	23	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-362-510A-41529 (1-394)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        344 TCTCTCATTACCTGGAATATT 364
```

RESULT 36
US-09-904-013-41529
; Sequence 41529, Application US/09904013
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-759CON1

```
;  CURRENT APPLICATION NUMBER: US/09/904,013
;  CURRENT FILING DATE:  2001-07-12
;  PRIOR APPLICATION NUMBER: 09/362,510
;  PRIOR FILING DATE: 1999-07-27
;  PRIOR APPLICATION NUMBER: US 09/221,820
;  PRIOR FILING DATE: 1998-12-30
;  NUMBER OF SEQ ID NOS: 62165
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 41529
;   LENGTH: 394
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-904-013-41529

Alignment Scores:
Pred. No.:                1e+04          Length:          394
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       35             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-904-013-41529 (1-394)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        344 TCTCTCATTACCTGGAATATT 364

RESULT 37
US-09-528-409-24161
; Sequence 24161, Application US/09528409
; GENERAL INFORMATION:
;  APPLICANT:  Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION:  Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE:  774
;  CURRENT APPLICATION NUMBER: US/09/528,409
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER:  60/125,453
;  PRIOR FILING DATE:  1999-03-19
;  NUMBER OF SEQ ID NOS:  116231
;  SOFTWARE:  Hy-patent.pl Version 3.1
; SEQ ID NO 24161
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-528-409-24161

Alignment Scores:
Pred. No.:                1.01e+04       Length:          396
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       25             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-528-409-24161 (1-396)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        159 TCTCTCATTACCTGGAATATT 179

RESULT 38
US-09-933-524-24161
; Sequence 24161, Application US/09933524
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 24161
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524-24161

Alignment Scores:
Pred. No.:                1.01e+04       Length:          396
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       36             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524-24161 (1-396)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        159 TCTCTCATTACCTGGAATATT 179

RESULT 39
US-09-933-524A-24161
; Sequence 24161, Application US/09933524A
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524A
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 24161
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524A-24161

Alignment Scores:
Pred. No.:                1.01e+04       Length:          396
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
```

DB: 36 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524A-24161 (1-396)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        159 TCTCTCATTACCTGGAATATT 179
```

RESULT 40
US-09-528-409-26306
; Sequence 26306, Application US/09528409
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/528,409
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,453
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 26306
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-528-409-26306

Alignment Scores:

Pred. No.:	1.01e+04	Length:	397
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	25	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-528-409-26306 (1-397)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        192 TCTCTCATTACCTGGAATATT 212
```

RESULT 41
US-09-933-524-26306
; Sequence 26306, Application US/09933524
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 26306
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-933-524-26306

Alignment Scores:

Pred. No.:	1.01e+04	Length:	397
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	36	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524-26306 (1-397)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        192 TCTCTCATTACCTGGAATATT 212
```

RESULT 42
US-09-933-524A-26306
; Sequence 26306, Application US/09933524A
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524A
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 26306
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-933-524A-26306

Alignment Scores:

Pred. No.:	1.01e+04	Length:	397
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	36	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524A-26306 (1-397)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        192 TCTCTCATTACCTGGAATATT 212
```

RESULT 43
US-09-489-036-16373
; Sequence 16373, Application US/09489036
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 783
; CURRENT APPLICATION NUMBER: US/09/489,036
; CURRENT FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 35324
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 16373
; LENGTH: 398
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-036-16373

Alignment Scores:

Pred. No.:	1.01e+04	Length:	398
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0

```
DB:                      24            Gaps:          0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-489-036-16373 (1-398)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        354 TCTCTCATTACCTGGAATATT 374

RESULT 44
US-09-943-143-16373
; Sequence 16373, Application US/09943143
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 783
;  CURRENT APPLICATION NUMBER: US/09/943,143
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/489,036
;  PRIOR FILING DATE: 2000-01-19
;  NUMBER OF SEQ ID NOS: 35324
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 16373
;   LENGTH: 398
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-943-143-16373

Alignment Scores:
Pred. No.:                1.01e+04       Length:          398
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       37             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-943-143-16373 (1-398)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        354 TCTCTCATTACCTGGAATATT 374

RESULT 45
US-09-489-036-16721
; Sequence 16721, Application US/09489036
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION: From Various Libraries
;  FILE REFERENCE: 783
;  CURRENT APPLICATION NUMBER: US/09/489,036
;  CURRENT FILING DATE: 2000-01-19
;  NUMBER OF SEQ ID NOS: 35324
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 16721
;   LENGTH: 403
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-489-036-16721

Alignment Scores:
Pred. No.:                1.03e+04       Length:          403
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       24             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-489-036-16721 (1-403)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        361 TCTCTCATTACCTGGAATATT 381

RESULT 46
US-09-943-143-16721
; Sequence 16721, Application US/09943143
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 783
;  CURRENT APPLICATION NUMBER: US/09/943,143
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/489,036
;  PRIOR FILING DATE: 2000-01-19
;  NUMBER OF SEQ ID NOS: 35324
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 16721
;   LENGTH: 403
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-943-143-16721

Alignment Scores:
Pred. No.:                1.03e+04       Length:          403
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       37             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-943-143-16721 (1-403)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        361 TCTCTCATTACCTGGAATATT 381

RESULT 47
US-09-489-036-16392
; Sequence 16392, Application US/09489036
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION: From Various Libraries
;  FILE REFERENCE: 783
;  CURRENT APPLICATION NUMBER: US/09/489,036
;  CURRENT FILING DATE: 2000-01-19
;  NUMBER OF SEQ ID NOS: 35324
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 16392
;   LENGTH: 409
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-489-036-16392

Alignment Scores:
Pred. No.:                1.05e+04       Length:          409
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       24             Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-489-036-16392 (1-409)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        355 TCTCTCATTACCTGGAATATT 375

RESULT 48
US-09-943-143-16392
; Sequence 16392, Application US/09943143
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
```

; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 783
; CURRENT APPLICATION NUMBER: US/09/943,143
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/489,036
; PRIOR FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 35324
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 16392
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-943-143-16392

Alignment Scores:

Pred. No.:	1.05e+04	Length:	409
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	37	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-943-143-16392 (1-409)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        355 TCTCTCATTACCTGGAATATT 375
```

RESULT 49
US-09-404-549-2518
; Sequence 2518, Application US/09404549
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN FETAL THYMUS LIBRARY
; FILE REFERENCE: MLN98-49pM
; CURRENT APPLICATION NUMBER: US/09/404,549
; CURRENT FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60/101,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106,434
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107.258
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 6007
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2518
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-549-2518

Alignment Scores:

Pred. No.:	1.06e+04	Length:	414
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	24	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-404-549-2518 (1-414)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         13 TCTCTCATTACCTGGAATATT 33
```

RESULT 50
US-09-404-549A-2518
; Sequence 2518, Application US/09404549A
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN FETAL THYMUS LIBRARY
; FILE REFERENCE: MLN98-49pM
; CURRENT APPLICATION NUMBER: US/09/404,549A
; CURRENT FILING DATE: 1999-09-24
; PRIOR APPLICATION NUMBER: 60/101,643
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106,434
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107.258
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 6007
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2518
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-404-549A-2518

Alignment Scores:

Pred. No.:	1.06e+04	Length:	414
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	24	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-404-549A-2518 (1-414)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         13 TCTCTCATTACCTGGAATATT 33
```

RESULT 51
US-09-528-409-24160
; Sequence 24160, Application US/09528409
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/528,409
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,453
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 24160
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(420)
; OTHER INFORMATION: n = A,T,C or G
US-09-528-409-24160

Alignment Scores:

Pred. No.:	1.08e+04	Length:	420
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	25	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-528-409-24160 (1-420)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
```

```
Db        158 TCTCTCATTACCTGGAATATT 178

RESULT 52
US-09-933-524-24160
; Sequence 24160, Application US/09933524
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 24160
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(420)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524-24160

Alignment Scores:
Pred. No.:                1.08e+04        Length:          420
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       36              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524-24160 (1-420)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        158 TCTCTCATTACCTGGAATATT 178

RESULT 53
US-09-933-524A-24160
; Sequence 24160, Application US/09933524A
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524A
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 24160
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(420)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524A-24160

Alignment Scores:
Pred. No.:                1.08e+04        Length:          420
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       36              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524A-24160 (1-420)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        158 TCTCTCATTACCTGGAATATT 178

RESULT 54
US-09-489-036-26234
; Sequence 26234, Application US/09489036
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION: From Various Libraries
;  FILE REFERENCE: 783
;  CURRENT APPLICATION NUMBER: US/09/489,036
;  CURRENT FILING DATE: 2000-01-19
;  NUMBER OF SEQ ID NOS: 35324
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 26234
;   LENGTH: 430
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(430)
;   OTHER INFORMATION: n = A,T,C or G
US-09-489-036-26234

Alignment Scores:
Pred. No.:                1.11e+04        Length:          430
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       24              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-489-036-26234 (1-430)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        349 TCTCTCATTACCTGGAATATT 369

RESULT 55
US-09-943-143-26234
; Sequence 26234, Application US/09943143
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 783
;  CURRENT APPLICATION NUMBER: US/09/943,143
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/489,036
;  PRIOR FILING DATE: 2000-01-19
;  NUMBER OF SEQ ID NOS: 35324
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 26234
;   LENGTH: 430
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(430)
;   OTHER INFORMATION: n = A,T,C or G
US-09-943-143-26234
```

```
Alignment Scores:
Pred. No.:                1.11e+04        Length:          430
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       37              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-943-143-26234 (1-430)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        349 TCTCTCATTACCTGGAATATT 369

RESULT 56
US-09-528-409-74096
; Sequence 74096, Application US/09528409
; GENERAL INFORMATION:
;  APPLICANT:  Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION:  Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE:  774
;  CURRENT APPLICATION NUMBER: US/09/528,409
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER:  60/125,453
;  PRIOR FILING DATE:  1999-03-19
;  NUMBER OF SEQ ID NOS:  116231
;  SOFTWARE:  Hy-patent.pl Version 3.1
; SEQ ID NO 74096
;   LENGTH: 433
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-528-409-74096

Alignment Scores:
Pred. No.:                1.11e+04        Length:          433
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       25              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-528-409-74096 (1-433)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        263 TCTCTCATTACCTGGAATATT 283

RESULT 57
US-09-933-524-74096
; Sequence 74096, Application US/09933524
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 74096
;   LENGTH: 433
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-933-524-74096

Alignment Scores:
Pred. No.:                1.11e+04        Length:          433
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       36              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524-74096 (1-433)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        263 TCTCTCATTACCTGGAATATT 283

RESULT 58
US-09-933-524A-74096
; Sequence 74096, Application US/09933524A
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524A
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 74096
;   LENGTH: 433
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-933-524A-74096

Alignment Scores:
Pred. No.:                1.11e+04        Length:          433
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       36              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524A-74096 (1-433)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        263 TCTCTCATTACCTGGAATATT 283

RESULT 59
US-09-353-690-10783
; Sequence 10783, Application US/09353690
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION: From cDNA Libraries Of Fetal Liver-Spleen and Infant Brain
;  FILE REFERENCE: 20411-726CON2
;  CURRENT APPLICATION NUMBER: US/09/353,690
;  CURRENT FILING DATE: 1999-07-14
;  EARLIER APPLICATION NUMBER: US 09/217,517
;  EARLIER FILING DATE: 1998-12-22
;  EARLIER APPLICATION NUMBER: US 09/004,182
;  EARLIER FILING DATE: 1998-01-07
;  NUMBER OF SEQ ID NOS: 12181
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10783
;   LENGTH: 438
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(438)
;   OTHER INFORMATION: n = A,T,C or G
US-09-353-690-10783

Alignment Scores:
Pred. No.:                1.13e+04        Length:         438
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       23              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-353-690-10783 (1-438)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        352 TCTCTCATTACCTGGAATATT 372

RESULT 60
US-09-528-409-3652
; Sequence 3652, Application US/09528409
; GENERAL INFORMATION:
;  APPLICANT:  Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION:  Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE:  774
;  CURRENT APPLICATION NUMBER: US/09/528,409
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER:  60/125,453
;  PRIOR FILING DATE:  1999-03-19
;  NUMBER OF SEQ ID NOS:  116231
;  SOFTWARE:  Hy-patent.pl Version 3.1
; SEQ ID NO 3652
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(438)
;   OTHER INFORMATION: n = A,T,C or G
US-09-528-409-3652

Alignment Scores:
Pred. No.:                1.13e+04        Length:         438
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       25              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-528-409-3652 (1-438)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        360 TCTCTCATTACCTGGAATATT 380

RESULT 61
US-09-922-340-10783
; Sequence 10783, Application US/09922340
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From cDNA Libraries Of Fetal Liver-Spleen and Infant Brain
;  FILE REFERENCE: 20411-726CON2
;  CURRENT APPLICATION NUMBER: US/09/922,340
;  CURRENT FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: 09/353,690
;  PRIOR FILING DATE: 1999-07-14
;  PRIOR APPLICATION NUMBER: US 09/004,182
;  PRIOR FILING DATE: 1998-01-07
;  NUMBER OF SEQ ID NOS: 12181
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10783
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(438)
;   OTHER INFORMATION: n = A,T,C or G
US-09-922-340-10783

Alignment Scores:
Pred. No.:                1.13e+04        Length:         438
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       35              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-922-340-10783 (1-438)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        352 TCTCTCATTACCTGGAATATT 372

RESULT 62
US-09-922-340A-10783
; Sequence 10783, Application US/09922340A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From cDNA Libraries Of Fetal Liver-Spleen and Infant Brain
;  FILE REFERENCE: 20411-726CON2
;  CURRENT APPLICATION NUMBER: US/09/922,340A
;  CURRENT FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: US/09/353,690
;  PRIOR FILING DATE: 1999-07-14
;  PRIOR APPLICATION NUMBER: US 09/217,517
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: US 09/004,182
;  PRIOR FILING DATE: 1998-01-07
;  NUMBER OF SEQ ID NOS: 12181
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10783
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(438)
;   OTHER INFORMATION: n = A,T,C or G
US-09-922-340A-10783

Alignment Scores:
Pred. No.:                1.13e+04        Length:         438
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       35              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-922-340A-10783 (1-438)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        352 TCTCTCATTACCTGGAATATT 372
```

RESULT 63
US-09-933-524-3652
; Sequence 3652, Application US/09933524
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 3652
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(438)
; OTHER INFORMATION: n = A,T,C or G
US-09-933-524-3652

Alignment Scores:

Pred. No.:	1.13e+04	Length:	438
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	36	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524-3652 (1-438)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        360 TCTCTCATTACCTGGAATATT 380
```

RESULT 64
US-09-933-524A-3652
; Sequence 3652, Application US/09933524A
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524A
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 3652
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(438)
; OTHER INFORMATION: n = A,T,C or G
US-09-933-524A-3652

Alignment Scores:

Pred. No.:	1.13e+04	Length:	438
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	36	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524A-3652 (1-438)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        360 TCTCTCATTACCTGGAATATT 380
```

RESULT 65
US-09-321-214-11519
; Sequence 11519, Application US/09321214
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Ford, John
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-740
; CURRENT APPLICATION NUMBER: US/09/321,214
; CURRENT FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: US 60/088,041
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-321-214-11519

Alignment Scores:

Pred. No.:	1.13e+04	Length:	439
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	23	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-321-214-11519 (1-439)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        353 TCTCTCATTACCTGGAATATT 373
```

RESULT 66
US-09-516-335-11519
; Sequence 11519, Application US/09516335
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron

; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/516,335
; CURRENT FILING DATE: 2000-03-01
; EARLIER APPLICATION NUMBER: 09/321,214
; EARLIER FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: 60/088,041
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-516-335-11519

Alignment Scores:

Pred. No.:	1.13e+04	Length:	439
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	25	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-516-335-11519 (1-439)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        353 TCTCTCATTACCTGGAATATT 373
```

RESULT 67
US-09-733-811-11519
; Sequence 11519, Application US/09733811
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/733,811
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/321,214
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-733-811-11519

Alignment Scores:

Pred. No.:	1.13e+04	Length:	439
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	31	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-733-811-11519 (1-439)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        353 TCTCTCATTACCTGGAATATT 373
```

RESULT 68
US-09-733-811A-11519
; Sequence 11519, Application US/09733811A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Ford, John
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-740
; CURRENT APPLICATION NUMBER: US/09/733,811A
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/321,214

; PRIOR FILING DATE: 1999-05-26
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-733-811A-11519

Alignment Scores:

Pred. No.:	1.13e+04	Length:	439
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	31	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-733-811A-11519 (1-439)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        353 TCTCTCATTACCTGGAATATT 373
```

RESULT 69
US-09-975-640-11519
; Sequence 11519, Application US/09975640
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/975,640
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/733,811
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-975-640-11519

Alignment Scores:

Pred. No.:	1.13e+04	Length:	439
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	38	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-975-640-11519 (1-439)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        353 TCTCTCATTACCTGGAATATT 373
```

RESULT 70
US-09-975-640A-11519
; Sequence 11519, Application US/09975640A
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjiwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/975,640A
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/733,811

; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-975-640A-11519

Alignment Scores:

Pred. No.:	1.13e+04	Length:	439
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	38	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-975-640A-11519 (1-439)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        353 TCTCTCATTACCTGGAATATT 373
```

RESULT 71
US-09-528-409-14287
; Sequence 14287, Application US/09528409
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/528,409
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,453
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 14287
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(440)
; OTHER INFORMATION: n = A,T,C or G
US-09-528-409-14287

Alignment Scores:

Pred. No.:	1.13e+04	Length:	440
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	25	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-528-409-14287 (1-440)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         87 TCTCTCATTACCTGGAATATT 107
```

RESULT 72
US-09-933-524-14287
; Sequence 14287, Application US/09933524
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 14287
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(440)
; OTHER INFORMATION: n = A,T,C or G
US-09-933-524-14287

Alignment Scores:

Pred. No.:	1.13e+04	Length:	440
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	36	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524-14287 (1-440)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         87 TCTCTCATTACCTGGAATATT 107
```

RESULT 73
US-09-933-524A-14287
; Sequence 14287, Application US/09933524A
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/933,524A
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/528,409
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 14287
; LENGTH: 440
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(440)
; OTHER INFORMATION: n = A,T,C or G
US-09-933-524A-14287

Alignment Scores:

Pred. No.:	1.13e+04	Length:	440
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0

```
Query Match:            100.0%          Indels:         0
DB:                     36              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-933-524A-14287 (1-440)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         87 TCTCTCATTACCTGGAATATT 107

RESULT 74
US-09-304-649-421
; Sequence 421, Application US/09304649
; GENERAL INFORMATION:
;  APPLICANT: Gearing, David P.
;  APPLICANT: McCarthy, Sean A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
;  TITLE OF INVENTION: HUMAN MICROVASCULAR ENDOTHELIAL LIBRARY
;  FILE REFERENCE: MLN98-10pM
;  CURRENT APPLICATION NUMBER: US/09/304,649
;  CURRENT FILING DATE: 1999-05-04
;  EARLIER APPLICATION NUMBER: 60/084,098
;  EARLIER FILING DATE: 1998-05-04
;  EARLIER APPLICATION NUMBER: 60/123,523
;  EARLIER FILING DATE: 1999-03-08
;  EARLIER APPLICATION NUMBER: 60/126,974
;  EARLIER FILING DATE: 1999-03-30
;  NUMBER OF SEQ ID NOS: 1311
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 421
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(449)
;   OTHER INFORMATION: n = A,T,C or G
US-09-304-649-421

Alignment Scores:
Pred. No.:              1.16e+04        Length:         449
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     23              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-304-649-421 (1-449)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         47 TCTCTCATTACCTGGAATATT 67

RESULT 75
US-09-840-424-421
; Sequence 421, Application US/09840424
; GENERAL INFORMATION:
;  APPLICANT: Gearing, David P.
;  APPLICANT:  McCarthy, Sean A.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
;  TITLE OF INVENTION:  HUMAN MICROVASCULAR ENDOTHELIAL LIBRARY
;  FILE REFERENCE: 1600.1010-002
;  CURRENT APPLICATION NUMBER: US/09/840,424
;  CURRENT FILING DATE:  2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/084,098
;  PRIOR FILING DATE: 1998-05-04
;  PRIOR APPLICATION NUMBER: US 60/123,523
;  PRIOR FILING DATE: 1999-03-08
;  PRIOR APPLICATION NUMBER: US 60/126,974
;  PRIOR FILING DATE: 1999-03-30
;  PRIOR APPLICATION NUMBER: US 09/304,649
;  PRIOR FILING DATE: 1999-05-04
;  NUMBER OF SEQ ID NOS: 1311
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 421
;   LENGTH: 449
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(449)
;   OTHER INFORMATION: n = A,T,C or G
US-09-840-424-421

Alignment Scores:
Pred. No.:              1.16e+04        Length:         449
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     33              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x US-09-840-424-421 (1-449)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         47 TCTCTCATTACCTGGAATATT 67


Search completed: March 11, 2006, 12:05:42
Job time : 296.547 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:19:50 ; Search time 3575.87 Seconds
(without alignments)
1345.232 Million cell updates/sec

Title: US-10-757-745-2_COPY_54_140
Perfect score: 448
Sequence: 1 MERALNSYFEPPVEESALER.....IDGLDLNNLSERARGVCSYL 87

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170333_4165/app_query.fasta_1
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=rnpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=absso2h -USER=US10757745_@CGN_1_1_20983_@runat_09032006_170333_4165
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCTUSB_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/PCTUSC_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US075_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US076_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US077_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US078_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US079_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US101_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US104_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US105_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US112_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US117_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US600_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US601_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US603_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US605_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US606_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US607_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	448	100.0	351	25	US-09-515-128-10314	Sequence 10314, A
2	448	100.0	351	31	US-09-721-544-10314	Sequence 10314, A
3	448	100.0	430	24	US-09-489-036-26234	Sequence 26234, A
4	448	100.0	430	37	US-09-943-143-26234	Sequence 26234, A
5	448	100.0	433	25	US-09-528-409-74096	Sequence 74096, A
6	448	100.0	433	36	US-09-933-524-74096	Sequence 74096, A
7	448	100.0	433	36	US-09-933-524A-74096	Sequence 74096, A
8	448	100.0	438	23	US-09-353-690-10783	Sequence 10783, A

9	448	100.0	438	25	US-09-528-409-3652	Sequence 3652, Ap
10	448	100.0	438	35	US-09-922-340-10783	Sequence 10783, A
11	448	100.0	438	35	US-09-922-340A-10783	Sequence 10783, A
12	448	100.0	438	36	US-09-933-524-3652	Sequence 3652, Ap
13	448	100.0	438	36	US-09-933-524A-3652	Sequence 3652, Ap
14	448	100.0	439	23	US-09-321-214-11519	Sequence 11519, A
15	448	100.0	439	25	US-09-516-335-11519	Sequence 11519, A
16	448	100.0	439	31	US-09-733-811-11519	Sequence 11519, A
17	448	100.0	439	31	US-09-733-811A-11519	Sequence 11519, A
18	448	100.0	439	38	US-09-975-640-11519	Sequence 11519, A
19	448	100.0	439	38	US-09-975-640A-11519	Sequence 11519, A
20	448	100.0	483	19	US-08-905-144-190	Sequence 190, App
21	448	100.0	483	27	US-09-547-599C-825	Sequence 825, App
22	448	100.0	483	54	US-10-631-402-825	Sequence 825, App
23	448	100.0	483	54	US-10-631-441-825	Sequence 825, App
24	448	100.0	483	54	US-10-631-441A-825	Sequence 825, App
25	448	100.0	501	29	US-09-652-914-4701	Sequence 4701, Ap
26	448	100.0	505	22	US-09-205-070-10302	Sequence 10302, A
27	448	100.0	505	23	US-09-340-623-10302	Sequence 10302, A
28	448	100.0	505	34	US-09-898-888-10302	Sequence 10302, A
29	448	100.0	505	34	US-09-898-888A-10302	Sequence 10302, A
30	448	100.0	505	48	US-10-302-689A-10302	Sequence 10302, A
31	448	100.0	517	31	US-09-726-802-1182	Sequence 1182, Ap
32	448	100.0	1296	1	PCT-US00-05883-170	Sequence 170, App
33	448	100.0	1296	36	US-09-925-299-170	Sequence 170, App
c 34	448	100.0	1849	42	US-10-170-235-26834	Sequence 26834, A
35	448	100.0	1852	29	US-09-644-868-7772	Sequence 7772, Ap
36	448	100.0	1852	29	US-09-649-161-8751	Sequence 8751, Ap
37	448	100.0	1852	29	US-09-652-109-8165	Sequence 8165, Ap
38	448	100.0	1852	29	US-09-652-128-7441	Sequence 7441, Ap
39	448	100.0	1852	29	US-09-652-816-8491	Sequence 8491, Ap
40	448	100.0	1852	29	US-09-652-914-8450	Sequence 8450, Ap
41	448	100.0	1919	42	US-10-170-235-27345	Sequence 27345, A
42	448	100.0	1920	61	US-10-757-745-1	Sequence 1, Appli
43	448	100.0	1940	3	PCT-US04-24424-921	Sequence 921, App
44	448	100.0	1940	61	US-10-783-271-28	Sequence 28, Appl
45	448	100.0	1940	80	US-60-490-890-921	Sequence 921, App
46	448	100.0	1940	81	US-60-568-073-582	Sequence 582, App
47	448	100.0	1940	82	US-60-671-501-156	Sequence 156, App
48	448	100.0	1940	82	US-60-685-372-523	Sequence 523, App
49	448	100.0	1940	83	US-60-704-740-28	Sequence 28, Appl
50	448	100.0	1940	83	US-60-710-726-741	Sequence 741, App
51	448	100.0	1948	24	US-09-488-725A-1200	Sequence 1200, Ap
52	448	100.0	1948	40	US-10-037-270-889	Sequence 889, App
53	448	100.0	1948	42	US-10-105-837-889	Sequence 889, App
54	448	100.0	1948	42	US-10-117-722-889	Sequence 889, App
55	448	100.0	1948	42	US-10-122-851-889	Sequence 889, App
56	448	100.0	1948	43	US-10-258-898A-1200	Sequence 1200, Ap
57	448	100.0	1948	43	US-10-286-897-1200	Sequence 1200, Ap
58	448	100.0	1958	24	US-09-471-275-1314	Sequence 1314, Ap
59	448	100.0	1958	24	US-09-488-725A-4772	Sequence 4772, Ap
60	448	100.0	1958	24	US-09-488-725B-8295	Sequence 8295, Ap
61	448	100.0	1958	27	US-09-552-317-8295	Sequence 8295, Ap
62	448	100.0	1958	43	US-10-258-898A-4772	Sequence 4772, Ap
63	448	100.0	1958	43	US-10-286-897-4772	Sequence 4772, Ap
64	448	100.0	1958	48	US-10-302-689A-59957	Sequence 59957, A
65	448	100.0	1958	48	US-10-302-689A-77389	Sequence 77389, A
66	448	100.0	2262	28	US-09-606-680-2868	Sequence 2868, Ap
67	448	100.0	2262	28	US-09-606-776-4085	Sequence 4085, Ap
68	448	100.0	2262	29	US-09-644-869-9498	Sequence 9498, Ap
69	448	100.0	2262	29	US-09-652-121-6794	Sequence 6794, Ap
70	448	100.0	2262	29	US-09-652-122-3756	Sequence 3756, Ap
71	448	100.0	2262	29	US-09-652-128-9329	Sequence 9329, Ap
72	448	100.0	2262	29	US-09-652-816-9572	Sequence 9572, Ap
73	448	100.0	2262	31	US-09-700-000-6419	Sequence 6419, Ap
74	448	100.0	2262	31	US-09-710-280-2045	Sequence 2045, Ap
75	448	100.0	2262	31	US-09-710-282-3926	Sequence 3926, Ap

ALIGNMENTS

RESULT 1
US-09-515-128-10314

; Sequence 10314, Application US/09515128
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jessen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fei
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/515,128
; CURRENT FILING DATE: 2000-02-28
; EARLIER APPLICATION NUMBER: 09/034,341
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10314
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-515-128-10314

Alignment Scores:

Pred. No.:	2.56e-63	Length:	351
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	25	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-515-128-10314 (1-351)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 124

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 184

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 244

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 304

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        305 CGAGGGGTGTGTTCCTACTTA 325
```

```
RESULT 2
US-09-721-544-10314
; Sequence 10314, Application US/09721544
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT:  Asghari, Vida
;  APPLICANT:  Damavandi, Simin
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drake, Jim
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Engleman, Carrie
;  APPLICANT:  Faulkner, Brandy
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Giedt, Gretchen
;  APPLICANT:  Hunter, Kelly
;  APPLICANT:  Jessen, Aaron
;  APPLICANT:  Jones, Lee
;  APPLICANT:  Kita, David
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Laroya, Mimi
;  APPLICANT:  Lomelli, Michelle
;  APPLICANT:  Nguyen, Phuong
;  APPLICANT:  Nogra, Margie
;  APPLICANT:  Palencia, Servando
;  APPLICANT:  Raisi, Fariba
;  APPLICANT:  Smith, Benjamin
;  APPLICANT:  Tkach, Joe
;  APPLICANT:  Tran, Lien
;  APPLICANT:  Verna, Ron
;  APPLICANT:  Yang, Fei
;  APPLICANT:  Yim, Kenneth
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From a cDNA Library of Fetal Liver-Spleen
;  FILE REFERENCE: 728CIP
;  CURRENT APPLICATION NUMBER: US/09/721,544
;  CURRENT FILING DATE:  2000-11-21
;  PRIOR APPLICATION NUMBER: 09/515,128
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: 09/034,341
;  PRIOR FILING DATE: 1998-02-13
;  NUMBER OF SEQ ID NOS: 24489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10314
;   LENGTH: 351
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-721-544-10314

Alignment Scores:
Pred. No.:              2.56e-63      Length:          351
Score:                  448.00        Matches:         87
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     31            Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-721-544-10314 (1-351)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 124

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 184

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 244

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        245 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 304

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        305 CGAGGGGTGTGTTCCTACTTA 325

RESULT 3
US-09-489-036-26234
; Sequence 26234, Application US/09489036
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION: From Various Libraries
;  FILE REFERENCE: 783
;  CURRENT APPLICATION NUMBER: US/09/489,036
;  CURRENT FILING DATE: 2000-01-19
;  NUMBER OF SEQ ID NOS: 35324
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 26234
;   LENGTH: 430
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(430)
;   OTHER INFORMATION: n = A,T,C or G
US-09-489-036-26234

Alignment Scores:
Pred. No.:              3.61e-63      Length:          430
Score:                  448.00        Matches:         87
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     24            Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-489-036-26234 (1-430)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 225

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 285

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 345

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 405

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        406 CGAGGGGTGTGTTCCTACTTA 426

RESULT 4
US-09-943-143-26234
; Sequence 26234, Application US/09943143
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 783
;  CURRENT APPLICATION NUMBER: US/09/943,143
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/489,036
;  PRIOR FILING DATE: 2000-01-19
;  NUMBER OF SEQ ID NOS: 35324
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 26234
```

```
;   LENGTH: 430
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(430)
;   OTHER INFORMATION: n = A,T,C or G
US-09-943-143-26234

Alignment Scores:
Pred. No.:                3.61e-63        Length:         430
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       37              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-943-143-26234 (1-430)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 225

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 285

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 345

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 405

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        406 CGAGGGGTGTGTTCCTACTTA 426

RESULT 5
US-09-528-409-74096
; Sequence 74096, Application US/09528409
; GENERAL INFORMATION:
;  APPLICANT:  Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION:  Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE:  774
;  CURRENT APPLICATION NUMBER: US/09/528,409
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER:  60/125,453
;  PRIOR FILING DATE:  1999-03-19
;  NUMBER OF SEQ ID NOS:  116231
;  SOFTWARE:  Hy-patent.pl Version 3.1
; SEQ ID NO 74096
;   LENGTH: 433
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-528-409-74096

Alignment Scores:
Pred. No.:                3.65e-63        Length:         433
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       25              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-528-409-74096 (1-433)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 139

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 199

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 259

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 319

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        320 CGAGGGGTGTGTTCCTACTTA 340

RESULT 6
US-09-933-524-74096
; Sequence 74096, Application US/09933524
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 74096
;   LENGTH: 433
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-933-524-74096

Alignment Scores:
Pred. No.:                3.65e-63        Length:         433
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       36              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-933-524-74096 (1-433)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 139

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 199

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 259

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 319

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        320 CGAGGGGTGTGTTCCTACTTA 340
```

```
RESULT 7
US-09-933-524A-74096
; Sequence 74096, Application US/09933524A
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524A
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 74096
;   LENGTH: 433
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-933-524A-74096

Alignment Scores:
Pred. No.:              3.65e-63        Length:         433
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     36              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-933-524A-74096 (1-433)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 139

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 199

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 259

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 319

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        320 CGAGGGGTGTGTTCCTACTTA 340

RESULT 8
US-09-353-690-10783
; Sequence 10783, Application US/09353690
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION: From cDNA Libraries Of Fetal Liver-Spleen and Infant Brain
;  FILE REFERENCE: 20411-726CON2
;  CURRENT APPLICATION NUMBER: US/09/353,690
;  CURRENT FILING DATE: 1999-07-14
;  EARLIER APPLICATION NUMBER: US 09/217,517
;  EARLIER FILING DATE: 1998-12-22
;  EARLIER APPLICATION NUMBER: US 09/004,182
;  EARLIER FILING DATE: 1998-01-07
;  NUMBER OF SEQ ID NOS: 12181
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10783
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(438)
;   OTHER INFORMATION: n = A,T,C or G
US-09-353-690-10783

Alignment Scores:
Pred. No.:              3.72e-63        Length:         438
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     23              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-353-690-10783 (1-438)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 228

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 288

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 348

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 408

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        409 CGAGGGGTGTGTTCCTACTTA 429

RESULT 9
US-09-528-409-3652
; Sequence 3652, Application US/09528409
; GENERAL INFORMATION:
;  APPLICANT:  Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION:  Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE:  774
;  CURRENT APPLICATION NUMBER: US/09/528,409
;  CURRENT FILING DATE:  2000-03-17
;  PRIOR APPLICATION NUMBER:  60/125,453
;  PRIOR FILING DATE:  1999-03-19
;  NUMBER OF SEQ ID NOS:  116231
;  SOFTWARE:  Hy-patent.pl Version 3.1
; SEQ ID NO 3652
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(438)
;   OTHER INFORMATION: n = A,T,C or G
US-09-528-409-3652

Alignment Scores:
Pred. No.:              3.72e-63        Length:         438
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     25              Gaps:           0
```

```
US-10-757-745-2_COPY_54_140 (1-87) x US-09-528-409-3652 (1-438)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 236

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 296

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 356

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 416

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        417 CGAGGGGTGTGTTCCTACTTA 437

RESULT 10
US-09-922-340-10783
; Sequence 10783, Application US/09922340
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From cDNA Libraries Of Fetal Liver-Spleen and Infant Brain
;  FILE REFERENCE: 20411-726CON2
;  CURRENT APPLICATION NUMBER: US/09/922,340
;  CURRENT FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: 09/353,690
;  PRIOR FILING DATE: 1999-07-14
;  PRIOR APPLICATION NUMBER: US 09/004,182
;  PRIOR FILING DATE: 1998-01-07
;  NUMBER OF SEQ ID NOS: 12181
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10783
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(438)
;   OTHER INFORMATION: n = A,T,C or G
US-09-922-340-10783

Alignment Scores:
Pred. No.:                3.72e-63        Length:          438
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       35              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-922-340-10783 (1-438)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 228

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 288

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 348

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 408

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        409 CGAGGGGTGTGTTCCTACTTA 429

RESULT 11
US-09-922-340A-10783
; Sequence 10783, Application US/09922340A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From cDNA Libraries Of Fetal Liver-Spleen and Infant Brain
;  FILE REFERENCE: 20411-726CON2
;  CURRENT APPLICATION NUMBER: US/09/922,340A
;  CURRENT FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: US/09/353,690
;  PRIOR FILING DATE: 1999-07-14
;  PRIOR APPLICATION NUMBER: US 09/217,517
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: US 09/004,182
;  PRIOR FILING DATE: 1998-01-07
;  NUMBER OF SEQ ID NOS: 12181
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10783
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(438)
;   OTHER INFORMATION: n = A,T,C or G
US-09-922-340A-10783

Alignment Scores:
Pred. No.:                3.72e-63        Length:          438
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       35              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-922-340A-10783 (1-438)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 228

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 288

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 348

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 408

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        409 CGAGGGGTGTGTTCCTACTTA 429

RESULT 12
US-09-933-524-3652
; Sequence 3652, Application US/09933524
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
```

```
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-03-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 3652
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(438)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524-3652

Alignment Scores:
Pred. No.:              3.72e-63        Length:         438
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     36              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-933-524-3652 (1-438)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 236

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 296

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 356

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 416

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        417 CGAGGGGTGTGTTCCTACTTA 437

RESULT 13
US-09-933-524A-3652
; Sequence 3652, Application US/09933524A
; GENERAL INFORMATION:
;  APPLICANT: Drmanac, Radoje T.
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Jones, Lee W.
;  TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 774
;  CURRENT APPLICATION NUMBER: US/09/933,524A
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 09/528,409
;  PRIOR FILING DATE: 2000-05-17
;  NUMBER OF SEQ ID NOS: 116231
;  SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 3652
;   LENGTH: 438
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(438)
;   OTHER INFORMATION: n = A,T,C or G
US-09-933-524A-3652

Alignment Scores:
Pred. No.:              3.72e-63        Length:         438
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     36              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-933-524A-3652 (1-438)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 236

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 296

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 356

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 416

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        417 CGAGGGGTGTGTTCCTACTTA 437

RESULT 14
US-09-321-214-11519
; Sequence 11519, Application US/09321214
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  APPLICANT: Drmanac, Radoje
;  APPLICANT: Crkvenjakov, Radomir
;  APPLICANT: Dickson, Mark
;  APPLICANT: Drmanac, Snezana
;  APPLICANT: Labat, Ivan
;  APPLICANT: Leshkowitz, Dena
;  APPLICANT: Kita, David
;  APPLICANT: Ford, John
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-740
;  CURRENT APPLICATION NUMBER: US/09/321,214
;  CURRENT FILING DATE: 1999-05-26
;  EARLIER APPLICATION NUMBER: US 60/088,041
;  EARLIER FILING DATE: 1998-06-02
;  NUMBER OF SEQ ID NOS: 31906
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(439)
;   OTHER INFORMATION: n = A,T,C or G
US-09-321-214-11519

Alignment Scores:
Pred. No.:              3.74e-63        Length:         439
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     23              Gaps:           0
```

```
US-10-757-745-2_COPY_54_140 (1-87) x US-09-321-214-11519 (1-439)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 15
US-09-516-335-11519
; Sequence 11519, Application US/09516335
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT: Asghari, Vida
;  APPLICANT: Damavandi, Simin
;  APPLICANT: Dickson, Mark
;  APPLICANT: Drake, Jim
;  APPLICANT: Drmanac, Radoje
;  APPLICANT: Engleman, Carrie
;  APPLICANT: Faulkner, Brandy
;  APPLICANT: Fox, Melvin
;  APPLICANT: Garcia, Veronica
;  APPLICANT: Giedt, Gretchen
;  APPLICANT: Jessen, Aaron
;  APPLICANT: Jomek, Leni
;  APPLICANT: Jones, Lee
;  APPLICANT: Kita, David
;  APPLICANT: Labat, Ivan
;  APPLICANT: Laroya, Mimi
;  APPLICANT: Lomelli, Michelle
;  APPLICANT: Nelken, Sarah
;  APPLICANT: Nguyen, Kody
;  APPLICANT: Nguyen, Lynne
;  APPLICANT: Nguyen, Phuong
;  APPLICANT: Nogra, Margie
;  APPLICANT: Palencia, Servando
;  APPLICANT: Raisi, Fariba
;  APPLICANT: Randhwa, Gurpreet
;  APPLICANT: Sidhu, Navjiwan
;  APPLICANT: Smith, Benjamin
;  APPLICANT: Smythe, Ashleigh
;  APPLICANT: Tkach, Joe
;  APPLICANT: Tran, Lien
;  APPLICANT: Verna, Ron
;  APPLICANT: Wachter, Adam
;  APPLICANT: Wu, James
;  APPLICANT: Yim, Kenneth
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 740CIP
;  CURRENT APPLICATION NUMBER: US/09/516,335
;  CURRENT FILING DATE: 2000-03-01
;  EARLIER APPLICATION NUMBER: 09/321,214
;  EARLIER FILING DATE: 1999-05-26
;  EARLIER APPLICATION NUMBER: 60/088,041
;  EARLIER FILING DATE: 1998-06-02
;  NUMBER OF SEQ ID NOS: 31906
;  SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 11519
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(439)
;   OTHER INFORMATION: n = A,T,C or G
US-09-516-335-11519

Alignment Scores:
Pred. No.:                3.74e-63        Length:          439
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       25              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-516-335-11519 (1-439)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 16
US-09-733-811-11519
; Sequence 11519, Application US/09733811
; GENERAL INFORMATION:
;  APPLICANT:  Arterburn, Matthew
;  APPLICANT:  Asghari, Vida
;  APPLICANT:  Damavandi, Simin
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drake, Jim
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Engleman, Carrie
;  APPLICANT:  Faulkner, Brandy
;  APPLICANT:  Fox, Melvin
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Giedt, Gretchen
;  APPLICANT:  Jessen, Aaron
;  APPLICANT:  Jomek, Leni
;  APPLICANT:  Jones, Lee
;  APPLICANT:  Kita, David
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Laroya, Mimi
;  APPLICANT:  Lomelli, Michelle
;  APPLICANT:  Nelken, Sarah
;  APPLICANT:  Nguyen, Kody
;  APPLICANT:  Nguyen, Lynne
;  APPLICANT:  Nguyen, Phuong
;  APPLICANT:  Nogra, Margie
;  APPLICANT:  Palencia, Servando
;  APPLICANT:  Raisi, Fariba
;  APPLICANT:  Randhwa, Gurpreet
;  APPLICANT:  Sidhu, Navjiwan
;  APPLICANT:  Smith, Benjamin
;  APPLICANT:  Smythe, Ashleigh
```

```
;  APPLICANT:  Tkach, Joe
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE:  740CIP
;  CURRENT APPLICATION NUMBER: US/09/733,811
;  CURRENT FILING DATE:  2000-12-08
;  PRIOR APPLICATION NUMBER:  09/321,214
;  PRIOR FILING DATE:  1999-05-26
;  PRIOR APPLICATION NUMBER:  60/088,041
;  PRIOR FILING DATE:  1998-06-02
;  NUMBER OF SEQ ID NOS:  31906
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(439)
;   OTHER INFORMATION: n = A,T,C or G
US-09-733-811-11519

Alignment Scores:
Pred. No.:              3.74e-63        Length:         439
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-733-811-11519 (1-439)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 17
US-09-733-811A-11519
; Sequence 11519, Application US/09733811A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Crkvenjakov, Radomir
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drmanac, Snezana
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Kita, David
;  APPLICANT:  Ford, John
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-740
;  CURRENT APPLICATION NUMBER: US/09/733,811A
;  CURRENT FILING DATE:  2000-12-08
;  PRIOR APPLICATION NUMBER: 09/321,214
;  PRIOR FILING DATE: 1999-05-26
;  NUMBER OF SEQ ID NOS: 31906
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(439)
;   OTHER INFORMATION: n = A,T,C or G
US-09-733-811A-11519

Alignment Scores:
Pred. No.:              3.74e-63        Length:         439
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-733-811A-11519 (1-439)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 18
US-09-975-640-11519
; Sequence 11519, Application US/09975640
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT:  Asghari, Vida
;  APPLICANT:  Damavandi, Simin
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drake, Jim
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Engleman, Carrie
;  APPLICANT:  Faulkner, Brandy
;  APPLICANT:  Fox, Melvin
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Giedt, Gretchen
;  APPLICANT:  Jessen, Aaron
;  APPLICANT:  Jomek, Leni
;  APPLICANT:  Jones, Lee
;  APPLICANT:  Kita, David
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Laroya, Mimi
;  APPLICANT:  Lomelli, Michelle
;  APPLICANT:  Nelken, Sarah
;  APPLICANT:  Nguyen, Kody
;  APPLICANT:  Nguyen, Lynne
;  APPLICANT:  Nguyen, Phuong
;  APPLICANT:  Nogra, Margie
;  APPLICANT:  Palencia, Servando
;  APPLICANT:  Raisi, Fariba
;  APPLICANT:  Randhwa, Gurpreet
;  APPLICANT:  Sidhu, Navjiwan
;  APPLICANT:  Smith, Benjamin
```

```
;  APPLICANT:  Smythe, Ashleigh
;  APPLICANT:  Tkach, Joe
;  APPLICANT:  Tran, Lien
;  APPLICANT:  Verna, Ron
;  APPLICANT:  Wachter, Adam
;  APPLICANT:  Wu, James
;  APPLICANT:  Yim, Kenneth
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 740CIP
;  CURRENT APPLICATION NUMBER: US/09/975,640
;  CURRENT FILING DATE:  2001-10-11
;  PRIOR APPLICATION NUMBER: 09/733,811
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/088,041
;  PRIOR FILING DATE: 1998-06-02
;  NUMBER OF SEQ ID NOS: 31906
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(439)
;   OTHER INFORMATION: n = A,T,C or G
US-09-975-640-11519

Alignment Scores:
Pred. No.:                3.74e-63        Length:          439
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       38              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-975-640-11519 (1-439)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 19
US-09-975-640A-11519
; Sequence 11519, Application US/09975640A
; GENERAL INFORMATION:
;  APPLICANT: Arterburn, Matthew
;  APPLICANT:  Asghari, Vida
;  APPLICANT:  Damavandi, Simin
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drake, Jim
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Engleman, Carrie
;  APPLICANT:  Faulkner, Brandy
;  APPLICANT:  Fox, Melvin
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Giedt, Gretchen
;  APPLICANT:  Jessen, Aaron
;  APPLICANT:  Jomek, Leni
;  APPLICANT:  Jones, Lee
;  APPLICANT:  Kita, David
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Laroya, Mimi
;  APPLICANT:  Lomelli, Michelle
;  APPLICANT:  Nelken, Sarah
;  APPLICANT:  Nguyen, Kody
;  APPLICANT:  Nguyen, Lynne
;  APPLICANT:  Nguyen, Phuong
;  APPLICANT:  Nogra, Margie
;  APPLICANT:  Palencia, Servando
;  APPLICANT:  Raisi, Fariba
;  APPLICANT:  Randhwa, Gurpreet
;  APPLICANT:  Sidhu, Navjiwan
;  APPLICANT:  Smith, Benjamin
;  APPLICANT:  Smythe, Ashleigh
;  APPLICANT:  Tkach, Joe
;  APPLICANT:  Tran, Lien
;  APPLICANT:  Verna, Ron
;  APPLICANT:  Wachter, Adam
;  APPLICANT:  Wu, James
;  APPLICANT:  Yim, Kenneth
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 740CIP
;  CURRENT APPLICATION NUMBER: US/09/975,640A
;  CURRENT FILING DATE:  2001-10-11
;  PRIOR APPLICATION NUMBER: 09/733,811
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/088,041
;  PRIOR FILING DATE: 1998-06-02
;  NUMBER OF SEQ ID NOS: 31906
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 11519
;   LENGTH: 439
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(439)
;   OTHER INFORMATION: n = A,T,C or G
US-09-975-640A-11519

Alignment Scores:
Pred. No.:                3.74e-63        Length:          439
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       38              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-975-640A-11519 (1-439)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
```

```
Db        410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 20
US-08-905-144-190
; Sequence 190, Application US/08905144
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Jean-Baptiste D.
;    APPLICANT:  Duelert, Aymeric
;    APPLICANT:  Lacroix, Bruno
;    TITLE OF INVENTION:  5' ESTs FOR SECRETED PROTEINS
;    TITLE OF INVENTION:  EXPRESSED IN PROSTATE
;    NUMBER OF SEQUENCES:  593
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  501 West Broadway
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-3505
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Win95
;      SOFTWARE:  Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/905,144
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned A.
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  GENSET.010A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  190:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  483 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  DOUBLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  CDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo Sapiens
;      TISSUE TYPE:  Normal prostate
;    FEATURE:
;      NAME/KEY:  other
;      LOCATION:  105..414
;      IDENTIFICATION METHOD:  blastn
;      OTHER INFORMATION:  identity 99
;      OTHER INFORMATION:  region 1..310
;      OTHER INFORMATION:  id T26956
;      OTHER INFORMATION:  est
;    FEATURE:
;      NAME/KEY:  other
;      LOCATION:  45..359
;      IDENTIFICATION METHOD:  blastn
;      OTHER INFORMATION:  identity 98
;      OTHER INFORMATION:  region 1..315
;      OTHER INFORMATION:  id T31666
;      OTHER INFORMATION:  est
;    FEATURE:
;      NAME/KEY:  other
;      LOCATION:  202..332
;      IDENTIFICATION METHOD:  blastn
;      OTHER INFORMATION:  identity 100
;      OTHER INFORMATION:  region 137..267
;      OTHER INFORMATION:  id R14990
;      OTHER INFORMATION:  est
;    FEATURE:
;      NAME/KEY:  other
;      LOCATION:  127..201
;      IDENTIFICATION METHOD:  blastn
;      OTHER INFORMATION:  identity 100
;      OTHER INFORMATION:  region 63..137
;      OTHER INFORMATION:  id R14990
;      OTHER INFORMATION:  est
;    FEATURE:
;      NAME/KEY:  other
;      LOCATION:  65..114
;      IDENTIFICATION METHOD:  blastn
;      OTHER INFORMATION:  identity 100
;      OTHER INFORMATION:  region 1..50
;      OTHER INFORMATION:  id R14990
;      OTHER INFORMATION:  est
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  1..120
;      IDENTIFICATION METHOD:  Von Heijne matrix
;      OTHER INFORMATION:  score 5.2
;      OTHER INFORMATION:  seq LCVEFASVASCDA/AV
US-08-905-144-190

Alignment Scores:
Pred. No.:                4.39e-63        Length:          483
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       19              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-08-905-144-190 (1-483)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGTTCCTACTTA 420

RESULT 21
US-09-547-599C-825
; Sequence 825, Application US/09547599C
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Duclert, Aymeric
;  APPLICANT:  Lacroix, Bruno
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
;  FILE REFERENCE: GEN-T119C1
;  CURRENT APPLICATION NUMBER: US/09/547,599C
;  CURRENT FILING DATE:  2000-04-11
;  PRIOR APPLICATION NUMBER: US 08/905,223
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,135
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,051
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,144
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,279
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/904,468
;  PRIOR FILING DATE: 1997-08-01
```

```
;  PRIOR APPLICATION NUMBER: US 08/905,134
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,133
;  PRIOR FILING DATE: 1997-08-01
;  NUMBER OF SEQ ID NOS: 3475
;  SOFTWARE: Patent.pm
; SEQ ID NO 825
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: Normal prostate
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 1..120
;   OTHER INFORMATION: Von Heijne matrix
;   OTHER INFORMATION: score 5.20
;   FEATURE:
;   OTHER INFORMATION: seq LCVEFASVASCDA/AV
;   NAME/KEY: misc_feature
;   LOCATION: 105..414
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 99
;   OTHER INFORMATION: region 1..310
;   OTHER INFORMATION: id T26956
;   FEATURE:
;   OTHER INFORMATION: est
;   NAME/KEY: misc_feature
;   LOCATION: 45..359
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 98
;   OTHER INFORMATION: region 1..315
;   OTHER INFORMATION: id T31666
;   FEATURE:
;   OTHER INFORMATION: est
;   NAME/KEY: misc_feature
;   LOCATION: 202..332
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 137..267
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   OTHER INFORMATION: est
;   NAME/KEY: misc_feature
;   LOCATION: 127..201
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 63..137
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   OTHER INFORMATION: est
;   NAME/KEY: misc_feature
;   LOCATION: 65..114
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 1..50
;   OTHER INFORMATION: id R14990
;   OTHER INFORMATION: est
US-09-547-599C-825

Alignment Scores:
Pred. No.:                4.39e-63        Length:          483
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       27              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-547-599C-825 (1-483)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGTTCCTACTTA 420

RESULT 22
US-10-631-402-825
; Sequence 825, Application US/10631402
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Duclert, Aymeric
;  APPLICANT:  Lacroix, Bruno
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
;  FILE REFERENCE: GEN-T119C1
;  CURRENT APPLICATION NUMBER: US/10/631,402
;  CURRENT FILING DATE:  2003-07-30
;  PRIOR APPLICATION NUMBER: US/09/547,599C
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 08/905,223
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,135
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,051
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,144
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,279
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/904,468
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,134
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,133
;  PRIOR FILING DATE: 1997-08-01
;  NUMBER OF SEQ ID NOS: 3475
;  SOFTWARE: Patent.pm
; SEQ ID NO 825
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: Normal prostate
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 1..120
;   OTHER INFORMATION: Von Heijne matrix
;   FEATURE:
;   OTHER INFORMATION: score 5.20
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 105..414
;   OTHER INFORMATION: seq LCVEFASVASCDA/AV
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 99
;   OTHER INFORMATION: region 1..310
;   FEATURE:
;   OTHER INFORMATION: id T26956
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 45..359
;   OTHER INFORMATION: est
```

```
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 98
;   OTHER INFORMATION: region 1..315
;   FEATURE:
;   OTHER INFORMATION: id T31666
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 202..332
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 137..267
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 127..201
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 63..137
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 65..114
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 1..50
;   OTHER INFORMATION: id R14990
;   OTHER INFORMATION: est
US-10-631-402-825

Alignment Scores:
Pred. No.:                4.39e-63        Length:          483
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       54              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-631-402-825 (1-483)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGTTCCTACTTA 420

RESULT 23
US-10-631-441-825
; Sequence 825, Application US/10631441
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Duclert, Aymeric
;  APPLICANT:  Lacroix, Bruno
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
;  FILE REFERENCE: GEN-T119C1
;  CURRENT APPLICATION NUMBER: US/10/631,441
;  CURRENT FILING DATE:  2003-07-30
;  PRIOR APPLICATION NUMBER: US/09/547,599C
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 08/905,223
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,135
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,051
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,144
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,279
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/904,468
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,134
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,133
;  PRIOR FILING DATE: 1997-08-01
;  NUMBER OF SEQ ID NOS: 3475
;  SOFTWARE: Patent.pm
; SEQ ID NO 825
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: Normal prostate
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 1..120
;   OTHER INFORMATION: Von Heijne matrix
;   FEATURE:
;   OTHER INFORMATION: score 5.20
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 105..414
;   OTHER INFORMATION: seq LCVEFASVASCDA/AV
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 99
;   OTHER INFORMATION: region 1..310
;   FEATURE:
;   OTHER INFORMATION: id T26956
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 45..359
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 98
;   OTHER INFORMATION: region 1..315
;   FEATURE:
;   OTHER INFORMATION: id T31666
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 202..332
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 137..267
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 127..201
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 63..137
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 65..114
```

```
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 1..50
;   OTHER INFORMATION: id R14990
;   OTHER INFORMATION: est
US-10-631-441-825

Alignment Scores:
Pred. No.:              4.39e-63        Length:         483
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     54              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-631-441-825 (1-483)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGTTCCTACTTA 420

RESULT 24
US-10-631-441A-825
; Sequence 825, Application US/10631441A
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Duclert, Aymeric
;  APPLICANT:  Lacroix, Bruno
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
;  FILE REFERENCE: GEN-T119C1
;  CURRENT APPLICATION NUMBER: US/10/631,441A
;  CURRENT FILING DATE:  2003-07-30
;  PRIOR APPLICATION NUMBER: US/09/547,599C
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 08/905,223
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,135
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,051
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,144
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,279
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/904,468
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,134
;  PRIOR FILING DATE: 1997-08-01
;  PRIOR APPLICATION NUMBER: US 08/905,133
;  PRIOR FILING DATE: 1997-08-01
;  NUMBER OF SEQ ID NOS: 3475
;  SOFTWARE: Patent.pm
; SEQ ID NO 825
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: Normal prostate
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: 1..120
;   OTHER INFORMATION: Von Heijne matrix
;   FEATURE:
;   OTHER INFORMATION: score 5.20
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 105..414
;   OTHER INFORMATION: seq LCVEFASVASCDA/AV
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 99
;   OTHER INFORMATION: region 1..310
;   FEATURE:
;   OTHER INFORMATION: id T26956
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 45..359
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 98
;   OTHER INFORMATION: region 1..315
;   FEATURE:
;   OTHER INFORMATION: id T31666
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 202..332
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 137..267
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 127..201
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 63..137
;   FEATURE:
;   OTHER INFORMATION: id R14990
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 65..114
;   OTHER INFORMATION: est
;   OTHER INFORMATION: blastn
;   OTHER INFORMATION: identity 100
;   OTHER INFORMATION: region 1..50
;   OTHER INFORMATION: id R14990
;   OTHER INFORMATION: est
US-10-631-441A-825

Alignment Scores:
Pred. No.:              4.39e-63        Length:         483
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     54              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-631-441A-825 (1-483)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279
```

```
Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         400 CGAGGGGTGTGTTCCTACTTA 420

RESULT 25
US-09-652-914-4701
; Sequence 4701, Application US/09652914
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1193-001
;  CURRENT APPLICATION NUMBER: US/09/652,914
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER:  60/152,112
;  PRIOR FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  9677
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 4701
;   LENGTH: 501
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(501)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-914-4701

Alignment Scores:
Pred. No.:                4.67e-63        Length:         501
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       29              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-652-914-4701 (1-501)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 26
US-09-205-070-10302
; Sequence 10302, Application US/09205070
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
;  TITLE OF INVENTION: LIBRARIES
;  FILE REFERENCE: 20411-748
;  CURRENT APPLICATION NUMBER: US/09/205,070
;  CURRENT FILING DATE: 1998-12-03
;  NUMBER OF SEQ ID NOS: 45207
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10302
;   LENGTH: 505
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-205-070-10302

Alignment Scores:
Pred. No.:                4.74e-63        Length:         505
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       22              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-205-070-10302 (1-505)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 27
US-09-340-623-10302
; Sequence 10302, Application US/09340623
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
;  TITLE OF INVENTION: LIBRARIES
;  FILE REFERENCE: 20411-748CON1
;  CURRENT APPLICATION NUMBER: US/09/340,623
;  CURRENT FILING DATE: 1999-06-28
;  EARLIER APPLICATION NUMBER: US 09/205,070
;  EARLIER FILING DATE: 1998-12-03
;  NUMBER OF SEQ ID NOS: 45207
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10302
;   LENGTH: 505
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-340-623-10302

Alignment Scores:
Pred. No.:                4.74e-63        Length:         505
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       23              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-340-623-10302 (1-505)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 28
US-09-898-888-10302
; Sequence 10302, Application US/09898888
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
;  TITLE OF INVENTION:  LIBRARIES
;  FILE REFERENCE: 20411-748CON1
;  CURRENT APPLICATION NUMBER: US/09/898,888
;  CURRENT FILING DATE:  2001-07-03
;  PRIOR APPLICATION NUMBER: 09/340,623
;  PRIOR FILING DATE: 1999-06-28
;  NUMBER OF SEQ ID NOS: 45207
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10302
;   LENGTH: 505
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-898-888-10302

Alignment Scores:
Pred. No.:                4.74e-63        Length:          505
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       34              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-898-888-10302 (1-505)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 29
US-09-898-888A-10302
; Sequence 10302, Application US/09898888A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
;  TITLE OF INVENTION:  LIBRARIES
;  FILE REFERENCE: 20411-748CON1
;  CURRENT APPLICATION NUMBER: US/09/898,888A
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/340,623
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: US 09/205,070
;  PRIOR FILING DATE: 1998-12-03
;  NUMBER OF SEQ ID NOS: 45207
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10302
;   LENGTH: 505
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-898-888A-10302

Alignment Scores:
Pred. No.:                4.74e-63        Length:          505
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       34              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-898-888A-10302 (1-505)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 30
US-10-302-689A-10302
; Sequence 10302, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Asundi, Vinod
;  APPLICANT: Ballinger, Dennis
;  APPLICANT: Labat, Ivan
;  APPLICANT: Leshkowitz, Dena
;  APPLICANT: Liu, Jin
;  APPLICANT: Loeb, Deborah
;  APPLICANT: Montgomery, Julia, R.
;  APPLICANT: Pace, Ann M.
;  APPLICANT: Sheridan, James P.
;  APPLICANT: Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
```

```
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 10302
;   LENGTH: 505
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-302-689A-10302

Alignment Scores:
Pred. No.:              4.74e-63        Length:         505
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     48              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-302-689A-10302 (1-505)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 289

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 349

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 409

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        410 CGAGGGGTGTGTTCCTACTTA 430

RESULT 31
US-09-726-802-1182
; Sequence 1182, Application US/09726802
; GENERAL INFORMATION:
;  APPLICANT:  Acton, Susan L.
;  APPLICANT:  Palb, Dean A.
;  APPLICANT:  .
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2025-001
;  CURRENT APPLICATION NUMBER: US/09/726,802
;  CURRENT FILING DATE:  2000-11-30
;  PRIOR APPLICATION NUMBER:  60/168,012
;  PRIOR FILING DATE:  1999-11-30
;  NUMBER OF SEQ ID NOS:  2872
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 1182
;   LENGTH: 517
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(517)
;   OTHER INFORMATION: n = A,T,C or G
US-09-726-802-1182

Alignment Scores:
Pred. No.:              4.93e-63        Length:         517
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-726-802-1182 (1-517)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 254

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 314

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 374

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 434

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        435 CGAGGGGTGTGTTCCTACTTA 455

RESULT 32
PCT-US00-05883-170
; Sequence 170, Application PC/TUS0005883
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT: Steve Ruben
;  TITLE OF INVENTION: Human Colon Cancer Associated Gene Sequences and Polypeptides
;  FILE REFERENCE: PA102PCT
;  CURRENT APPLICATION NUMBER: PCT/US00/05883
;  CURRENT FILING DATE: 2000-03-08
;  EARLIER APPLICATION NUMBER: 60/124,270
;  EARLIER FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
PCT-US00-05883-170

Alignment Scores:
Pred. No.:              2.33e-62        Length:         1296
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x PCT-US00-05883-170 (1-1296)
```

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 242

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 302

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 362

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 422

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        423 CGAGGGGTGTGTTCCTACTTA 443
```

RESULT 33
US-09-925-299-170
; Sequence 170, Application US/09925299
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1261)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (1276)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:

Pred. No.:	2.33e-62	Length:	1296
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	36	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-299-170 (1-1296)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 242

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 302

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 362

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 422

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        423 CGAGGGGTGTGTTCCTACTTA 443
```

RESULT 34
US-10-170-235-26834/c
; Sequence 26834, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 26834
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-26834

Alignment Scores:

Pred. No.:	4.26e-62	Length:	1849
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	42	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-170-235-26834 (1-1849)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        795 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 736

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        735 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 676

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 616

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 556

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        555 CGAGGGGTGTGTTCCTACTTA 535
```

RESULT 35
US-09-644-868-7772
; Sequence 7772, Application US/09644868
; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1168-001
; CURRENT APPLICATION NUMBER: US/09/644,868
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,063
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 10075
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7772
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-644-868-7772

Alignment Scores:
Pred. No.:              4.27e-62        Length:         1852
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     29              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-644-868-7772 (1-1852)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 251

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 311

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 371

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 431

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        432 CGAGGGGTGTGTTCCTACTTA 452

RESULT 36
US-09-649-161-8751
; Sequence 8751, Application US/09649161
; GENERAL INFORMATION:
;  APPLICANT:  Gutierrez-Ramos, Jose-Carlos
;  APPLICANT:  Villeval, Jean-Luc M.G.
;  APPLICANT:  Fraser, Christopher C.
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Hodge, Martin R.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1178-001
;  CURRENT APPLICATION NUMBER: US/09/649,161
;  CURRENT FILING DATE:  2000-08-28
;  PRIOR APPLICATION NUMBER:  60/151,060
;  PRIOR FILING DATE:  1999-08-27
;  NUMBER OF SEQ ID NOS:  10788
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8751
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-649-161-8751

Alignment Scores:
Pred. No.:              4.27e-62        Length:         1852
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     29              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-649-161-8751 (1-1852)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 251

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 311

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 371

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 431

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        432 CGAGGGGTGTGTTCCTACTTA 452

RESULT 37
US-09-652-109-8165
; Sequence 8165, Application US/09652109
; GENERAL INFORMATION:
;  APPLICANT:  McCarthy, Sean A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1180-001
;  CURRENT APPLICATION NUMBER: US/09/652,109
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,128
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  10105
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8165
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-109-8165

Alignment Scores:
Pred. No.:              4.27e-62        Length:         1852
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     29              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-652-109-8165 (1-1852)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 251

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 311

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 371

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 431

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        432 CGAGGGGTGTGTTCCTACTTA 452
```

```
RESULT 38
US-09-652-128-7441
; Sequence 7441, Application US/09652128
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  APPLICANT:  Holtzman, Douglas A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1171-001
;  CURRENT APPLICATION NUMBER: US/09/652,128
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,133
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  10265
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 7441
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-128-7441

Alignment Scores:
Pred. No.:                4.27e-62        Length:          1852
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       29              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-652-128-7441 (1-1852)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 251

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 311

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 371

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 431

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        432 CGAGGGGTGTGTTCCTACTTA 452

RESULT 39
US-09-652-816-8491
; Sequence 8491, Application US/09652816
; GENERAL INFORMATION:
;  APPLICANT:  Gutierrez-Ramos, Jose-Carlos
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1177-001
;  CURRENT APPLICATION NUMBER: US/09/652,816
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER:  60/152,111
;  PRIOR FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  9647
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8491
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-816-8491

Alignment Scores:
Pred. No.:                4.27e-62        Length:          1852
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       29              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-652-816-8491 (1-1852)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 251

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 311

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 371

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 431

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        432 CGAGGGGTGTGTTCCTACTTA 452

RESULT 40
US-09-652-914-8450
; Sequence 8450, Application US/09652914
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1193-001
;  CURRENT APPLICATION NUMBER: US/09/652,914
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER:  60/152,112
;  PRIOR FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  9677
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 8450
;   LENGTH: 1852
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1852)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-914-8450

Alignment Scores:
Pred. No.:                4.27e-62        Length:          1852
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       29              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-652-914-8450 (1-1852)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        192 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 251

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 311

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 371

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 431

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        432 CGAGGGGTGTGTTCCTACTTA 452

RESULT 41
US-10-170-235-27345
; Sequence 27345, Application US/10170235
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig
;  TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
;  TITLE OF INVENTION:  TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
;  FILE REFERENCE: CL001380
;  CURRENT APPLICATION NUMBER: US/10/170,235
;  CURRENT FILING DATE:  2003-03-17
;  NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 27345
;   LENGTH: 1919
;   TYPE: DNA
;   ORGANISM: HUMAN
US-10-170-235-27345

Alignment Scores:
Pred. No.:                4.53e-62        Length:         1919
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       42              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-170-235-27345 (1-1919)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 236

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 296

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 356

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 416

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        417 CGAGGGGTGTGTTCCTACTTA 437

RESULT 42
US-10-757-745-1
; Sequence 1, Application US/10757745
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1627)..(1627)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1108)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1849)..(1849)
;   OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

Alignment Scores:
Pred. No.:                4.54e-62        Length:         1920
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       61              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-757-745-1 (1-1920)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 238

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 298

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 358

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 418

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        419 CGAGGGGTGTGTTCCTACTTA 439

RESULT 43
PCT-US04-24424-921
; Sequence 921, Application PC/TUS0424424
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PCT
;  CURRENT APPLICATION NUMBER: PCT/US04/24424
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US 60/490,890
;  PRIOR FILING DATE: 2003-07-29
;  NUMBER OF SEQ ID NOS: 2786
```

```
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US04-24424-921

Alignment Scores:
Pred. No.:                4.62e-62        Length:          1940
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x PCT-US04-24424-921 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 44
US-10-783-271-28
; Sequence 28, Application US/10783271
; GENERAL INFORMATION:
;  APPLICANT: Veridex, LLC
;  APPLICANT:  Wang, Yixin
;  TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
;  FILE REFERENCE: VDX-5003 USNP
;  CURRENT APPLICATION NUMBER: US/10/783,271
;  CURRENT FILING DATE:  2004-02-20
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: human
US-10-783-271-28

Alignment Scores:
Pred. No.:                4.62e-62        Length:          1940
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       61              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-783-271-28 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 45
US-60-490-890-921
; Sequence 921, Application US/60490890
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PSP
;  CURRENT APPLICATION NUMBER: US/60/490,890
;  CURRENT FILING DATE:  2003-07-29
;  NUMBER OF SEQ ID NOS: 2779
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-490-890-921

Alignment Scores:
Pred. No.:                4.62e-62        Length:          1940
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       80              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-60-490-890-921 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 46
US-60-568-073-582
; Sequence 582, Application US/60568073
; GENERAL INFORMATION:
;  APPLICANT: Huang, Fei
;  APPLICANT:  Han, Xia
;  APPLICANT:  Shaw, Peter
;  APPLICANT:  Clark, Edwin
;  TITLE OF INVENTION: COLON CANCER BIOMARKERS
;  FILE REFERENCE: 10205 PSP
```

```
;  CURRENT APPLICATION NUMBER: US/60/568,073
;  CURRENT FILING DATE:  2004-05-04
;  NUMBER OF SEQ ID NOS: 1305
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 582
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-568-073-582

Alignment Scores:
Pred. No.:              4.62e-62      Length:         1940
Score:                  448.00        Matches:        87
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     81            Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-60-568-073-582 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 47
US-60-671-501-156
; Sequence 156, Application US/60671501
; GENERAL INFORMATION:
;  APPLICANT: Wim Van Criekinge
;  APPLICANT:  Josef Straub
;  TITLE OF INVENTION: METHYLATION MARKERS FOR DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF CANCERS
;  FILE REFERENCE: 006791.00001
;  CURRENT APPLICATION NUMBER: US/60/671,501
;  CURRENT FILING DATE:  2005-04-15
;  NUMBER OF SEQ ID NOS: 420
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 156
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-671-501-156

Alignment Scores:
Pred. No.:              4.62e-62      Length:         1940
Score:                  448.00        Matches:        87
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     82            Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-60-671-501-156 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 48
US-60-685-372-523
; Sequence 523, Application US/60685372
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John V
;  APPLICANT:  Bradley, Walter E
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall D
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul V
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: GeneMap of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 059908-5005-PR
;  CURRENT APPLICATION NUMBER: US/60/685,372
;  CURRENT FILING DATE:  2005-05-31
;  NUMBER OF SEQ ID NOS: 2738
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 523
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-685-372-523

Alignment Scores:
Pred. No.:              4.62e-62      Length:         1940
Score:                  448.00        Matches:        87
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     82            Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-60-685-372-523 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445
```

```
RESULT 49
US-60-704-740-28
; Sequence 28, Application US/60704740
; GENERAL INFORMATION:
;  APPLICANT: Wang, Yixin
;  APPLICANT:  Zhang, Yi
;  APPLICANT:  Atkins, David
;  APPLICANT:  Sieuwerts, Anieta M.
;  APPLICANT:  Smid, Marcel
;  APPLICANT:  Klijn, Jan G.M.
;  APPLICANT:  Martens, John W.M.
;  APPLICANT:  Foekens, John A.
;  TITLE OF INVENTION: Predicting Bone Relapse of Breast Cancer
;  FILE REFERENCE: VDX 5026 USPSP
;  CURRENT APPLICATION NUMBER: US/60/704,740
;  CURRENT FILING DATE:  2005-08-02
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: human
US-60-704-740-28

Alignment Scores:
Pred. No.:                4.62e-62        Length:         1940
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       83              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-60-704-740-28 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 50
US-60-710-726-741
; Sequence 741, Application US/60710726
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John V
;  APPLICANT:  Bradley, Walter E
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Fournier, Helene
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Serre, Valerie
;  APPLICANT:  Eerdewegh, Paul V
;  APPLICANT:  Little, Randall D
;  APPLICANT:  Keith , Tim
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: GeneMap of the human genes associated with Crohn's disease
;  FILE REFERENCE: 059908-5009-PR
;  CURRENT APPLICATION NUMBER: US/60/710,726
;  CURRENT FILING DATE:  2005-08-24
;  NUMBER OF SEQ ID NOS: 20166
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 741
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homosapiens
US-60-710-726-741

Alignment Scores:
Pred. No.:                4.62e-62        Length:         1940
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       83              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-60-710-726-741 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 51
US-09-488-725A-1200
; Sequence 1200, Application US/09488725A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/09/488,725A
;  CURRENT FILING DATE:  2000-12-22
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7144
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
```

```
;   LOCATION: (39)..(1127)
US-09-488-725A-1200

Alignment Scores:
Pred. No.:                4.65e-62        Length:         1948
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       24              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-488-725A-1200 (1-1948)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 52
US-10-037-270-889
; Sequence 889, Application US/10037270
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/10/037,270
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-037-270-889

Alignment Scores:
Pred. No.:                4.65e-62        Length:         1948
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       40              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-037-270-889 (1-1948)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 53
US-10-105-837-889
; Sequence 889, Application US/10105837
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BDIVA
;  CURRENT APPLICATION NUMBER: US/10/105,837
;  CURRENT FILING DATE:  2002-03-21
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-105-837-889

Alignment Scores:
Pred. No.:                4.65e-62        Length:         1948
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       42              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-105-837-889 (1-1948)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
```

```
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40

Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60

Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80

Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87

Db        438 CGAGGGGTGTGTTCCTACTTA 458
```

RESULT 54
US-10-117-722-889
; Sequence 889, Application US/10117722
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-117-722-889

Alignment Scores:

Pred. No.:	4.65e-62	Length:	1948
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	42	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-117-722-889 (1-1948)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20

Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40

Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60

Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80

Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87

Db        438 CGAGGGGTGTGTTCCTACTTA 458
```

RESULT 55
US-10-122-851-889
; Sequence 889, Application US/10122851
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BDV3
; CURRENT APPLICATION NUMBER: US/10/122,851
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-122-851-889

Alignment Scores:

Pred. No.:	4.65e-62	Length:	1948
Score:	448.00	Matches:	87
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	42	Gaps:	0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-122-851-889 (1-1948)

```
Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20

Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40

Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60

Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80

Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87

Db        438 CGAGGGGTGTGTTCCTACTTA 458
```

RESULT 56
US-10-258-898A-1200
; Sequence 1200, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc

```
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/258,898A
;  CURRENT FILING DATE:  2002-10-29
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-258-898A-1200

Alignment Scores:
Pred. No.:                4.65e-62        Length:          1948
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       43              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-258-898A-1200 (1-1948)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 57
US-10-286-897-1200
; Sequence 1200, Application US/10286897
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/286,897
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US/09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US/09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US/09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US/09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US/09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US/09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US/09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1200
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-286-897-1200

Alignment Scores:
Pred. No.:                4.65e-62        Length:          1948
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       43              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-286-897-1200 (1-1948)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 257

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 377

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 437

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        438 CGAGGGGTGTGTTCCTACTTA 458

RESULT 58
US-09-471-275-1314
; Sequence 1314, Application US/09471275
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: Novel Contigs Obtained
;  TITLE OF INVENTION: From Various Libraries
;  FILE REFERENCE: 782
;  CURRENT APPLICATION NUMBER: US/09/471,275
;  CURRENT FILING DATE: 1999-12-23
;  EARLIER APPLICATION NUMBER: US 09/235,076
;  EARLIER FILING DATE: 1999-01-20
;  EARLIER APPLICATION NUMBER: US 09/234,611
;  EARLIER FILING DATE: 1999-01-22
;  EARLIER APPLICATION NUMBER: US 09/240,371
;  EARLIER FILING DATE: 1999-01-29
;  EARLIER APPLICATION NUMBER: US 09/277,227
;  EARLIER FILING DATE: 1999-03-25
;  EARLIER APPLICATION NUMBER: US 09/271,490
```

```
;  EARLIER FILING DATE: 1999-03-18
;  EARLIER APPLICATION NUMBER: US 09/293,972
;  EARLIER FILING DATE: 1999-04-15
;  EARLIER APPLICATION NUMBER: US 09/274,861
;  EARLIER FILING DATE: 1999-03-23
;  EARLIER APPLICATION NUMBER: US 60/125,453
;  EARLIER FILING DATE: 1999-03-19
;  EARLIER APPLICATION NUMBER: US 60/126,605
;  EARLIER FILING DATE: 1999-03-26
;  EARLIER APPLICATION NUMBER: US 09/306,350
;  EARLIER FILING DATE: 1999-05-07
;  EARLIER APPLICATION NUMBER: US 09/399,720
;  EARLIER FILING DATE: 1999-09-21
;  EARLIER APPLICATION NUMBER: US 09/404,284
;  EARLIER FILING DATE: 1999-09-21
;  EARLIER APPLICATION NUMBER:
;  EARLIER FILING DATE: 1999-12-16
;  NUMBER OF SEQ ID NOS: 10451
;  SOFTWARE: pt_CT_genes Version 1.0
; SEQ ID NO 1314
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-471-275-1314

Alignment Scores:
Pred. No.:                4.69e-62        Length:         1958
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       24              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-471-275-1314 (1-1958)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 269

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 329

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 389

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 449

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        450 CGAGGGGTGTGTTCCTACTTA 470

RESULT 59
US-09-488-725A-4772
; Sequence 4772, Application US/09488725A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/09/488,725A
;  CURRENT FILING DATE:  2000-12-22
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7144
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-488-725A-4772

Alignment Scores:
Pred. No.:                4.69e-62        Length:         1958
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       24              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-488-725A-4772 (1-1958)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 269

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 329

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 389

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 449

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        450 CGAGGGGTGTGTTCCTACTTA 470

RESULT 60
US-09-488-725B-8295
; Sequence 8295, Application US/09488725B
; GENERAL INFORMATION:
;  APPLICANT: Yuanhua T. Tang
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Ankura Sinku
;  APPLICANT:  Chenghua Liu
;  APPLICANT:  Radoje T. Drmanac
;  TITLE OF INVENTION: Novel Contigs Obtained
;  TITLE OF INVENTION:  From Various Libraries
;  FILE REFERENCE: 784
;  CURRENT APPLICATION NUMBER: US/09/488,725B
;  CURRENT FILING DATE:  2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/004,182
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: US 09/034,341
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 09/045,400
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: US 09/321,214
;  PRIOR FILING DATE: 1999-05-26
;  PRIOR APPLICATION NUMBER: US 09/131,598
```

```
;  PRIOR FILING DATE: 1998-08-10
;  PRIOR APPLICATION NUMBER: US 09/170,294
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: US 09/179,473
;  PRIOR FILING DATE: 1998-10-27
;  PRIOR APPLICATION NUMBER: US 09/181,430
;  PRIOR FILING DATE: 1998-10-28
;  PRIOR APPLICATION NUMBER: US 09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  PRIOR APPLICATION NUMBER: US 09/234,611
;  PRIOR FILING DATE: 1999-01-22
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10289
;  SOFTWARE: pt_CT_genes Version 1.01
; SEQ ID NO 8295
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-488-725B-8295

Alignment Scores:
Pred. No.:              4.69e-62      Length:          1958
Score:                  448.00        Matches:         87
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     24            Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-488-725B-8295 (1-1958)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 269

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 329

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 389

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 449

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        450 CGAGGGGTGTGTTCCTACTTA 470

RESULT 61
US-09-552-317-8295
; Sequence 8295, Application US/09552317
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Tillinghast, John
;  APPLICANT:  Sinku, Ankura
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids
;  TITLE OF INVENTION:  and Polypeptides
;  FILE REFERENCE: 784CIP
;  CURRENT APPLICATION NUMBER: US/09/552,317
;  CURRENT FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 10289
;  SOFTWARE: pt_CT_genes Version 1.01
; SEQ ID NO 8295
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)...(1139)
;   OTHER INFORMATION: similar to gi4210353 in the genepept database release 114,
;   OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-552-317-8295

Alignment Scores:
Pred. No.:              4.69e-62      Length:          1958
Score:                  448.00        Matches:         87
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     27            Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-552-317-8295 (1-1958)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 269

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 329

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 389

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 449

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        450 CGAGGGGTGTGTTCCTACTTA 470

RESULT 62
US-10-258-898A-4772
; Sequence 4772, Application US/10258898A
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/258,898A
;  CURRENT FILING DATE:  2002-10-29
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-258-898A-4772
```

```
Alignment Scores:
Pred. No.:                4.69e-62        Length:          1958
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       43              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-258-898A-4772 (1-1958)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 269

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 329

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 389

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 449

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        450 CGAGGGGTGTGTTCCTACTTA 470

RESULT 63
US-10-286-897-4772
; Sequence 4772, Application US/10286897
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
;  FILE REFERENCE: 784FLPCT
;  CURRENT APPLICATION NUMBER: US/10/286,897
;  CURRENT FILING DATE:  2002-11-01
;  PRIOR APPLICATION NUMBER: US/09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US/09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US/09/598,042
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: US/09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US/09/653,450
;  PRIOR FILING DATE: 2000-08-31
;  PRIOR APPLICATION NUMBER: US/09/662,191
;  PRIOR FILING DATE: 2000-09-14
;  PRIOR APPLICATION NUMBER: US/09/693,036
;  PRIOR FILING DATE: 2000-10-19
;  PRIOR APPLICATION NUMBER: US/09/727,344
;  PRIOR FILING DATE: 2000-11-29
;  NUMBER OF SEQ ID NOS: 7143
;  SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 4772
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-286-897-4772

Alignment Scores:
Pred. No.:                4.69e-62        Length:          1958
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       43              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-286-897-4772 (1-1958)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 269

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 329

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 389

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 449

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        450 CGAGGGGTGTGTTCCTACTTA 470

RESULT 64
US-10-302-689A-59957
; Sequence 59957, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ballinger, Dennis
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Liu, Jin
;  APPLICANT:  Loeb, Deborah
;  APPLICANT:  Montgomery, Julia, R.
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 59957
;   LENGTH: 1958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-302-689A-59957

Alignment Scores:
Pred. No.:                4.69e-62        Length:          1958
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
```

```
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     48              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-302-689A-59957 (1-1958)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         210 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 269

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         270 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 329

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         330 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 389

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         390 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 449

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         450 CGAGGGGTGTGTTCCTACTTA 470

RESULT 65
US-10-302-689A-77389
; Sequence 77389, Application US/10302689A
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ballinger, Dennis
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Liu, Jin
;  APPLICANT:  Loeb, Deborah
;  APPLICANT:  Montgomery, Julia, R.
;  APPLICANT:  Pace, Ann M.
;  APPLICANT:  Sheridan, James P.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 502CIP
;  CURRENT APPLICATION NUMBER: US/10/302,689A
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: 10/273,573
;  PRIOR FILING DATE: 2002-10-18
;  PRIOR APPLICATION NUMBER: 10/084,643
;  PRIOR FILING DATE: 2002-02-26
;  PRIOR APPLICATION NUMBER: 09/989,660
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 10/014,487
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: 09/952,981
;  PRIOR FILING DATE: 2001-09-14
;  PRIOR APPLICATION NUMBER: 09/922,279
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: 09/905,059
;  PRIOR FILING DATE: 2001-07-12
;  PRIOR APPLICATION NUMBER: 09/898,888
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 09/919,002
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 09/770,160
;  PRIOR FILING DATE: 2001-01-26
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 158931
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 77389
;    LENGTH: 1958
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-302-689A-77389

Alignment Scores:
Pred. No.:              4.69e-62        Length:         1958
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     48              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-302-689A-77389 (1-1958)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         210 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 269

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         270 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 329

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         330 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 389

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         390 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 449

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         450 CGAGGGGTGTGTTCCTACTTA 470

RESULT 66
US-09-606-680-2868
; Sequence 2868, Application US/09606680
; GENERAL INFORMATION:
;  APPLICANT: Lloyd, Clare M.
;  APPLICANT:  Williamson, Mark
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 1600.1131-001
;  CURRENT APPLICATION NUMBER: US/09/606,680
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER: 60/141,227
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 60/141,226
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 4394
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2868
;    LENGTH: 2262
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)...(2262)
;    OTHER INFORMATION: n = A,T,C or G
US-09-606-680-2868

Alignment Scores:
Pred. No.:              5.99e-62        Length:         2262
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     28              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-606-680-2868 (1-2262)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         474 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 533
```

```
Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 593

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 653

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 713

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        714 CGAGGGGTGTGTTCCTACTTA 734

RESULT 67
US-09-606-776-4085
; Sequence 4085, Application US/09606776
; GENERAL INFORMATION:
;  APPLICANT: Holtzman, Douglas A.
;  APPLICANT:  Myers, Paul
;  APPLICANT:  Gearing, David P.
;  APPLICANT:  Pan, Yang
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 1600.1129-001
;  CURRENT APPLICATION NUMBER: US/09/606,776
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER: 60/141,578
;  PRIOR FILING DATE: 1999-06-29
;  PRIOR APPLICATION NUMBER: 60/141,379
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: 60/141,138
;  PRIOR FILING DATE: 1999-06-28
;  PRIOR APPLICATION NUMBER: 60/141,581
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 5415
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4085
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-606-776-4085

Alignment Scores:
Pred. No.:                5.99e-62        Length:          2262
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       28              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-606-776-4085 (1-2262)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 533

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 593

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 653

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 713

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        714 CGAGGGGTGTGTTCCTACTTA 734

RESULT 68
US-09-644-869-9498
; Sequence 9498, Application US/09644869
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  APPLICANT:  McCarthy, Sean A.
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Monahan, John
;  APPLICANT:  Richardson, Jennifer
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1182-001
;  CURRENT APPLICATION NUMBER: US/09/644,869
;  CURRENT FILING DATE:  2000-08-28
;  PRIOR APPLICATION NUMBER:  60/151,062
;  PRIOR FILING DATE:  1999-08-27
;  NUMBER OF SEQ ID NOS:  9708
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9498
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-644-869-9498

Alignment Scores:
Pred. No.:                5.99e-62        Length:          2262
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       29              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-644-869-9498 (1-2262)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 533

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 593

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 653

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 713

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        714 CGAGGGGTGTGTTCCTACTTA 734

RESULT 69
US-09-652-121-6794
; Sequence 6794, Application US/09652121
; GENERAL INFORMATION:
;  APPLICANT:  DiStefano, Peter
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1188-001
;  CURRENT APPLICATION NUMBER: US/09/652,121
```

```
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,129
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  7615
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6794
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-121-6794

Alignment Scores:
Pred. No.:                5.99e-62        Length:         2262
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       29              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-652-121-6794 (1-2262)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 533

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 593

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 653

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 713

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        714 CGAGGGGTGTGTTCCTACTTA 734

RESULT 70
US-09-652-122-3756
; Sequence 3756, Application US/09652122
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1194-001
;  CURRENT APPLICATION NUMBER: US/09/652,122
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,421
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  5020
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3756
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-122-3756

Alignment Scores:
Pred. No.:                5.99e-62        Length:         2262
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       29              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-652-122-3756 (1-2262)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 533

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 593

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 653

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 713

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        714 CGAGGGGTGTGTTCCTACTTA 734

RESULT 71
US-09-652-128-9329
; Sequence 9329, Application US/09652128
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  APPLICANT:  Holtzman, Douglas A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1171-001
;  CURRENT APPLICATION NUMBER: US/09/652,128
;  CURRENT FILING DATE:  2000-08-30
;  PRIOR APPLICATION NUMBER:  60/151,133
;  PRIOR FILING DATE:  1999-08-30
;  NUMBER OF SEQ ID NOS:  10265
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9329
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-128-9329

Alignment Scores:
Pred. No.:                5.99e-62        Length:         2262
Score:                    448.00          Matches:        87
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       29              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-652-128-9329 (1-2262)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 533

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 593

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 653
```

```
Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         654 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 713

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         714 CGAGGGGTGTGTTCCTACTTA 734

RESULT 72
US-09-652-816-9572
; Sequence 9572, Application US/09652816
; GENERAL INFORMATION:
;  APPLICANT:  Gutierrez-Ramos, Jose-Carlos
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.1177-001
;  CURRENT APPLICATION NUMBER: US/09/652,816
;  CURRENT FILING DATE:  2000-08-31
;  PRIOR APPLICATION NUMBER:  60/152,111
;  PRIOR FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  9647
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9572
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-652-816-9572

Alignment Scores:
Pred. No.:                5.99e-62        Length:          2262
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       29              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-652-816-9572 (1-2262)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         474 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 533

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         534 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 593

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         594 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 653

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         654 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 713

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         714 CGAGGGGTGTGTTCCTACTTA 734

RESULT 73
US-09-700-000-6419
; Sequence 6419, Application US/09700000
; GENERAL INFORMATION:
;  APPLICANT:  Richardson, Jennifer
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2022-001
;  CURRENT APPLICATION NUMBER: US/09/700,000
;  CURRENT FILING DATE:  2000-10-30
;  PRIOR APPLICATION NUMBER:  60/162,619
;  PRIOR FILING DATE:  1999-10-29
;  NUMBER OF SEQ ID NOS:  7171
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 6419
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-700-000-6419

Alignment Scores:
Pred. No.:                5.99e-62        Length:          2262
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       31              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-700-000-6419 (1-2262)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         474 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 533

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         534 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 593

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         594 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 653

Qy          61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         654 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 713

Qy          81 ArgGlyValCysSerTyrLeu 87
               |||||||||||||||||||||
Db         714 CGAGGGGTGTGTTCCTACTTA 734

RESULT 74
US-09-710-280-2045
; Sequence 2045, Application US/09710280
; GENERAL INFORMATION:
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Gearing, David P.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2007-001
;  CURRENT APPLICATION NUMBER: US/09/710,280
;  CURRENT FILING DATE:  2000-11-10
;  PRIOR APPLICATION NUMBER:  60/164,537
;  PRIOR FILING DATE:  1999-11-10
;  NUMBER OF SEQ ID NOS:  2763
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2045
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-710-280-2045

Alignment Scores:
Pred. No.:                5.99e-62        Length:          2262
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
```

```
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-710-280-2045 (1-2262)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 533

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 593

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 653

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 713

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        714 CGAGGGGTGTGTTCCTACTTA 734

RESULT 75
US-09-710-282-3926
; Sequence 3926, Application US/09710282
; GENERAL INFORMATION:
;  APPLICANT:  Gearing, David P.
;  APPLICANT:  Holtzman, Douglas A.
;  APPLICANT:  Kingsbury, Gillian A.
;  TITLE OF INVENTION:  NOVEL NUCLEIC ACID MOLECULES AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  1600.2004-001
;  CURRENT APPLICATION NUMBER: US/09/710,282
;  CURRENT FILING DATE:  2000-11-10
;  PRIOR APPLICATION NUMBER:  60/164,256
;  PRIOR FILING DATE:  1999-11-09
;  NUMBER OF SEQ ID NOS:  4662
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3926
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2262)
;   OTHER INFORMATION: n = A,T,C or G
US-09-710-282-3926

Alignment Scores:
Pred. No.:              5.99e-62        Length:         2262
Score:                  448.00          Matches:        87
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     31              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-710-282-3926 (1-2262)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 533

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 593

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 653

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 713

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        714 CGAGGGGTGTGTTCCTACTTA 734

Search completed: March 11, 2006, 12:06:37
Job time : 3588.7 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 23:22:00 ; Search time 203.552 Seconds
(without alignments)
1013.401 Million cell updates/sec

Title: US-10-757-745-2_COPY_54_140
Perfect score: 448
Sequence: 1 MERALNSYFEPPVEESALER.....IDGLDLNNLSERARGVCSYL 87

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5007642 seqs, 1185514400 residues

Total number of hits satisfying chosen parameters: 10015284

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170335_4211/app_query.fasta_1
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_240_@runat_09032006_170335_4211 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq1:*
9: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	448	100.0	1940	6	US-10-567-867-921	Sequence 921, App
	2	448	100.0	1940	11	US-60-742-219-2071	Sequence 2071, Ap
	3	440	98.2	1288	9	US-11-266-748A-226184	Sequence 226184,
	4	439	98.0	752	6	US-10-917-503B-4908	Sequence 4908, Ap
	5	439	98.0	1898	6	US-10-917-503B-13209	Sequence 13209, A
c	6	145	32.4	2300	11	US-60-762-056-35503	Sequence 35503, A
c	7	78	17.4	767	9	US-11-266-748A-72123	Sequence 72123, A
c	8	78	17.4	767	9	US-11-266-748A-107071	Sequence 107071,
	9	78	17.4	767	9	US-11-266-748A-124934	Sequence 124934,
	10	78	17.4	1548	9	US-11-266-748A-161851	Sequence 161851,
	11	78	17.4	1982	9	US-11-266-748A-184381	Sequence 184381,
	12	78	17.4	1982	9	US-11-266-748A-191888	Sequence 191888,
	13	78	17.4	2405	9	US-11-266-748A-184374	Sequence 184374,
	14	78	17.4	2405	9	US-11-266-748A-191881	Sequence 191881,
	15	78	17.4	2745	6	US-10-120-780-143	Sequence 143, App
	16	78	17.4	2792	9	US-11-266-748A-72116	Sequence 72116, A
	17	78	17.4	2792	9	US-11-266-748A-107064	Sequence 107064,
c	18	78	17.4	2792	9	US-11-266-748A-124927	Sequence 124927,
	19	78	17.4	4482	11	US-60-741-051-378	Sequence 378, App
	20	78	17.4	4482	11	US-60-741-051-379	Sequence 379, App
	21	78	17.4	6316	7	US-11-287-334-53	Sequence 53, Appl
	22	75.5	16.9	10506	1	PCT-US03-37356-1	Sequence 1, Appli
	23	75.5	16.9	10617	1	PCT-US03-37356-14	Sequence 14, Appl
	24	73.5	16.4	5713	6	US-10-556-637-6	Sequence 6, Appli
	25	71.5	16.0	933	6	US-10-513-024-863	Sequence 863, App
	26	71.5	16.0	936	10	US-11-348-413-1579	Sequence 1579, Ap
c	27	71	15.8	2944528	9	US-11-045-004-1	Sequence 1, Appli
	28	70	15.6	5048	6	US-10-567-867-2265	Sequence 2265, Ap
c	29	68.5	15.3	503	7	US-11-353-150-46076	Sequence 46076, A
c	30	68.5	15.3	1335	6	US-10-953-349-3481	Sequence 3481, Ap
	31	67.5	15.1	465	7	US-11-331-019-24356	Sequence 24356, A
	32	67.5	15.1	3594	9	US-11-266-748A-25375	Sequence 25375, A
	33	67.5	15.1	4148	9	US-11-266-748A-30209	Sequence 30209, A
	34	67.5	15.1	4148	11	US-60-751-420-2669	Sequence 2669, Ap
	35	67	15.0	399	7	US-11-331-019-39117	Sequence 39117, A
	36	67	15.0	400	7	US-11-331-019-38689	Sequence 38689, A
	37	67	15.0	424	7	US-11-331-019-37593	Sequence 37593, A
	38	67	15.0	442	7	US-11-331-019-39617	Sequence 39617, A
	39	67	15.0	453	7	US-11-331-019-38858	Sequence 38858, A
	40	67	15.0	1047	9	US-11-266-748A-349404	Sequence 349404,
	41	67	15.0	1047	9	US-11-266-748A-382318	Sequence 382318,
c	42	67	15.0	1047	9	US-11-266-748A-432783	Sequence 432783,
	43	67	15.0	1280	9	US-11-266-748A-349405	Sequence 349405,
	44	67	15.0	1280	9	US-11-266-748A-382319	Sequence 382319,
c	45	67	15.0	1280	9	US-11-266-748A-432784	Sequence 432784,
	46	67	15.0	1410	6	US-10-953-349-32751	Sequence 32751, A
	47	67	15.0	1500	11	US-60-748-114-12809	Sequence 12809, A
	48	67	15.0	1518	6	US-10-342-276B-20	Sequence 20, Appl
c	49	67	15.0	2300	11	US-60-762-056-44165	Sequence 44165, A
	50	67	15.0	2625	9	US-11-266-748A-349401	Sequence 349401,
	51	67	15.0	2625	9	US-11-266-748A-382315	Sequence 382315,
c	52	67	15.0	2625	9	US-11-266-748A-432780	Sequence 432780,
	53	67	15.0	2636	6	US-10-342-276B-21	Sequence 21, Appl
c	54	66.5	14.8	393	7	US-11-331-032-8894	Sequence 8894, Ap
c	55	66.5	14.8	445	7	US-11-331-032-7835	Sequence 7835, Ap
c	56	66.5	14.8	446	7	US-11-331-032-20623	Sequence 20623, A
c	57	66.5	14.8	570	7	US-11-331-032-3347	Sequence 3347, Ap
	58	66.5	14.8	7632	9	US-11-266-748A-31062	Sequence 31062, A
	59	66.5	14.8	2944528	9	US-11-045-004-1	Sequence 1, Appli
	60	66	14.7	371	7	US-11-331-032-2004	Sequence 2004, Ap
c	61	66	14.7	382	7	US-11-330-364-5748	Sequence 5748, Ap
	62	66	14.7	550	7	US-11-331-032-15190	Sequence 15190, A
	63	66	14.7	1632	6	US-10-461-673-3010	Sequence 3010, Ap
	64	66	14.7	1687	6	US-10-953-349-36708	Sequence 36708, A
	65	66	14.7	2032	4	US-08-765-108-5	Sequence 5, Appli
	66	66	14.7	2032	4	US-08-765-108-5	Sequence 5, Appli
	67	66	14.7	2032	4	US-08-765-108-5	Sequence 5, Appli
c	68	66	14.7	4775	9	US-11-266-748A-56767	Sequence 56767, A
c	69	66	14.7	4775	11	US-60-742-219-3003	Sequence 3003, Ap
c	70	65.5	14.6	401	7	US-11-331-032-22187	Sequence 22187, A
c	71	65.5	14.6	408	7	US-11-331-032-21617	Sequence 21617, A
c	72	65.5	14.6	485	7	US-11-331-032-11044	Sequence 11044, A
	73	65.5	14.6	855	6	US-10-917-503B-620	Sequence 620, App
	74	65.5	14.6	2082	6	US-10-917-503B-10820	Sequence 10820, A
	75	65.5	14.6	2211	6	US-10-558-119-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1

```
US-10-567-867-921
; Sequence 921, Application US/10567867
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PCT
;  CURRENT APPLICATION NUMBER: US/10/567,867
;  CURRENT FILING DATE:  2006-01-27
;  PRIOR APPLICATION NUMBER: US 60/490,890
;  PRIOR FILING DATE: 2003-07-29
;  NUMBER OF SEQ ID NOS: 2786
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 921
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-567-867-921

Alignment Scores:
Pred. No.:                1.84e-52        Length:          1940
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-567-867-921 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 2
US-60-742-219-2071
; Sequence 2071, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2071
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-2071

Alignment Scores:
Pred. No.:                1.84e-52        Length:          1940
Score:                    448.00          Matches:         87
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       11              Gaps:            0

US-10-757-745-2_COPY_54_140 (1-87) x US-60-742-219-2071 (1-1940)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 244

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 304

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 364

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 424

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        425 CGAGGGGTGTGTTCCTACTTA 445

RESULT 3
US-11-266-748A-226184
; Sequence 226184, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 226184
;   LENGTH: 1288
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-226184
```

```
Alignment Scores:
Pred. No.:              1.39e-51        Length:         1288
Score:                  440.00          Matches:        85
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  98.8%           Mismatches:     0
Query Match:            98.2%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-226184 (1-1288)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 74

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 134

Qy         42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTC 194

Qy         62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 254

Qy         82 GlyValCysSerTyrLeu 87
              |||||||||||||||||
Db        255 GGGGTGTGTTCCTACTTA 272

RESULT 4
US-10-917-503B-4908
; Sequence 4908, Application US/10917503B
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4908
;   LENGTH: 752
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (685)..(685)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (694)..(694)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (740)..(740)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-10-917-503B-4908

Alignment Scores:
Pred. No.:              9.04e-52        Length:         752
Score:                  439.00          Matches:        86
Percent Similarity:     98.9%           Conservative:   0
Best Local Similarity:  98.9%           Mismatches:     1
Query Match:            98.0%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-917-503B-4908 (1-752)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||
Db        400 CGAGGGGTGTGCTCCTACTTA 420

RESULT 5
US-10-917-503B-13209
; Sequence 13209, Application US/10917503B
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
```

```
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13209
;   LENGTH: 1898
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1086)
US-10-917-503B-13209

Alignment Scores:
Pred. No.:                3.32e-51        Length:         1898
Score:                    439.00          Matches:        86
Percent Similarity:       98.9%           Conservative:   0
Best Local Similarity:    98.9%           Mismatches:     1
Query Match:              98.0%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-917-503B-13209 (1-1898)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        160 ATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGC 219

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339

Qy         61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 399

Qy         81 ArgGlyValCysSerTyrLeu 87
              |||||||||||||||||||||
Db        400 CGAGGGGTGTGCTCCTACTTA 420

RESULT 6
US-60-762-056-35503/c
; Sequence 35503, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 35503
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-35503

Alignment Scores:
Pred. No.:                1.3e-09         Length:         2300
Score:                    145.00          Matches:        28
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    96.6%           Mismatches:     0
Query Match:              32.4%           Indels:         0
DB:                       11              Gaps:           0

US-10-757-745-2_COPY_54_140 (1-87) x US-60-762-056-35503 (1-2300)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1668 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 1609

Qy         22 ProGluThrIleSerGluProLysThr 30
              ||||||||||||||||||||||||||||
Db       1608 CCTGAAACCATCTCTGAGCCCAAGACC 1582

RESULT 7
US-11-266-748A-72123/c
; Sequence 72123, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 72123
;   LENGTH: 767
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (66)..(66)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (95)..(95)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (101)..(101)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (109)..(110)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (118)..(118)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-72123

Alignment Scores:
Pred. No.:                0.784           Length:         767
Score:                    78.00           Matches:        26
Percent Similarity:       48.5%           Conservative:   7
Best Local Similarity:    38.2%           Mismatches:     32
Query Match:              17.4%           Indels:         3
DB:                       9               Gaps:           1
```

```
US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-72123 (1-767)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::   |||       |||
Db        565 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 506

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::     |||   |||   |||      |||:::|||         |||||||||
Db        505 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 446

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||          |||   ||||||:::|||:::
Db        445 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 386

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::     |||
Db        385 TGTCTGGACTTTGAGCAGGCTT 364

RESULT 8
US-11-266-748A-107071/c
; Sequence 107071, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 107071
;   LENGTH: 767
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (66)..(66)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (95)..(95)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (101)..(101)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (109)..(110)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (118)..(118)
;   OTHER INFORMATION: n is a, c, g, or t

US-11-266-748A-107071

Alignment Scores:
Pred. No.:                0.784          Length:          767
Score:                    78.00          Matches:         26
Percent Similarity:       48.5%          Conservative:    7
Best Local Similarity:    38.2%          Mismatches:      32
Query Match:              17.4%          Indels:          3
DB:                       9              Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-107071 (1-767)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::   |||       |||
Db        565 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 506

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::     |||   |||   |||      |||:::|||         |||||||||
Db        505 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 446

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||          |||   ||||||:::|||:::
Db        445 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 386

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::     |||
Db        385 TGTCTGGACTTTGAGCAGGCTT 364

RESULT 9
US-11-266-748A-124934
; Sequence 124934, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 124934
;   LENGTH: 767
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (650)..(650)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (658)..(659)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (667)..(667)
```

```
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (673)..(673)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (702)..(702)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-124934
```

Alignment Scores:

Pred. No.:	0.784	Length:	767
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	9	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-124934 (1-767)

```
Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db        203 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 262

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::     |||   |||  |||     |||:::|||          |||||||||
Db        263 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 322

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||          |||   ||||||:::|||:::
Db        323 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 382

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db        383 TGTCTGGACTTTGAGCAGGCTT 404
```

```
RESULT 10
US-11-266-748A-161851
; Sequence 161851, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 161851
;   LENGTH: 1548
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-161851
```

Alignment Scores:

Pred. No.:	2.11	Length:	1548
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	9	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-161851 (1-1548)

```
Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||       |||
Db        871 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 930

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::     |||   |||  |||     |||:::|||          |||||||||
Db        931 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 990

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||          |||   ||||||:::|||:::
Db        991 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1050

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       1051 TGTCTGGACTTTGAGCAGGCTT 1072
```

```
RESULT 11
US-11-266-748A-184381
; Sequence 184381, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 184381
;   LENGTH: 1982
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1765)..(1812)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-184381
```

Alignment Scores:

Pred. No.:	2.98	Length:	1982
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	9	Gaps:	1

```
US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-184381 (1-1982)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||      |||
Db         90 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 149

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||         |||||||||
Db        150 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 209

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||         |||   ||||||:::|||:::
Db        210 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 269

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                |||   :::      |||
Db        270 TGTCTGGACTTTGAGCAGGCTT 291

RESULT 12
US-11-266-748A-191888
; Sequence 191888, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 191888
;   LENGTH: 1982
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1765)..(1812)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-191888

Alignment Scores:
Pred. No.:              2.98          Length:        1982
Score:                  78.00         Matches:       26
Percent Similarity:     48.5%         Conservative:  7
Best Local Similarity:  38.2%         Mismatches:    32
Query Match:            17.4%         Indels:        3
DB:                     9             Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-191888 (1-1982)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||      |||
Db         90 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 149

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||         |||||||||
Db        150 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 209

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||         |||   ||||||:::|||:::
Db        210 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 269

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                |||   :::      |||
Db        270 TGTCTGGACTTTGAGCAGGCTT 291

RESULT 13
US-11-266-748A-184374
; Sequence 184374, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 184374
;   LENGTH: 2405
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (58)..(94)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-184374

Alignment Scores:
Pred. No.:              3.91          Length:        2405
Score:                  78.00         Matches:       26
Percent Similarity:     48.5%         Conservative:  7
Best Local Similarity:  38.2%         Mismatches:    32
Query Match:            17.4%         Indels:        3
DB:                     9             Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-184374 (1-2405)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::    |||      |||
Db       1575 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1634

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||         |||||||||
Db       1635 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1694

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||         |||   ||||||:::|||:::
Db       1695 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1754
```

```
Qy          64 eThrTrpAsnIleAspGlyLeu 71
                  |||   :::       |||
Db        1755 TGTCTGGACTTTGAGCAGGCTT 1776

RESULT 14
US-11-266-748A-191881
; Sequence 191881, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 191881
;   LENGTH: 2405
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (58)..(94)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-191881

Alignment Scores:
Pred. No.:                3.91          Length:          2405
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       9             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-191881 (1-2405)

Qy           7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
               |||   |||   ||||||:::           ||||||:::   |||       |||
Db        1575 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1634

Qy          26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
                 :::      |||   |||   |||      |||:::|||         |||||||||
Db        1635 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1694

Qy          46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                   |||   ||||||   |||          |||   ||||||:::|||:::
Db        1695 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1754

Qy          64 eThrTrpAsnIleAspGlyLeu 71
                  |||   :::       |||
Db        1755 TGTCTGGACTTTGAGCAGGCTT 1776

RESULT 15
US-10-120-780-143
; Sequence 143, Application US/10120780
; GENERAL INFORMATION:
;  APPLICANT: KIYAMA, RYOITI
;  APPLICANT:  OGUCHI, SHINOBU
;  TITLE OF INVENTION: MICROARRAYS AND METHODS FOR EVALUATING ACTIVITY OF COMPOUNDS
;  TITLE OF INVENTION:  HAVING ESTROGEN-LIKE ACTIVITY
;  FILE REFERENCE: INF-A0101-USC1
;  CURRENT APPLICATION NUMBER: US/10/120,780
;  CURRENT FILING DATE:  2002-04-12
;  PRIOR APPLICATION NUMBER: JP 2000-313403
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 09/975,316
;  PRIOR FILING DATE: 2001-10-12
;  PRIOR APPLICATION NUMBER: JP 2001-317152
;  PRIOR FILING DATE: 2001-10-15
;  PRIOR APPLICATION NUMBER: US 60/300,835
;  PRIOR FILING DATE: 2001-06-27
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 143
;   LENGTH: 2745
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-120-780-143

Alignment Scores:
Pred. No.:                4.71          Length:          2745
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       6             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-120-780-143 (1-2745)

Qy           7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
               |||   |||   ||||||:::           ||||||:::   |||       |||
Db         833 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 892

Qy          26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
                 :::      |||   |||   |||      |||:::|||         |||||||||
Db         893 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 952

Qy          46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                   |||   ||||||   |||          |||   ||||||:::|||:::
Db         953 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1012

Qy          64 eThrTrpAsnIleAspGlyLeu 71
                  |||   :::       |||
Db        1013 TGTCTGGACTTTGAGCAGGCTT 1034

RESULT 16
US-11-266-748A-72116
; Sequence 72116, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
```

```
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 72116
;   LENGTH: 2792
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2445)..(2492)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2674)..(2707)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-72116
```

Alignment Scores:

Pred. No.:	4.83	Length:	2792
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	9	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-72116 (1-2792)

```
Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::   |||      |||
Db        770 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 829

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||         |||||||||
Db        830 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 889

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||       |||   ||||||:::|||:::
Db        890 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 949

Qy         64 eThrTrpAsnIleAspGlyLeu 71
              |||   :::      |||
Db        950 TGTCTGGACTTTGAGCAGGCTT 971
```

```
RESULT 17
US-11-266-748A-107064
; Sequence 107064, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 107064
;   LENGTH: 2792
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2445)..(2492)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2674)..(2707)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-107064
```

Alignment Scores:

Pred. No.:	4.83	Length:	2792
Score:	78.00	Matches:	26
Percent Similarity:	48.5%	Conservative:	7
Best Local Similarity:	38.2%	Mismatches:	32
Query Match:	17.4%	Indels:	3
DB:	9	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-107064 (1-2792)

```
Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::           ||||||:::   |||      |||
Db        770 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 829

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||      |||:::|||         |||||||||
Db        830 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 889

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                    |||   ||||||   |||       |||   ||||||:::|||:::
Db        890 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 949

Qy         64 eThrTrpAsnIleAspGlyLeu 71
              |||   :::      |||
Db        950 TGTCTGGACTTTGAGCAGGCTT 971
```

```
RESULT 18
US-11-266-748A-124927/c
; Sequence 124927, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
```

```
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 124927
;   LENGTH: 2792
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (86)..(119)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (301)..(348)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-124927

Alignment Scores:
Pred. No.:                4.83          Length:          2792
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       9             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-124927 (1-2792)

Qy           7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
               |||   |||   ||||||:::          ||||||:::    |||       |||
Db        2023 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1964

Qy          26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
                :::     |||  |||   |||      |||:::|||           |||||||||
Db        1963 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 1904

Qy          46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                 |||   ||||||   |||          |||   ||||||:::|||:::
Db        1903 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 1844

Qy          64 eThrTrpAsnIleAspGlyLeu 71
               |||   :::     |||
Db        1843 TGTCTGGACTTTGAGCAGGCTT 1822

RESULT 19
US-60-741-051-378
; Sequence 378, Application US/60741051
; GENERAL INFORMATION:
;  APPLICANT: BIRSE, Charles et al.
;  TITLE OF INVENTION: BREAST CANCER SECRETED TARGETS AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001625
;  CURRENT APPLICATION NUMBER: US/60/741,051
;  CURRENT FILING DATE:  2005-12-01
;  NUMBER OF SEQ ID NOS: 526
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 378
;   LENGTH: 4482
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-741-051-378

Alignment Scores:
Pred. No.:                9.39          Length:          4482
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       11            Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-60-741-051-378 (1-4482)

Qy           7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
               |||   |||   ||||||:::          ||||||:::    |||       |||
Db        3930 AGCATGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 3989

Qy          26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
                :::     |||  |||   |||      |||:::|||           |||||||||
Db        3990 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4049

Qy          46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                 |||   ||||||   |||          |||   ||||||:::|||:::
Db        4050 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4109

Qy          64 eThrTrpAsnIleAspGlyLeu 71
               |||   :::     |||
Db        4110 TGTCTGGACTTTGAGCAGGCTT 4131

RESULT 20
US-60-741-051-379
; Sequence 379, Application US/60741051
; GENERAL INFORMATION:
;  APPLICANT: BIRSE, Charles et al.
;  TITLE OF INVENTION: BREAST CANCER SECRETED TARGETS AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001625
;  CURRENT APPLICATION NUMBER: US/60/741,051
;  CURRENT FILING DATE:  2005-12-01
;  NUMBER OF SEQ ID NOS: 526
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 379
;   LENGTH: 4482
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-741-051-379

Alignment Scores:
Pred. No.:                9.39          Length:          4482
Score:                    78.00         Matches:         26
Percent Similarity:       48.5%         Conservative:    7
Best Local Similarity:    38.2%         Mismatches:      32
Query Match:              17.4%         Indels:          3
DB:                       11            Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-60-741-051-379 (1-4482)

Qy           7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
               |||   |||   ||||||:::          ||||||:::    |||       |||
Db        3930 AGCATGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 3989

Qy          26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
                :::     |||  |||   |||      |||:::|||           |||||||||
Db        3990 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4049

Qy          46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                 |||   ||||||   |||          |||   ||||||:::|||:::
Db        4050 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4109

Qy          64 eThrTrpAsnIleAspGlyLeu 71
               |||   :::     |||
Db        4110 TGTCTGGACTTTGAGCAGGCTT 4131

RESULT 21
US-11-287-334-53
; Sequence 53, Application US/11287334
; GENERAL INFORMATION:
;  APPLICANT: Gleave, Martin.
;  TITLE OF INVENTION: Treatment of cancer by inhibition of IGFBPs and Clusterin
;  FILE REFERENCE: UBC_P043
;  CURRENT APPLICATION NUMBER: US/11/287,334
;  CURRENT FILING DATE:  2005-11-23
;  PRIOR APPLICATION NUMBER: US 60/522,960
;  PRIOR FILING DATE: 2004-11-24
;  PRIOR APPLICATION NUMBER: US 60/522,948
;  PRIOR FILING DATE: 2004-11-23
;  PRIOR APPLICATION NUMBER: US 10/346,493
;  PRIOR FILING DATE: 2003-01-17
;  PRIOR APPLICATION NUMBER: 60/350,046
```

```
;  PRIOR FILING DATE: 2002-01-17
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 53
;   LENGTH: 6316
;   TYPE: DNA
;   ORGANISM: human
US-11-287-334-53

Alignment Scores:
Pred. No.:              15.2          Length:         6316
Score:                  78.00         Matches:        26
Percent Similarity:     48.5%         Conservative:   7
Best Local Similarity:  38.2%         Mismatches:     32
Query Match:            17.4%         Indels:         3
DB:                     7             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-287-334-53 (1-6316)

Qy          7 SerTyrPheGluProProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
              |||   |||   ||||||:::          ||||||:::    |||      |||
Db       4103 TCACTGTTCTTTCCACCCATTCTCCCTTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 4162

Qy         26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLy 46
               :::      |||   |||   |||       |||:::|||          |||||||||
Db       4163 CCAAGTGGATACTGGGGTGCCACTCCCCTGAGAAAAGACTGAGCCAGGAACTACAAGCTC 4222

Qy         46 s------IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIl 64
                  |||   ||||||   |||         |||   ||||||:::|||:::
Db       4223 CCCCCACATTCCTCCCAGCCTGGACCTAATTCTTGAGAGGGGCTCTCTCTTCACGGACTG 4282

Qy         64 eThrTrpAsnIleAspGlyLeu 71
                 |||   :::      |||
Db       4283 TGTCTGGACTTTGAGCAGGCTT 4304

RESULT 22
PCT-US03-37356-1
; Sequence 1, Application PC/TUS0337356
; GENERAL INFORMATION:
;  APPLICANT: Dahl, Gary A.
;  APPLICANT:  Jendrisak, Jerome J.
;  APPLICANT:  Davydova, Elena K.
;  APPLICANT:  Rothman-Denes, Lucia B.
;  APPLICANT:  Gerdes, Svetlana Y.
;  TITLE OF INVENTION: PREPARATION AND USE OF SINGLE-STRANDED TRANSCRIPTION SUBSTRATES
;  TITLE OF INVENTION:  FOR SYNTHESIS OF TRANSCRIPTION PRODUCTS CORRESPONDING TO TARGET
;  TITLE OF INVENTION:  SEQUENCES
;  FILE REFERENCE: ARCD:?
;  CURRENT APPLICATION NUMBER: PCT/US03/37356
;  CURRENT FILING DATE:  2003-11-21
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 10506
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
PCT-US03-37356-1

Alignment Scores:
Pred. No.:              70.1          Length:         10506
Score:                  75.50         Matches:        24
Percent Similarity:     46.1%         Conservative:   17
Best Local Similarity:  27.0%         Mismatches:     27
Query Match:            16.9%         Indels:         21
DB:                     1             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x PCT-US03-37356-1 (1-10506)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::           |||:::   |||   ||||||
Db       2851 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 2898

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                         |||      |||:::|||||||||||:::                |||
Db       2899 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 2958

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                    :::      :::|||   :::   :::|||               |||   :::
Db       2959 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3003

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       3004 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3063

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       3064 TCTAAAGCAGAGGGTATTAAGAACTAT 3090

RESULT 23
PCT-US03-37356-14
; Sequence 14, Application PC/TUS0337356
; GENERAL INFORMATION:
;  APPLICANT: Dahl, Gary A.
;  APPLICANT:  Jendrisak, Jerome J.
;  APPLICANT:  Davydova, Elena K.
;  APPLICANT:  Rothman-Denes, Lucia B.
;  APPLICANT:  Gerdes, Svetlana Y.
;  TITLE OF INVENTION: PREPARATION AND USE OF SINGLE-STRANDED TRANSCRIPTION SUBSTRATES
;  TITLE OF INVENTION:  FOR SYNTHESIS OF TRANSCRIPTION PRODUCTS CORRESPONDING TO TARGET
;  TITLE OF INVENTION:  SEQUENCES
;  FILE REFERENCE: ARCD:?
;  CURRENT APPLICATION NUMBER: PCT/US03/37356
;  CURRENT FILING DATE:  2003-11-21
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 10617
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
PCT-US03-37356-14

Alignment Scores:
Pred. No.:              71.1          Length:         10617
Score:                  75.50         Matches:        24
Percent Similarity:     46.1%         Conservative:   17
Best Local Similarity:  27.0%         Mismatches:     27
Query Match:            16.9%         Indels:         21
DB:                     1             Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x PCT-US03-37356-14 (1-10617)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28
              ||||||      :::||||||:::           |||:::   |||   ||||||
Db       2962 GAACCTACTACTAAAGAGTCTAGC-----------CCACAGGCTATAAAAGAACCTGTG 3009

Qy         29 -----------LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                         |||      |||:::|||||||||||:::                |||
Db       3010 AACCAGTCTGAAAAACAGGATGTTAACCTTACTAATGAGGACAACATCAAGCAACCTACT 3069

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                    :::      :::|||   :::   :::|||               |||   :::
Db       3070 GAATCTGTTAAAGAAACTGAAACTTCTACAAAAGAA--------------AGTACAGTT 3114

Qy         65 ThrTrpAsnIle---AspGlyLeuAsp-----------------LeuAsnAsnLeuSer 77
              |||      :::   :::|||:::|||                 :::
Db       3115 ACAGAAGAATTAAAAGAAGGTATTGATGCTGTTTACCCTTCATTGGTAGGTACTGCTGAT 3174

Qy         78 GluArgAlaArgGlyValCysSerTyr 86
                 :::|||   |||:::   :::|||
Db       3175 TCTAAAGCAGAGGGTATTAAGAACTAT 3201
```

```
RESULT 24
US-10-556-637-6
; Sequence 6, Application US/10556637
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: MYLKS AS MODIFIERS OF BRANCHING MORPHOGENESIS AND METHODS OF USE
;  FILE REFERENCE: EX04-042C-US
;  CURRENT APPLICATION NUMBER: US/10/556,637
;  CURRENT FILING DATE:  2005-11-10
;  PRIOR APPLICATION NUMBER: US 60/479,781
;  PRIOR FILING DATE: 2003-06-19
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
;   LENGTH: 5713
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-556-637-6

Alignment Scores:
Pred. No.:              57          Length:         5713
Score:                  73.50       Matches:        22
Percent Similarity:     47.9%       Conservative:   13
Best Local Similarity:  30.1%       Mismatches:     27
Query Match:            16.4%       Indels:         11
DB:                     6           Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-556-637-6 (1-5713)

Qy          3 ArgAlaLeuAsnSerTyr---PheGluProProValGluGluSerAlaLeuGluArgArg 21
              ||||||::|||   |||            |||   :::||||||   |||
Db       4339 CGTGCAATCAACGTGTATGGAACCAGTGAGCCAAGCCAGGAGTCTGAACTCACAACGGTA 4398

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu--------- 38
              |||       |||||||||      |||:::::::::::::::::
Db       4399 GGAGAGAAACCTGAAGAGCCGAAGGATGAAGTGGAGGTGTCAGATGATGATGAGAAGGAG 4458

Qy         39 ---------------------ThrThrAspSerThrThrSerLysIleSerProSerGlu 51
                                   |||   :::|||       |||:::|||
Db       4459 CCCGAGGTTGATTACCGGACAGTGACAATCAATACTGAACAAAAAGTATCTGACTTCTAC 4518

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              |||   ::::::      ||||||:::|||      |||
Db       4519 GACATTGAGGAGAGATTAGGATCGCTGTTTGAGCGCATC 4557

RESULT 25
US-10-513-024-863
; Sequence 863, Application US/10513024
; GENERAL INFORMATION:
;  APPLICANT: CHIRON SRL
;  APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
;  APPLICANT: TELFORD John
;  APPLICANT: MASIGNANI Vega
;  APPLICANT: MARGARIT Y ROS Immaculada
;  APPLICANT: GRANDI Guido
;  APPLICANT: FRASER Claire
;  APPLICANT: TETTELIN Herve
;  TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
;  FILE REFERENCE: P030497WO
;  CURRENT APPLICATION NUMBER: US/10/513,024
;  CURRENT FILING DATE:  2004-11-01
;  PRIOR APPLICATION NUMBER: GB-0210128.5
;  PRIOR FILING DATE: 2002-05-02
;  NUMBER OF SEQ ID NOS: 1373
;  SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 863
;   LENGTH: 933
;   TYPE: DNA
;   ORGANISM: Streptococcus pyogenes
US-10-513-024-863

Alignment Scores:
Pred. No.:              8.53        Length:         933
Score:                  71.50       Matches:        18
Percent Similarity:     58.9%       Conservative:   15
Best Local Similarity:  32.1%       Mismatches:     14
Query Match:            16.0%       Indels:         9
DB:                     6           Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-513-024-863 (1-933)

Qy         15 GluSerAlaLeuGluArgArgPro---GluThrIleSerGluProLysThrTyrValAsp 33
              ||||||:::::::::|||||||||   ||||||:::|||:::      :::   :::
Db        763 GAGTCAAGTGTCGATAGACGCCCTATGGAAACTGTTTCTAAA---GATAGTGGTATTCCT 819

Qy         34 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
              :::   :::|||      |||||||||          |||      |||   :::
Db        820 ATTTATTCTGAGATCTTTACAGATTCAATTGCTAAAAAAGGTAAACCTGGCGAT------ 873

Qy         54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 69
                            |||   ::::::::::::   ||||||:::|||
Db        874 --------------AGTTATTATGCTATGATGAAATGGAACCTTGAC 906

RESULT 26
US-11-348-413-1579
; Sequence 1579, Application US/11348413
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  APPLICANT:  Murphy, Ellen
;  APPLICANT:  Olmsted, Stephen
;  TITLE OF INVENTION: PROBE ARRAYS FOR DETECTING MULTIPLE STRAINS OF DIFFERENT SPECIES
;  FILE REFERENCE: 031896-084100 (AM 101724)
;  CURRENT APPLICATION NUMBER: US/11/348,413
;  CURRENT FILING DATE:  2006-02-07
;  PRIOR APPLICATION NUMBER: PCT/US05/035471
;  PRIOR FILING DATE: 2005-10-05
;  PRIOR APPLICATION NUMBER: US 11/243,445
;  PRIOR FILING DATE: 2005-10-05
;  PRIOR APPLICATION NUMBER: US 60/615,573
;  PRIOR FILING DATE: 2004-10-05
;  NUMBER OF SEQ ID NOS: 1276209
; SEQ ID NO 1579
;   LENGTH: 936
;   TYPE: DNA
;   ORGANISM: Streptococcus pyogenes
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(936)
;   OTHER INFORMATION: WAN01UKT5; ORF; 11111100000000; Cluster contains WAN01P6PA
;   OTHER INFORMATION: :SPy0453:mtsA:metal binding protein of ABC transporter (li
;   OTHER INFORMATION: poprotein):SF370:NC_002737.1
US-11-348-413-1579

Alignment Scores:
Pred. No.:              8.57        Length:         936
Score:                  71.50       Matches:        18
Percent Similarity:     58.9%       Conservative:   15
Best Local Similarity:  32.1%       Mismatches:     14
Query Match:            16.0%       Indels:         9
DB:                     10          Gaps:           3

US-10-757-745-2_COPY_54_140 (1-87) x US-11-348-413-1579 (1-936)

Qy         15 GluSerAlaLeuGluArgArgPro---GluThrIleSerGluProLysThrTyrValAsp 33
              ||||||:::::::::|||||||||   ||||||:::|||:::      :::   :::
Db        766 GAGTCAAGTGTCGATAGRCGCCCTATGGAAACTGTTTCTAAA---GATAGTGGTATTCCT 822

Qy         34 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
              :::   :::|||      |||||||||          |||      |||   :::
Db        823 ATTTATTCTGAGATCTTTACAGATTCAATTGCTAAAAAAGGTAAACCTGGCGAT------ 876

Qy         54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 69
                            |||   ::::::::::::   ||||||:::|||
Db        877 --------------AGTTATTATGCTATGATGAAATGGAACCTTGAC 909
```

```
RESULT 27
US-11-045-004-1/c
; Sequence 1, Application US/11045004
; GENERAL INFORMATION:
;  APPLICANT: BUCHRIESER, CARMEN
;  APPLICANT:  FRANGEUL, LIONEL
;  APPLICANT:  COUVE, ELISABETH
;  APPLICANT:  RUSNIOK, CHRISTOPHE
;  APPLICANT:  FSIHI, HAFIDA
;  APPLICANT:  DEHOUX, PIERRE
;  APPLICANT:  DUSSURGET, OLIVIER
;  APPLICANT:  CHETOUANI, FARID
;  APPLICANT:  NEDJARI, HAFED
;  APPLICANT:  GLASER, PHILIPPE
;  APPLICANT:  KUNST, FRANCK
;  APPLICANT:  COSSART, PASCALE
;  APPLICANT:  DANIELS, JUSTIN
;  APPLICANT:  GOEBEI, WERNER
;  APPLICANT:  KREFT, JURGEN
;  APPLICANT:  KUHN, MICHAEL
;  APPLICANT:  NG, EVA
;  APPLICANT:  VAZQUEZ-BOLAND, ANTONIO
;  APPLICANT:  DOMINGUEZ-BERNAL, GUSTAVO
;  APPLICANT:  GARRIDO-GARCIA, PATRICIA
;  APPLICANT:  TIERREZ-MARTINEZ, ALBERTO
;  APPLICANT:  AMEND, ALEXANDRA
;  APPLICANT:  CHAKRABORTY, TRINAD
;  APPLICANT:  DOMANN, EUGEN
;  APPLICANT:  HAIN, THORSTEN
;  APPLICANT:  BERCHE, PATRICK
;  APPLICANT:  CHARBIT, ALAIN
;  APPLICANT:  DURANT, LIONEL
;  APPLICANT:  PEREZ-DIAZ, JOSE-CLAUDIO
;  APPLICANT:  BAQUERO, FERNANDO
;  APPLICANT:  GARCIA DEL PORTILLO, FRANCISCO
;  APPLICANT:  GOMEZ-LOPEZ, NURIA
;  APPLICANT:  MADUENIO, ENCARNA
;  APPLICANT:  PABLOS, BETRIZ DE
;  APPLICANT:  WEHLAND, JURGEN
;  APPLICANT:  KARST, UWE
;  APPLICANT:  ENTIAN, KARL-DIETER
;  APPLICANT:  HAUF, JORG
;  APPLICANT:  ROSE, MATTHIAS
;  APPLICANT:  VOSS, HAMUT
;  TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
;  FILE REFERENCE: 05394.0018-02
;  CURRENT APPLICATION NUMBER: US/11/045,004
;  CURRENT FILING DATE:  2005-01-28
;  PRIOR APPLICATION NUMBER: 10/637,657
;  PRIOR FILING DATE: 2003-08-11
;  PRIOR APPLICATION NUMBER: 10/257,023
;  PRIOR FILING DATE: 2002-10-08
;  PRIOR APPLICATION NUMBER: PCT/FR01/01118
;  PRIOR FILING DATE: 2001-04-11
;  PRIOR APPLICATION NUMBER: FR 00/04,629
;  PRIOR FILING DATE: 2000-04-11
;  NUMBER OF SEQ ID NOS: 2854
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
;   LENGTH: 2944528
;   TYPE: DNA
;   ORGANISM: Listeria monocytogenes
US-11-045-004-1

Alignment Scores:
Pred. No.:                8.02e+05        Length:         2944528
Score:                    71.00           Matches:        28
Percent Similarity:       45.3%           Conservative:   15
Best Local Similarity:    29.5%           Mismatches:     20
Query Match:              15.8%           Indels:         32
DB:                       9               Gaps:           5

US-10-757-745-2_COPY_54_140 (1-87) x US-11-045-004-1 (1-2944528)

Qy         14 GluGluSerAlaLeuGluArgArg------ProGluThrIleSerGluProLysThrTyr 31
              :::::::::|||||||||||||||      |||   |||:::|||||||||   |||
Db    1028283 AAAGATGCCGCACTTGAACGACGCTTCCAACCAGTAACTGTAAGCGAACCATCCACAAAA 1028224

Qy         32 ValAspLeuThr----------------------AsnGluGlu-------------- 38
                  ||||||                      |||:::
Db    1028223 GAAACGCTAACTATTTTAAATGGTTTAAAACCAAAATATGAAGATTTCCATGAAGTAGTT 1028164

Qy         39 ---ThrThrAspSerThrThrSerLysIleSerProSerGlu---------------- 51
                 :::   ::::::   |||:::   :::      |||
Db    1028163 TATTCACCTGAAGCATTAACAGCAGCAGTTGAATTAAGCGCTCGTTATATCCAAGATCGT 1028104

Qy         52 --------------------AspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
                                  |||         |||   ||||||   ::::::|||
Db    1028103 CATTTGCCAGACAAAGCAATTGATTTAATGGATGAAGTTGGCTCAAAATATAATCTA--- 1028047

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
                 :::|||:::   ||||||   |||   :::|||||||||
Db    1028046 ------TCCATTGAAAAACTAGACGAAAATACGGTAAGCGAACGT 1028008

RESULT 28
US-10-567-867-2265
; Sequence 2265, Application US/10567867
; GENERAL INFORMATION:
;  APPLICANT: Li, Martha
;  APPLICANT:  Rupnow, Brent A.
;  APPLICANT:  Webster, Kevin R.
;  APPLICANT:  Jackson, Donald
;  APPLICANT:  Wong, Tai W.
;  TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
;  FILE REFERENCE: D0310 PCT
;  CURRENT APPLICATION NUMBER: US/10/567,867
;  CURRENT FILING DATE:  2006-01-27
;  PRIOR APPLICATION NUMBER: US 60/490,890
;  PRIOR FILING DATE: 2003-07-29
;  NUMBER OF SEQ ID NOS: 2786
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2265
;   LENGTH: 5048
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-567-867-2265

Alignment Scores:
Pred. No.:                149             Length:         5048
Score:                    70.00           Matches:        33
Percent Similarity:       39.8%           Conservative:   12
Best Local Similarity:    29.2%           Mismatches:     32
Query Match:              15.6%           Indels:         36
DB:                       6               Gaps:           8

US-10-757-745-2_COPY_54_140 (1-87) x US-10-567-867-2265 (1-5048)

Qy          6 AsnSerTyrPheGluPro--------------------------------ProValGlu 14
                |||   ::::::||||||                               ||||||:::
Db       1456 AATCCTTTTTATGAACCTAAATCAACTCCTCCTCCAAATAATTTGGTAAATCCTGTTCAA 1515

Qy         15 GluSerAlaLeuGluArgArg--------------------ProGluThrIleSerGlu 27
              |||         |||||||||                     |||      :::|||
Db       1516 GAACTAGAAACTGAAAGGCGAGTGAAAAGAAAGGCCCCGGCTCCACCAGTCCTCTCA--- 1572

Qy         28 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp--------------Ser 42
              |||||||||   |||        ||||||   |||                     |||
Db       1573 CCAAAAACAGGAGTA------TTAAATGAAAACACAGTTTCTGCAGGAAAAGATCTCTCT 1626

Qy         43 ThrThrSerLysIleSerProSerGluAspThr---------GlnGlnGluAsnGlySer 59
              |||:::   |||   ||||||                       ::::::   |||   |||
Db       1627 ACTTCTCCTAAGCCAAGCCCTATACCAAGTCCTGTTTTGGGGCGAAAGCCAAATGCTAGT 1686

Qy         60 MetPheSerLeuIleThrTrp-----------------AsnIleAspGlyLeuAspLeu 73
                |||||:::   |||                 |||      |||:::   :::
```

```
Db       1687 CAG---TCTTTGCTTGTATGGTGTAAAGAAGTTACAAAGAACTACCGAGGAGTAAAAATC 1743

Qy         74 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
                |||   :::       |||       |||:::
Db       1744 ACCAATTTTACTACATCGTGGAGAAATGGTTTATCTTTT 1782
```

RESULT 29
US-11-353-150-46076/c
; Sequence 46076, Application US/11353150
; GENERAL INFORMATION:
; APPLICANT: Andersen, Scott E.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Lalgudi, Raghunath V.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.365 - 38-21(15459)D/US
; CURRENT APPLICATION NUMBER: US/11/353,150
; CURRENT FILING DATE: 2006-02-14
; PRIOR APPLICATION NUMBER: 09/565,386
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/132,860
; PRIOR FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 83523
; SEQ ID NO 46076
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3066-032-Q1-E1-B4
US-11-353-150-46076

Alignment Scores:

Pred. No.:	9.48	Length:	503
Score:	68.50	Matches:	20
Percent Similarity:	50.0%	Conservative:	12
Best Local Similarity:	31.2%	Mismatches:	19
Query Match:	15.3%	Indels:	13
DB:	7	Gaps:	2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-353-150-46076 (1-503)

```
Qy         16 SerAlaLeuGluArgArgProGluThrIle-----------SerGluProLysThrTyr 31
              |||   |||   |||        ||||||            :::|||   ||||||
Db        266 AGCAAACTTACACGAAATCTTAACACGATCAACAAAACCTAAAACGAAAACAAAACCACA 207

Qy         32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
                :::||||||   |||     ::::::||||||||:::   :::|||   :::
Db        206 ACAGGCATGACCAATACCGAAAAACAAAACACCACAACCTCAAGGAAAACACCTAAAGAC 147

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
                 :::|||::::::              ::::::   |||
Db        146 CATCAGCGTCAACAGTCC-------------------------TCGCTCCGGGGCCCT 114

Qy         72 AspLeuAsnAsn 75
              |||   ||||||
Db        113 GACAAGAACAAC 102
```

RESULT 30
US-10-953-349-3481/c
; Sequence 3481, Application US/10953349
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCONDED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3481
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

US-10-953-349-3481

Alignment Scores:

Pred. No.:	37.4	Length:	1335
Score:	68.50	Matches:	23
Percent Similarity:	44.9%	Conservative:	12
Best Local Similarity:	29.5%	Mismatches:	32
Query Match:	15.3%	Indels:	11
DB:	6	Gaps:	2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-953-349-3481 (1-1335)

```
Qy          3 ArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgPro 22
              |||       |||   ||||||:::      :::       |||              |||
Db       1194 AGATGTTTTAATCTTTATTTTAAACATAAAATAACAATAAGCAAGAGACACAACCGCGGG 1135

Qy         23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
              :::|||   ::::::   :::|||                              |||:::   |||
Db       1134 AAAACAACAACTCAATAACGAACC-----------------------ACTTCCTATTCC 1099

Qy         43 ThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
              ||||||:::      |||          |||          |||      ||||||:::   |||
Db       1098 ACAACCACAAATTATTCTTCTCGTATTGATTTTTCTCAATCCAGAGGAAGTTTATTGAGC 1039

Qy         63 LeuIle---------ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer 77
              |||              :::         :::|||||:::          ||||||
Db       1038 TTGCGGTAACTGTAATCCTTGAAAGCAAACGGTTTAAACTTTGGAGGATTATCA 985
```

RESULT 31
US-11-331-019-24356
; Sequence 24356, Application US/11331019
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen L.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: McCarter, David W.
; APPLICANT: Pear, Julie R.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 16517.359 - 38-21(51375)C/US
; CURRENT APPLICATION NUMBER: US/11/331,019
; CURRENT FILING DATE: 2006-01-13
; PRIOR APPLICATION NUMBER: US 09/637,086
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/149,881
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 52949
; SEQ ID NO 24356
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3147-019-Q1-K1-G11
US-11-331-019-24356

Alignment Scores:

Pred. No.:	11.7	Length:	465
Score:	67.50	Matches:	17
Percent Similarity:	48.0%	Conservative:	7
Best Local Similarity:	34.0%	Mismatches:	19
Query Match:	15.1%	Indels:	7
DB:	7	Gaps:	1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-019-24356 (1-465)

```
Qy         17 AlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsn 36
              :::   |||      |||      ::::::   ||||||      |||      |||
Db         96 AGTTTTGAAGAGATCCCTGCCCTGCTGAATACCCCAAAAGAGGAACCAGACTCTCAAAAT 155

Qy         37 GluGluThrThrAspSerThrThrSerLysIle---------------------SerPro 49
              ::::::      |||:::      |||||||||                      |||
Db        156 GACAAGGTTGCTGATAATATTAAAAGCAAAATCATAAATGTAGACAAGATATTGGTGCCC 215

Qy         50 SerGluAspThrGlnGlnGluAsnGlySer 59
```

```
                    ||||||   |||||||||   |||:::
Db        216 AAAGAAGATCTGCAACAGGAAGTAGGCACT 245

RESULT 32
US-11-266-748A-25375
; Sequence 25375, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25375
;   LENGTH: 3594
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-25375

Alignment Scores:
Pred. No.:                208             Length:         3594
Score:                    67.50           Matches:        20
Percent Similarity:       49.3%           Conservative:   17
Best Local Similarity:    26.7%           Mismatches:     35
Query Match:              15.1%           Indels:         3
DB:                       9               Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-25375 (1-3594)

Qy          12 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
               |||:::|||:::         |||      |||||||||:::   :::   :::|||
Db        2118 CCCATGGAAAAGGACCCAGCAGAG------CCAGAAACAGTGCCTGATGGCGAAACCCCA 2171

Qy          32 ValAspLeuThr---AsnGluGluThrThrAspSerThrThrSerLysIleSerProSer 50
               |||           ||||||      |||:::::::::::|||:::
Db        2172 GAAGATGAAAATCCAACAGAGGAAGGAGCAGACAACTCTTCAGCAAAGATGGAAGAGGAG 2231

Qy          51 GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGly 70
               |||:::   ::::::|||                        |||      |||
Db        2232 GAGGAGGAAGAGGAAGAAGAAGAAGAGAGCCTCCCAGGATGTACTCTGTTTATTAAGAAT 2291

Qy          71 LeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 85
               |||:::   :::         |||:::   :::   |||   |||
Db        2292 CTCAATTTTGACACAACAGAAGAGAAGCTGAAGGAAGTGTTTTCA 2336

RESULT 33
US-11-266-748A-30209
; Sequence 30209, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30209
;   LENGTH: 4148
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-30209

Alignment Scores:
Pred. No.:                255             Length:         4148
Score:                    67.50           Matches:        20
Percent Similarity:       49.3%           Conservative:   17
Best Local Similarity:    26.7%           Mismatches:     35
Query Match:              15.1%           Indels:         3
DB:                       9               Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-30209 (1-4148)

Qy          12 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
               |||:::|||:::         |||      |||||||||:::   :::   :::|||
Db        2120 CCCATGGAAAAGGACCCAGCAGAG------CCAGAAACAGTGCCTGATGGCGAAACCCCA 2173

Qy          32 ValAspLeuThr---AsnGluGluThrThrAspSerThrThrSerLysIleSerProSer 50
               |||           ||||||      |||:::::::::::|||:::
Db        2174 GAAGATGAAAATCCAACAGAGGAAGGAGCAGACAACTCTTCAGCAAAGATGGAAGAGGAG 2233

Qy          51 GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGly 70
               |||:::   ::::::|||                        |||      |||
Db        2234 GAGGAGGAAGAGGAAGAAGAAGAAGAGAGCCTCCCAGGATGTACTCTGTTTATTAAGAAT 2293

Qy          71 LeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 85
               |||:::   :::         |||:::   :::   |||   |||
Db        2294 CTCAATTTTGACACAACAGAAGAGAAGCTGAAGGAAGTGTTTTCA 2338

RESULT 34
US-60-751-420-2669
; Sequence 2669, Application US/60751420
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Fournier, Helene
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Debrus, Sophie
;  APPLICANT:  Van Berdewegh, Paul
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: GeneMap of the human genes associated with Crohn's disease
;  FILE REFERENCE: GENI-011/00US
```

```
;  CURRENT APPLICATION NUMBER: US/60/751,420
;  CURRENT FILING DATE:  2005-12-19
;  NUMBER OF SEQ ID NOS: 27266
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2669
;   LENGTH: 4148
;   TYPE: DNA
;   ORGANISM: Homosapiens
US-60-751-420-2669

Alignment Scores:
Pred. No.:                255             Length:         4148
Score:                    67.50           Matches:        20
Percent Similarity:       49.3%           Conservative:   17
Best Local Similarity:    26.7%           Mismatches:     35
Query Match:              15.1%           Indels:         3
DB:                       11              Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-60-751-420-2669 (1-4148)

Qy          12 ProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyr 31
               |||:::|||:::          |||       |||||||||:::   :::   :::|||
Db        2120 CCCATGGAAAAGGACCCAGCAGAG------CCAGAAACAGTGCCTGATGGCGAAACCCCA 2173

Qy          32 ValAspLeuThr---AsnGluGluThrThrAspSerThrThrSerLysIleSerProSer 50
                  |||          ||||||      |||:::::::::::::|||:::
Db        2174 GAAGATGAAAATCCAACAGAGGAAGGAGCAGACAACTCTTCAGCAAAGATGGAAGAGGAG 2233

Qy          51 GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGly 70
               |||:::   ::::::|||                                 |||      |||
Db        2234 GAGGAGGAAGAGGAAGAAGAAGAAGAGAGCCTCCCAGGATGTACTCTGTTTATTAAGAAT 2293

Qy          71 LeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 85
               |||:::   :::          |||:::   :::   |||   |||
Db        2294 CTCAATTTTGACACAACAGAAGAGAAGCTGAAGGAAGTGTTTTCA 2338

RESULT 35
US-11-331-019-39117
; Sequence 39117, Application US/11331019
; GENERAL INFORMATION:
;  APPLICANT: Fincher, Karen L.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  McCarter, David W.
;  APPLICANT:  Pear, Julie R.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.359 - 38-21(51375)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,019
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/637,086
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: US 60/149,881
;  PRIOR FILING DATE: 1999-08-19
;  NUMBER OF SEQ ID NOS: 52949
; SEQ ID NO 39117
;   LENGTH: 399
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3165-034-P1-K6-G5
US-11-331-019-39117

Alignment Scores:
Pred. No.:                11.1            Length:         399
Score:                    67.00           Matches:        20
Percent Similarity:       43.9%           Conservative:   9
Best Local Similarity:    30.3%           Mismatches:     35
Query Match:              15.0%           Indels:         2
DB:                       7               Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-019-39117 (1-399)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
                  :::   :::|||   |||:::      |||   |||   |||:::          ||||||
Db         145 GTCCACAAGGCCTTCAACAACAAGGTTGAGTTTCCGGAAGAAGATCAGTGCACTGAGCGG 204

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               |||          |||       |||       |||   :::   :::   ||||||
Db         205 TCCCCAACGGCAAACTCCCGAAAGAAGGTTCAGGTG---GTTGAGCATGCAGAGACCGCT 261

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               |||   |||       |||   ::::::          |||:::
Db         262 GATGAAGCCACCCTTGAT---TCTCAAGCTGATAGCTACATCCAGCAGAAACACAAAGCC 318

Qy          61 PheSerLeuIleThrTrp 66
               |||   |||       |||
Db         319 TTTGAACTGTGCAAATGG 336

RESULT 36
US-11-331-019-38689
; Sequence 38689, Application US/11331019
; GENERAL INFORMATION:
;  APPLICANT: Fincher, Karen L.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  McCarter, David W.
;  APPLICANT:  Pear, Julie R.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.359 - 38-21(51375)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,019
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/637,086
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: US 60/149,881
;  PRIOR FILING DATE: 1999-08-19
;  NUMBER OF SEQ ID NOS: 52949
; SEQ ID NO 38689
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3165-034-P1-K1-G5
US-11-331-019-38689

Alignment Scores:
Pred. No.:                11.2            Length:         400
Score:                    67.00           Matches:        20
Percent Similarity:       43.9%           Conservative:   9
Best Local Similarity:    30.3%           Mismatches:     35
Query Match:              15.0%           Indels:         2
DB:                       7               Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-019-38689 (1-400)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
                  :::   :::|||   |||:::      |||   |||   |||:::          ||||||
Db         145 GTCCACAAGGCCTTCAACAACAAGGTTGAGTTTCCGGAAGAAGATCAGTGCACTGAGCGG 204

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
               |||          |||       |||       |||   :::   :::   ||||||
Db         205 TCCCCAACGGCAAACTCCCGAAAGAAGGTTCAGGTG---GTTGAGCATGCAGAGACCGTT 261

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               |||   |||       |||   ::::::          |||:::
Db         262 GATGAAGCCACCCTTGAT---TCTCAAGCTGATAGCTACATCCAGCAGAAACACAAAGCC 318

Qy          61 PheSerLeuIleThrTrp 66
               |||   |||       |||
Db         319 TTTGAACTGTGCAAATGG 336

RESULT 37
US-11-331-019-37593
; Sequence 37593, Application US/11331019
; GENERAL INFORMATION:
;  APPLICANT: Fincher, Karen L.
;  APPLICANT:  La Rosa, Thomas J.
```

```
;  APPLICANT:  McCarter, David W.
;  APPLICANT:  Pear, Julie R.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.359 - 38-21(51375)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,019
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/637,086
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: US 60/149,881
;  PRIOR FILING DATE: 1999-08-19
;  NUMBER OF SEQ ID NOS: 52949
; SEQ ID NO 37593
;   LENGTH: 424
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3165-018-P1-K1-G8
US-11-331-019-37593

Alignment Scores:
Pred. No.:              12.1          Length:         424
Score:                  67.00         Matches:        20
Percent Similarity:     43.9%         Conservative:   9
Best Local Similarity:  30.3%         Mismatches:     35
Query Match:            15.0%         Indels:         2
DB:                     7             Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-019-37593 (1-424)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              :::   :::|||   |||:::       |||   |||   |||:::         ||||||
Db        170 GTCCACAAGGCCTTCAACAACAAGGTTGAGTTTCCGGAAGAAGATCAGTGCACTGAGCGG 229

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||         |||       |||       |||   :::   :::   ||||||
Db        230 TCCCCAACGGCAAACTCCCGAAAGAAGGTTCAGGTG---GTTGAGCATGCAGAGACCGTT 286

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||       |||         |||   ::::::         |||:::
Db        287 GATGAAGCCACCCTTGAT---TCTCAAGCTGATAGCTACATCCAGCAAAAACACAAAGCC 343

Qy         61 PheSerLeuIleThrTrp 66
              |||   |||       |||
Db        344 TTTGAACTGTGCAAATGG 361

RESULT 38
US-11-331-019-39617
; Sequence 39617, Application US/11331019
; GENERAL INFORMATION:
;  APPLICANT: Fincher, Karen L.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  McCarter, David W.
;  APPLICANT:  Pear, Julie R.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.359 - 38-21(51375)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,019
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/637,086
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: US 60/149,881
;  PRIOR FILING DATE: 1999-08-19
;  NUMBER OF SEQ ID NOS: 52949
; SEQ ID NO 39617
;   LENGTH: 442
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3165-042-Q1-K1-D2
US-11-331-019-39617

Alignment Scores:
Pred. No.:              12.9          Length:         442
Score:                  67.00         Matches:        20
Percent Similarity:     43.9%         Conservative:   9
Best Local Similarity:  30.3%         Mismatches:     35
Query Match:            15.0%         Indels:         2
DB:                     7             Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-019-39617 (1-442)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              :::   :::|||   |||:::       |||   |||   |||:::         ||||||
Db        155 GTCCACAAGGCCTTCAACAACAAGGTTGAGTTTCCGGAAGAAGATCAGTGCACTGAGCGG 214

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||         |||       |||       |||   :::   :::   ||||||
Db        215 TCCCCAACGGCAAACTCCCGAAAGAAGGTTCAGGTG---GTTGAGCATGCAGAGACCGTT 271

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||       |||         |||   ::::::         |||:::
Db        272 GATGAAGCCACCCTTGAT---TCTCAAGCTGATAGCTACATCCAGCAGAAACACAAAGCC 328

Qy         61 PheSerLeuIleThrTrp 66
              |||   |||       |||
Db        329 TTTGAACTGTGCAAATGG 346

RESULT 39
US-11-331-019-38858
; Sequence 38858, Application US/11331019
; GENERAL INFORMATION:
;  APPLICANT: Fincher, Karen L.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  McCarter, David W.
;  APPLICANT:  Pear, Julie R.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.359 - 38-21(51375)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,019
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/637,086
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: US 60/149,881
;  PRIOR FILING DATE: 1999-08-19
;  NUMBER OF SEQ ID NOS: 52949
; SEQ ID NO 38858
;   LENGTH: 453
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3165-034-P1-K3-G5
US-11-331-019-38858

Alignment Scores:
Pred. No.:              13.3          Length:         453
Score:                  67.00         Matches:        20
Percent Similarity:     43.9%         Conservative:   9
Best Local Similarity:  30.3%         Mismatches:     35
Query Match:            15.0%         Indels:         2
DB:                     7             Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-019-38858 (1-453)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              :::   :::|||   |||:::       |||   |||   |||:::         ||||||
Db        145 GTCCACAAGGCCTTCAACAACAAGGTTGAGTTTCCGGAAGAAGATCAGTGCACTGAGCGG 204

Qy         21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
              |||         |||       |||       |||   :::   :::   ||||||
Db        205 TCCCCAACGGCAAACTCCCGAAAGAAGGTTCAGGTG---GTTGAGCATGCAGAGACCGTT 261

Qy         41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
              |||       |||         |||   ::::::         |||:::
Db        262 GATGAAGCCACCCTTGAT---TCTCAAGCTGATAGCTACATCCAGCAGAAACACAAAGCC 318
```

```
Qy          61 PheSerLeuIleThrTrp 66
               |||   |||       |||
Db         319 TTTGAACTGTGCAAATGG 336

RESULT 40
US-11-266-748A-349404
; Sequence 349404, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 349404
;   LENGTH: 1047
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (364)..(479)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (543)..(543)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (990)..(990)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (995)..(995)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1013)..(1013)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1017)..(1017)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1021)..(1021)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1030)..(1030)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1044)..(1044)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-349404

Alignment Scores:
Pred. No.:                43.3          Length:          1047
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       9             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-349404 (1-1047)

Qy           5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
               :::|||:::   :::||||||      ||||||              |||
Db         715 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 774

Qy          25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
               |||                          ::::::||||||      |||:::|||
Db         775 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 810

Qy          45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
               |||::::::   |||   |||      ::::::||||||
Db         811 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 849

RESULT 41
US-11-266-748A-382318
; Sequence 382318, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 382318
;   LENGTH: 1047
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (364)..(479)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (543)..(543)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (990)..(990)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (995)..(995)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1013)..(1013)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1017)..(1017)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1021)..(1021)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1030)..(1030)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1044)..(1044)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-382318

Alignment Scores:
Pred. No.:                43.3          Length:         1047
Score:                    67.00         Matches:        16
Percent Similarity:       49.1%         Conservative:   10
Best Local Similarity:    30.2%         Mismatches:     19
Query Match:              15.0%         Indels:         8
DB:                       9             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-382318 (1-1047)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                  |||
Db        715 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 774

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                |||                        ::::::||||||      |||:::|||
Db        775 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 810

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        811 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 849

RESULT 42
US-11-266-748A-432783/c
; Sequence 432783, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 432783
;   LENGTH: 1047
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (27)..(27)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (31)..(31)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (35)..(35)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (53)..(53)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (58)..(58)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (505)..(505)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (569)..(684)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-432783

Alignment Scores:
Pred. No.:                43.3          Length:         1047
Score:                    67.00         Matches:        16
Percent Similarity:       49.1%         Conservative:   10
Best Local Similarity:    30.2%         Mismatches:     19
Query Match:              15.0%         Indels:         8
DB:                       9             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-432783 (1-1047)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                  |||
Db        333 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 274

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                |||                        ::::::||||||      |||:::|||
Db        273 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 238

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        237 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 199

RESULT 43
US-11-266-748A-349405
; Sequence 349405, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
```

```
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 349405
;   LENGTH: 1280
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44)..(44)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (52)..(52)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (54)..(54)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (196)..(196)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (304)..(304)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (307)..(307)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (366)..(479)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1070)..(1070)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1120)..(1120)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1134)..(1134)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1146)..(1146)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1166)..(1166)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1216)..(1216)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1223)..(1223)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1238)..(1238)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-349405

Alignment Scores:
Pred. No.:                57.4          Length:          1280
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       9             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-349405 (1-1280)

Qy           5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
               :::|||:::   :::||||||      ||||||                    |||
Db         714 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 773

Qy          25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
               |||                          ::::::||||||     |||:::|||
Db         774 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 809

Qy          45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
               |||::::::  |||  |||      ::::::||||||
Db         810 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 848

RESULT 44
US-11-266-748A-382319
; Sequence 382319, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 382319
;   LENGTH: 1280
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44)..(44)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (52)..(52)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (54)..(54)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (196)..(196)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (304)..(304)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (307)..(307)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (366)..(479)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1070)..(1070)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1120)..(1120)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1134)..(1134)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1146)..(1146)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1166)..(1166)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1216)..(1216)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1223)..(1223)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1238)..(1238)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-382319

Alignment Scores:
Pred. No.:                57.4          Length:          1280
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       9             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-382319 (1-1280)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                        |||
Db        714 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 773

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
               |||                          ::::::||||||      |||:::|||
Db        774 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 809

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        810 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 848

RESULT 45
US-11-266-748A-432784/c
; Sequence 432784, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 432784
;   LENGTH: 1280
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (43)..(43)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (58)..(58)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65)..(65)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (115)..(115)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (135)..(135)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (147)..(147)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(161)
;   OTHER INFORMATION: n is a, c, g, or t
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (211)..(211)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (802)..(915)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (974)..(974)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (977)..(977)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1085)..(1085)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1227)..(1227)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1229)..(1229)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1237)..(1237)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-432784

Alignment Scores:
Pred. No.:                57.4          Length:          1280
Score:                    67.00         Matches:         16
Percent Similarity:       49.1%         Conservative:    10
Best Local Similarity:    30.2%         Mismatches:      19
Query Match:              15.0%         Indels:          8
DB:                       9             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-432784 (1-1280)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::  :::||||||     ||||||                  |||
Db        567 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 508

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              |||                      ::::::||||||      |||:::|||
Db        507 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 472

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||     ::::::||||||
Db        471 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 433

RESULT 46
US-10-953-349-32751
; Sequence 32751, Application US/10953349
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCONDED THERBY
;  FILE REFERENCE: 2750-1579PUS2
;  CURRENT APPLICATION NUMBER: US/10/953,349
;  CURRENT FILING DATE:  2004-09-30
;  NUMBER OF SEQ ID NOS: 40252
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 32751
;   LENGTH: 1410
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
US-10-953-349-32751

Alignment Scores:
Pred. No.:                65.8          Length:          1410
Score:                    67.00         Matches:         19
Percent Similarity:       57.4%         Conservative:    16
Best Local Similarity:    31.1%         Mismatches:      20
Query Match:              15.0%         Indels:          6
DB:                       6             Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-953-349-32751 (1-1410)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              |||         :::|||      :::   |||   :::|||:::|||   |||::::::
Db        424 GAGCCGCAGGCCGACAGCTCAGACAAGAAGCCGGAGCAGGAGACGGCCTCGGAGAAGAAG 483

Qy         22 Pro------GluThrIleSerGlu------ProLysThrTyrValAspLeuThrAsnGlu 37
              |||      ||||||   ||||||      |||:::|||        :::      :::|||
Db        484 CCGGAGCAGGAGACGGCCTCGGAGAAGAAGCCTGAGACGGAGAAGGAAGCAGAGTCTGAG 543

Qy         38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              :::      |||        :::|||      ||||||::::::   :::::::::
Db        544 AAGAAGGCTGACAAGGAGGAGGCGAAA------CCAAGCGACGAGGCAAAGAAAGACGGC 597

Qy         58 Gly 58
              |||
Db        598 GGC 600

RESULT 47
US-60-748-114-12809
; Sequence 12809, Application US/60748114
; GENERAL INFORMATION:
;  APPLICANT: Boukharov, Andrey
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Kovalic, David
;  APPLICANT:  Li, Ping
;  APPLICANT:  Lu, Maolong
;  APPLICANT:  Lutfiyya, Linda
;  APPLICANT:  Wang, Qi
;  APPLICANT:  Wu, Wei
;  TITLE OF INVENTION: GENOME-WIDE IDENTIFICATION AND CHARACTERIZATION OF GENE EXPRESSIO
;  TITLE OF INVENTION:  REGULATORY ELEMENTS IN ZEA MAYS FOR USE IN PLANTS
;  FILE REFERENCE: 3821(54047)A
;  CURRENT APPLICATION NUMBER: US/60/748,114
;  CURRENT FILING DATE:  2005-12-07
;  NUMBER OF SEQ ID NOS: 14151
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12809
;   LENGTH: 1500
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: TATA-box: Start position=635
US-60-748-114-12809

Alignment Scores:
Pred. No.:                71.7          Length:          1500
Score:                    67.00         Matches:         19
Percent Similarity:       38.4%         Conservative:    19
Best Local Similarity:    19.2%         Mismatches:      49
Query Match:              15.0%         Indels:          12
DB:                       11            Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-60-748-114-12809 (1-1500)

Qy          1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
              ||||||      |||   :::|||   ||||||          |||:::   |||
Db         27 ATGGAGCCGAAAAATAATCTTCACTTCACCCCTCCAACCAACCCTTGTGCCGTTGTGCGG 86

Qy         21 ArgProGlu-----------------------------------ThrIleSerGluPro 28
              |||:::                                        :::
Db         87 ACTCCGCAAGGCCGGCGGAGCGCGGGAGCAAGCGAGAGGCAGAACCGCAGAACGCGGGCG 146

Qy         29 LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSer 48
```

```
            :::||||||           ::::::                    |||
Db        147 CGGACCTACACCTACTCCGTTCACGACTTTAAATATAGCTGTTGCGCACAAAAATTTTTT 206

Qy         49 ProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle 68
                :::     :::::::::        :::   :::|||:::   |||
Db        207 TCTAAAAAAATAAAAAGAAAGAAAGAAAAACACTTGAGAACGTTGGTAGGGTGGCGGAGC 266

Qy         69 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
               :::|||          ||||||   ::::::   :::|||         |||:::
Db        267 AATGGCTCGAGTTATTGGAATTTGGAAGACAAGCAGAAAGGCACGGTCTGCTACGTA 323

RESULT 48
US-10-342-276B-20
; Sequence 20, Application US/10342276B
; GENERAL INFORMATION:
;  APPLICANT: FUJIWARA, TSUTOMU
;  APPLICANT:  WATANABE, TAKESHI
;  APPLICANT:  HORIE, MASATO
;  TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;  TITLE OF INVENTION:  SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;  FILE REFERENCE: Q-53599
;  CURRENT APPLICATION NUMBER: US/10/342,276B
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: US/09/976,165
;  PRIOR FILING DATE: 2001-10-15
;  PRIOR APPLICATION NUMBER: 09/565,538
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: 09/055,699
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: 08/820,170
;  PRIOR FILING DATE: 1997-03-19
;  PRIOR APPLICATION NUMBER: JP 63410/1996
;  PRIOR FILING DATE: 1996-03-19
;  PRIOR APPLICATION NUMBER: JP 69163/1997
;  PRIOR FILING DATE: 1997-03-05
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 1518
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-276B-20

Alignment Scores:
Pred. No.:              73          Length:         1518
Score:                  67.00       Matches:        16
Percent Similarity:     49.1%       Conservative:   10
Best Local Similarity:  30.2%       Mismatches:     19
Query Match:            15.0%       Indels:         8
DB:                     6           Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-342-276B-20 (1-1518)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                     |||
Db        451 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 510

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                |||                        ::::::||||||      |||:::|||
Db        511 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 546

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        547 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 585

RESULT 49
US-60-762-056-44165/c
; Sequence 44165, Application US/60762056
; GENERAL INFORMATION:
;  APPLICANT: Trinklein, Nathan D
;  APPLICANT:  Aldred, Shelley Force
;  APPLICANT:  Cooper, Sara J
;  APPLICANT:  Myers, Richard M
;  TITLE OF INVENTION: FUNCTIONAL ARRAYS FOR HIGH THROUGHPUT CHARACTERIZATION OF GENE
;  TITLE OF INVENTION:  EXPRESSION REGULATORY ELEMENTS
;  FILE REFERENCE: 33102-701.102
;  CURRENT APPLICATION NUMBER: US/60/762,056
;  CURRENT FILING DATE:  2006-01-24
;  NUMBER OF SEQ ID NOS: 45096
; SEQ ID NO 44165
;   LENGTH: 2300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-762-056-44165

Alignment Scores:
Pred. No.:              131         Length:         2300
Score:                  67.00       Matches:        16
Percent Similarity:     49.1%       Conservative:   10
Best Local Similarity:  30.2%       Mismatches:     19
Query Match:            15.0%       Indels:         8
DB:                     11          Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-60-762-056-44165 (1-2300)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                     |||
Db        749 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 690

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                |||                        ::::::||||||      |||:::|||
Db        689 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 654

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        653 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 615

RESULT 50
US-11-266-748A-349401
; Sequence 349401, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 349401
;   LENGTH: 2625
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (29)..(29)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (36)..(36)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (356)..(471)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2147)..(2147)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2189)..(2190)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-349401

Alignment Scores:
Pred. No.:                158             Length:          2625
Score:                    67.00           Matches:         16
Percent Similarity:       49.1%           Conservative:    10
Best Local Similarity:    30.2%           Mismatches:      19
Query Match:              15.0%           Indels:          8
DB:                       9               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-349401 (1-2625)

Qy           5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
               :::|||:::   :::||||||      ||||||                  |||
Db         706 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 765

Qy          25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                |||                        ::::::||||||      |||:::|||
Db         766 TTCAGC-----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 801

Qy          45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
               |||::::::   |||  |||      ::::::||||||
Db         802 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 840

RESULT 51
US-11-266-748A-382315
; Sequence 382315, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 382315
;   LENGTH: 2625
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (29)..(29)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36)..(36)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (356)..(471)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2147)..(2147)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2189)..(2190)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-382315

Alignment Scores:
Pred. No.:                158             Length:          2625
Score:                    67.00           Matches:         16
Percent Similarity:       49.1%           Conservative:    10
Best Local Similarity:    30.2%           Mismatches:      19
Query Match:              15.0%           Indels:          8
DB:                       9               Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-382315 (1-2625)

Qy           5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
               :::|||:::   :::||||||      ||||||                  |||
Db         706 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 765

Qy          25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
                |||                        ::::::||||||      |||:::|||
Db         766 TTCAGC-----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 801

Qy          45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
               |||::::::   |||  |||      ::::::||||||
Db         802 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 840

RESULT 52
US-11-266-748A-432780/c
; Sequence 432780, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
```

```
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 432780
;   LENGTH: 2625
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (436)..(437)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (479)..(479)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2155)..(2270)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2590)..(2590)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2597)..(2597)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-432780

Alignment Scores:
Pred. No.:              158          Length:        2625
Score:                  67.00        Matches:       16
Percent Similarity:     49.1%        Conservative:  10
Best Local Similarity:  30.2%        Mismatches:    19
Query Match:            15.0%        Indels:        8
DB:                     9            Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-432780 (1-2625)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                  |||
Db       1920 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 1861

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              |||                          ::::::||||||      |||:::|||
Db       1860 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 1825

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db       1824 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 1786

RESULT 53
US-10-342-276B-21
; Sequence 21, Application US/10342276B
; GENERAL INFORMATION:
;  APPLICANT: FUJIWARA, TSUTOMU
;  APPLICANT:  WATANABE, TAKESHI
;  APPLICANT:  HORIE, MASATO
;  TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;  TITLE OF INVENTION:  SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;  FILE REFERENCE: Q-53599
;  CURRENT APPLICATION NUMBER: US/10/342,276B
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: US/09/976,165
;  PRIOR FILING DATE: 2001-10-15
;  PRIOR APPLICATION NUMBER: 09/565,538
;  PRIOR FILING DATE: 2000-05-05
;  PRIOR APPLICATION NUMBER: 09/055,699
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: 08/820,170
;  PRIOR FILING DATE: 1997-03-19
;  PRIOR APPLICATION NUMBER: JP 63410/1996
;  PRIOR FILING DATE: 1996-03-19
;  PRIOR APPLICATION NUMBER: JP 69163/1997
;  PRIOR FILING DATE: 1997-03-05
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 2636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (266)..(1783)
US-10-342-276B-21

Alignment Scores:
Pred. No.:              159          Length:        2636
Score:                  67.00        Matches:       16
Percent Similarity:     49.1%        Conservative:  10
Best Local Similarity:  30.2%        Mismatches:    19
Query Match:            15.0%        Indels:        8
DB:                     6            Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-342-276B-21 (1-2636)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              :::|||:::   :::||||||      ||||||                  |||
Db        716 ATCAATGCAGAATACGAGCCTACAGAAGAAGAATGTGAATGGAATTCAGAGGATGAGGAG 775

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              |||                          ::::::||||||      |||:::|||
Db        776 TTCAGC----------------------AGTGATGAGGAGGTGCAGGATAACACCCCT 811

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
              |||::::::   |||   |||      ::::::||||||
Db        812 AGTGAAATGCCTCCCTTAGAGGGTGAGGAAGAAGAAAAC 850

RESULT 54
US-11-331-032-8894/c
; Sequence 8894, Application US/11331032
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Xiao, Jinhua
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.361 - 38-21(51462)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,032
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/666,355
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US 60/155,005
;  PRIOR FILING DATE: 1999-09-21
;  NUMBER OF SEQ ID NOS: 25978
; SEQ ID NO 8894
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: uC-gsronu33b025a09b1
US-11-331-032-8894

Alignment Scores:
Pred. No.:              12.8         Length:        393
Score:                  66.50        Matches:       19
Percent Similarity:     47.1%        Conservative:  14
Best Local Similarity:  27.1%        Mismatches:    28
Query Match:            14.8%        Indels:        9
DB:                     7            Gaps:          1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-032-8894 (1-393)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              ||||||   :::|||       |||
Db        270 CTCAACCATTTCTTTCATCAACCC-------------------------AAACGCATC 238

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
```

```
                ||| ||:|||:::   :::   |||:::   :::   :::   |||||||||:::|||   |||
Db        237 CTCTCGAAGTTTCGTGGATACCTTTTCGTATACCACCTCGAAACAACCGAAAGTCTCACA 178

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
               |||   ::::::|||         |||             ||||||         :::
Db        177 AGCCCGTTGAACCCTCTTGCACAAACGGTTCCACCCTTCGGTTCCTCGCTTCAAATCCGA 118

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsn 74
                  :::::|||   :::||||||
Db        117 TGAACGGAACTCAATGGCTGGGAACTCAAT 88

RESULT 55
US-11-331-032-7835/c
; Sequence 7835, Application US/11331032
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Xiao, Jinhua
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.361 - 38-21(51462)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,032
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/666,355
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US 60/155,005
;  PRIOR FILING DATE: 1999-09-21
;  NUMBER OF SEQ ID NOS: 25978
; SEQ ID NO 7835
;   LENGTH: 445
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(445)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: uC-gsronu33B132c09b1
US-11-331-032-7835

Alignment Scores:
Pred. No.:                15.3          Length:          445
Score:                    66.50         Matches:         19
Percent Similarity:       47.1%         Conservative:    14
Best Local Similarity:    27.1%         Mismatches:      28
Query Match:              14.8%         Indels:          9
DB:                       7             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-032-7835 (1-445)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              ||||||   :::|||       |||
Db        257 CTCAACCATTTCTTTCATCAACCC-------------------------AAACGCATC 225

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              :::|||:::   :::   |||:::   :::   :::   |||||||||:::|||   |||
Db        224 CTCTCGAAGTTTCGTGGATACCTTTTCGTATACCACCTCGAAACAACCGAAAGTCTCACA 165

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              |||   ::::::|||         |||             ||||||         :::
Db        164 AGCCCGTTGAACCCTCTTGCACAAACGGTTCCACCCTTCGGTTCCTCGCTTCAAATCCGA 105

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsn 74
                  :::::|||   :::||||||
Db        104 TGAACGGAACTCAATGGCTGGGAACTCAAT 75

RESULT 56
US-11-331-032-20623/c
; Sequence 20623, Application US/11331032
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Xiao, Jinhua
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.361 - 38-21(51462)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,032
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/666,355
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US 60/155,005
;  PRIOR FILING DATE: 1999-09-21
;  NUMBER OF SEQ ID NOS: 25978
; SEQ ID NO 20623
;   LENGTH: 446
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(446)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: uC-gsflnu33B026b11b1
US-11-331-032-20623

Alignment Scores:
Pred. No.:                15.3          Length:          446
Score:                    66.50         Matches:         19
Percent Similarity:       47.1%         Conservative:    14
Best Local Similarity:    27.1%         Mismatches:      28
Query Match:              14.8%         Indels:          9
DB:                       7             Gaps:            1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-032-20623 (1-446)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              ||||||   :::|||       |||
Db        265 CTCAACCATTTCTTTCATCAACCC-------------------------AAACGCATC 233

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              :::|||:::   :::   |||:::   :::   :::   |||||||||:::|||   |||
Db        232 CTCTCGAAGTTTCGTGGATACCTTTTCGTATACCACCTCGAAACAACCGAAAGTCTCACA 173

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              |||   ::::::|||         |||             ||||||         :::
Db        172 AGCCCGTTGAACCCTCTTGCACAAACGGTTCCACCCTTCGGTTCCTCGCTTCAAATCCGA 113

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsn 74
                  :::::|||   :::||||||
Db        112 TGAACGGAACTCAATGGCTGGGAACTCAAT 83

RESULT 57
US-11-331-032-3347/c
; Sequence 3347, Application US/11331032
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Xiao, Jinhua
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.361 - 38-21(51462)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,032
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/666,355
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US 60/155,005
;  PRIOR FILING DATE: 1999-09-21
;  NUMBER OF SEQ ID NOS: 25978
; SEQ ID NO 3347
;   LENGTH: 570
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(570)
;   OTHER INFORMATION: unsure at all n locations
```

```
;   FEATURE:
;   OTHER INFORMATION: Clone ID: uC-gsronu33B117d05b1
US-11-331-032-3347

Alignment Scores:
Pred. No.:                21.6          Length:         570
Score:                    66.50         Matches:        19
Percent Similarity:       47.1%         Conservative:   14
Best Local Similarity:    27.1%         Mismatches:     28
Query Match:              14.8%         Indels:         9
DB:                       7             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-032-3347 (1-570)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              ||||||   :::|||       |||
Db        257 CTCAACCATTTCTTTCATCAACCC-------------------------AAACGCATC 225

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              :::|||:::   :::   |||:::  :::   :::   |||||||||:::|||   |||
Db        224 CTCTCGAAGTTTCGTGGATACCTTTTCGTATACCACCTCGAAACAACCGAAAGTCTCACA 165

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              |||   ::::::|||       |||          ||||||         :::
Db        164 AGCCCGTTGAACCCTCTTGCACAAACGGTTCCACCCTTCGGTTCCTCGCTTCAAATCCGA 105

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsn 74
                  ::::::|||   :::||||||
Db        104 TGAACGGAACTCAATGGCTGGGAACTCAAT 75

RESULT 58
US-11-266-748A-31062
; Sequence 31062, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31062
;   LENGTH: 7632
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-31062

Alignment Scores:
Pred. No.:                832           Length:         7632
Score:                    66.50         Matches:        21
Percent Similarity:       46.5%         Conservative:   12
Best Local Similarity:    29.6%         Mismatches:     27
Query Match:              14.8%         Indels:         11
DB:                       9             Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-31062 (1-7632)

Qy          3 ArgAlaLeuAsnSerTyr---PheGluProProValGluGluSerAlaLeuGluArgArg 21
              ||||||:::|||   |||           |||   :::||||||   |||
Db       4495 CGTGCAATCAACGTGTATGGAACCAGTGAGCCAAGCCAGGAGTCTGAACTCACAACGGTA 4554

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGlu--------- 38
                 |||         |||||||||      |||:::::::::::::::::::
Db       4555 GGAGAGAAACCTGAAGAGCCGAAGGATGAAGTGGAGGTGTCAGATGATGATGAGAAGGAG 4614

Qy         39 ---------------------ThrThrAspSerThrThrSerLysIleSerProSerGlu 51
                                       |||   :::|||      ||:::|||
Db       4615 CCCGAGGTTGATTACCGGACAGTGACAATCAATACTGAACAAAAAGTATCTGACTTCTAC 4674

Qy         52 AspThrGlnGlnGluAsnGlySerMetPheSer 62
              |||   ::::::      ||||||     |||
Db       4675 GACATTGAGGAGAGATTAGGATCCTGTTTGAGC 4707

RESULT 59
US-11-045-004-1
; Sequence 1, Application US/11045004
; GENERAL INFORMATION:
;  APPLICANT: BUCHRIESER, CARMEN
;  APPLICANT:  FRANGEUL, LIONEL
;  APPLICANT:  COUVE, ELISABETH
;  APPLICANT:  RUSNIOK, CHRISTOPHE
;  APPLICANT:  FSIHI, HAFIDA
;  APPLICANT:  DEHOUX, PIERRE
;  APPLICANT:  DUSSURGET, OLIVIER
;  APPLICANT:  CHETOUANI, FARID
;  APPLICANT:  NEDJARI, HAFED
;  APPLICANT:  GLASER, PHILIPPE
;  APPLICANT:  KUNST, FRANCK
;  APPLICANT:  COSSART, PASCALE
;  APPLICANT:  DANIELS, JUSTIN
;  APPLICANT:  GOEBEI, WERNER
;  APPLICANT:  KREFT, JURGEN
;  APPLICANT:  KUHN, MICHAEL
;  APPLICANT:  NG, EVA
;  APPLICANT:  VAZQUEZ-BOLAND, ANTONIO
;  APPLICANT:  DOMINGUEZ-BERNAL, GUSTAVO
;  APPLICANT:  GARRIDO-GARCIA, PATRICIA
;  APPLICANT:  TIERREZ-MARTINEZ, ALBERTO
;  APPLICANT:  AMEND, ALEXANDRA
;  APPLICANT:  CHAKRABORTY, TRINAD
;  APPLICANT:  DOMANN, EUGEN
;  APPLICANT:  HAIN, THORSTEN
;  APPLICANT:  BERCHE, PATRICK
;  APPLICANT:  CHARBIT, ALAIN
;  APPLICANT:  DURANT, LIONEL
;  APPLICANT:  PEREZ-DIAZ, JOSE-CLAUDIO
;  APPLICANT:  BAQUERO, FERNANDO
;  APPLICANT:  GARCIA DEL PORTILLO, FRANCISCO
;  APPLICANT:  GOMEZ-LOPEZ, NURIA
;  APPLICANT:  MADUENIO, ENCARNA
;  APPLICANT:  PABLOS, BETRIZ DE
;  APPLICANT:  WEHLAND, JURGEN
;  APPLICANT:  KARST, UWE
;  APPLICANT:  ENTIAN, KARL-DIETER
;  APPLICANT:  HAUF, JORG
;  APPLICANT:  ROSE, MATTHIAS
;  APPLICANT:  VOSS, HAMUT
;  TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
;  FILE REFERENCE: 05394.0018-02
;  CURRENT APPLICATION NUMBER: US/11/045,004
;  CURRENT FILING DATE:  2005-01-28
;  PRIOR APPLICATION NUMBER: 10/637,657
;  PRIOR FILING DATE: 2003-08-11
;  PRIOR APPLICATION NUMBER: 10/257,023
;  PRIOR FILING DATE: 2002-10-08
;  PRIOR APPLICATION NUMBER: PCT/FR01/01118
;  PRIOR FILING DATE: 2001-04-11
```

```
;  PRIOR APPLICATION NUMBER: FR 00/04,629
;  PRIOR FILING DATE: 2000-04-11
;  NUMBER OF SEQ ID NOS: 2854
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
;   LENGTH: 2944528
;   TYPE: DNA
;   ORGANISM: Listeria monocytogenes
US-11-045-004-1

Alignment Scores:
Pred. No.:                3.03e+06        Length:          2944528
Score:                    66.50           Matches:         21
Percent Similarity:       50.7%           Conservative:    14
Best Local Similarity:    30.4%           Mismatches:      29
Query Match:              14.8%           Indels:          5
DB:                       9               Gaps:            2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-045-004-1 (1-2944528)

Qy          14 GluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAsp 33
               |||      |||:::         :::|||:::      ||||||||||:::|||
Db      128323 GAAATTAGAGTGTTAAAAAATGTAGATGATACAGTTTTCTATCCGAAGACACATGTA--- 128379

Qy          34 LeuThrAsnGluGluThrThrAspSer---------ThrThrSerLysIleSerProSer 50
               |||      |||    ||||||   |||||||||   :::   |||:::
Db      128380 ---ACGGCTGTGGAAGGTTTGGACTCGGCTACAACTACTACATCTGGATTAATGCCGGCC 128436

Qy          51 GluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGly 70
               |||      :::   ||||||        :::          |||:::
Db      128437 AGCGACAAAACGAAATTAAATGGAATCGAAGCTAATGCAGAAAAAAACAATGTGACTGCA 128496

Qy          71 LeuAspLeuAsnAsnLeuSerGluArg 79
               :::|||:::   |||   :::::::::
Db      128497 ATCGATATTGCCAATTGGAATAAAAAA 128523

RESULT 60
US-11-331-032-2004
; Sequence 2004, Application US/11331032
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Xiao, Jinhua
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.361 - 38-21(51462)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,032
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/666,355
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US 60/155,005
;  PRIOR FILING DATE: 1999-09-21
;  NUMBER OF SEQ ID NOS: 25978
; SEQ ID NO 2004
;   LENGTH: 371
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(371)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: uC-gsronu33B079c03b1
US-11-331-032-2004

Alignment Scores:
Pred. No.:                13.9            Length:          371
Score:                    66.00           Matches:         20
Percent Similarity:       41.5%           Conservative:    7
Best Local Similarity:    30.8%           Mismatches:      20
Query Match:              14.7%           Indels:          18
DB:                       7               Gaps:            3

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-032-2004 (1-371)

Qy          10 GluProProValGluGlu---SerAlaLeuGluArgArgProGluThrIleSerGluPro 28
               :::||||||   :::|||            |||:::|||||:::          |||
Db         117 CAACCCCCAGCCAAAGAACCAAAAACAAAAGAAAAACGCCCAAAAAAAAACCAAACCCCA 176

Qy          29 LysThrTyrValAspLeuThr---------------------------------------- 35
               |||||||||           |||
Db         177 AAAACATACAAACAACACACAAAACAAAAAACCAACAAATAAAAACACCAAGCAAAAAAA 236

Qy          36 ---------AsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 52
                        |||   :::|||||||      |||   :::|||      |||:::
Db         237 AAACACGCCAACCACAAAACAACCAAGAAAACCCCCGCAAAA---CCCCCAACCCCAAAA 293

Qy          53 ThrGlnGlnGluAsn 57
               ||||||   |||
Db         294 GCCCAACAAACAAAC 308

RESULT 61
US-11-330-364-5748/c
; Sequence 5748, Application US/11330364
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S.
;  APPLICANT:  Andersen, Scott E.
;  APPLICANT:  Dubois, Patrice
;  APPLICANT:  Mahadeo, Debbie A.
;  APPLICANT:  Masucci, James D.
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.352 - 38-21(51721)C/US
;  CURRENT APPLICATION NUMBER: US/11/330,364
;  CURRENT FILING DATE:  2006-01-12
;  PRIOR APPLICATION NUMBER: US 09/696,664
;  PRIOR FILING DATE: 2000-10-25
;  PRIOR APPLICATION NUMBER: US 60/161,619
;  PRIOR FILING DATE: 1999-10-26
;  NUMBER OF SEQ ID NOS: 17472
; SEQ ID NO 5748
;   LENGTH: 382
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(382)
;   OTHER INFORMATION: unsure at all n locations
;   OTHER INFORMATION: Clone ID: LIB3279-053-P1-K1-B9
US-11-330-364-5748

Alignment Scores:
Pred. No.:                14.5            Length:          382
Score:                    66.00           Matches:         23
Percent Similarity:       43.2%           Conservative:    12
Best Local Similarity:    28.4%           Mismatches:      34
Query Match:              14.7%           Indels:          12
DB:                       7               Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x US-11-330-364-5748 (1-382)

Qy          11 ProProValGluGluSerAlaLeu--------------GluArgArgProGluThrIle 25
               ||||||:::      ||||||                 ::::::::::|||
Db         312 CCGCCCATCAGCTCTTCCGCCACGCCGTGGTTCACCAGCCAGAAGAAGCCCACCGTCCCG 253

Qy          26 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp---SerThrThr 44
               :::      |||        ::::::          ||||||   |||||||||
Db         252 GCCGCGTGCCTCACCTGCTCGACTAGGTCCTCGCGCCGCGTCACCGACAGGTCCACCACC 193

Qy          45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer------MetPheSer 62
                  :::|||      |||              |||:::      :::   |||
Db         192 GGGATGGCGGCGCCGGAGGGCGACGACGACGGCAGAGGCGGCGCCGGAGAAATTATGTCG 133

Qy          63 LeuIleThr------------TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 78
               |||                |||   :::|||||||:::|||||||          |||
Db         132 CTCAGGGCTTCCGGCGGGTGGTGGAAGATGGACGGGATGGACTTGACGCCGGAGTCCACC 73
```

```
Qy          79 Arg 79
               |||
Db          72 AGG 70

RESULT 62
US-11-331-032-15190
; Sequence 15190, Application US/11331032
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Xiao, Jinhua
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.361 - 38-21(51462)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,032
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/666,355
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US 60/155,005
;  PRIOR FILING DATE: 1999-09-21
;  NUMBER OF SEQ ID NOS: 25978
; SEQ ID NO 15190
;   LENGTH: 550
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(550)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: uC-gsflnu33B063b06b1
US-11-331-032-15190

Alignment Scores:
Pred. No.:              24.2          Length:         550
Score:                  66.00         Matches:        20
Percent Similarity:     43.9%         Conservative:   9
Best Local Similarity:  30.3%         Mismatches:     35
Query Match:            14.7%         Indels:         2
DB:                     7             Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-032-15190 (1-550)

Qy           1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
               :::   :::|||   |||:::      |||   |||   |||:::          ||||||
Db          25 GTCCACAAGGCCTTCAACAACAAGGTTGAGTTTCCGGATGAAGATCAGTGCACTGAGCGG 84

Qy          21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
                 |||         |||      |||      |||   :::   :::   ||||||
Db          85 TCCCCAACGGCAAACTCCCGAAAGAAGGTTCAGGTG---GTTGAGCATGCAGAGACCGTT 141

Qy          41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
               |||       |||         |||   ::::::          |||:::
Db         142 GATGAAGCCACCCTTGAT---TCTCAAGCTGATAGCTACATCCAGCAGAAACACAAAGCC 198

Qy          61 PheSerLeuIleThrTrp 66
               |||   |||       |||
Db         199 TTTGAACTGTGCAAATGG 216

RESULT 63
US-10-461-673-3010
; Sequence 3010, Application US/10461673
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Xue, Aidong
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Zhi Wei
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Goodrich, Ryle W.
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Yamazaki, Victoria
;  APPLICANT:  Ujwal, Manusha L.
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Chen, Rui-Hong
;  APPLICANT:  Ghosh, Malabika
;  APPLICANT:  Weng, Gezhi
;  APPLICANT:  Haley-Vicente, Dana
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;  FILE REFERENCE: 823
;  CURRENT APPLICATION NUMBER: US/10/461,673
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/US02/29964
;  PRIOR FILING DATE: 2002-09-19
;  PRIOR APPLICATION NUMBER: US 10/245,014
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,739
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: PCT/US02/29636
;  PRIOR FILING DATE: 2002-09-18
;  PRIOR APPLICATION NUMBER: US 10/245,817
;  PRIOR FILING DATE: 2002-09-16
;  PRIOR APPLICATION NUMBER: US 60/323,349
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: PCT/US02/29001
;  PRIOR FILING DATE: 2002-09-13
;  PRIOR APPLICATION NUMBER: US 10/243,552
;  PRIOR FILING DATE: 2002-09-12
;  PRIOR APPLICATION NUMBER: US 60/322,511
;  PRIOR FILING DATE: 2001-09-13
;  PRIOR APPLICATION NUMBER: PCT/US02/25485
;  PRIOR FILING DATE: 2002-08-09
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 17116
;  SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 3010
;   LENGTH: 1632
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (107)..(976)
US-10-461-673-3010

Alignment Scores:
Pred. No.:              112           Length:         1632
Score:                  66.00         Matches:        18
Percent Similarity:     49.0%         Conservative:   7
Best Local Similarity:  35.3%         Mismatches:     19
Query Match:            14.7%         Indels:         7
DB:                     6             Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-461-673-3010 (1-1632)

Qy          23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSer 42
               :::   :::::::::      |||      ||||||  ||||||          |||
Db         722 CAGAACTTGGCACAGACTTATACTCTTACTGATCTGATGAATGAGATCAAAGAAAGTAGT 781

Qy          43 ThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer 62
               |||      ||||||  ||||||  |||:::|||   :::
Db         782 ACAAAACTCAAGATTTTTCCTTCCTCAGACTCTCAGCTGGACCCA-------------- 826

Qy          63 LeuIleThr-TrpAsnIleAspGlyLeuAsp 72
                  :::    ||||||     ||||||||
Db         827 ---TTGGTGGTGGAACTTTCAAGGCCTGGAC 854

RESULT 64
US-10-953-349-36708
```

```
; Sequence 36708, Application US/10953349
; GENERAL INFORMATION:
;  APPLICANT: ALEXANDROV, Nickolai et al.
;  TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
;  TITLE OF INVENTION:  ENCONDED THERBY
;  FILE REFERENCE: 2750-1579PUS2
;  CURRENT APPLICATION NUMBER: US/10/953,349
;  CURRENT FILING DATE:  2004-09-30
;  NUMBER OF SEQ ID NOS: 40252
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36708
;   LENGTH: 1687
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1587)..(1588)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-953-349-36708

Alignment Scores:
Pred. No.:                117             Length:          1687
Score:                    66.00           Matches:         15
Percent Similarity:       50.0%           Conservative:    12
Best Local Similarity:    27.8%           Mismatches:      23
Query Match:              14.7%           Indels:          4
DB:                       6               Gaps:            2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-953-349-36708 (1-1687)

Qy          2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
              :::|||   |||:::      :::   ||||||           |||   :::||||||
Db        182 CAGCGCGTCCTCAGCCAGCGCTGGAGGCCGCCC------AGCCGCGCCTGCGACCGCAGA 235

Qy         22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
              |||      :::::::::|||        :::                     |||
Db        236 CCGCGCGTCCTCACTCAGCCAGCGCTGGAGATCTCCTGTCCGCCCAGCCGCGCCTGCGAC 295

Qy         42 SerThrThrSerLysIleSerProSer------GluAspThr 53
              |||   |||:::   :::   |||:::      :::|||||
Db        296 AGTGCGACCGCCCTCCTGCCCCCAGCACTGCAAGATGATACT 337

RESULT 65
US-08-765-108-5
; Sequence 5, Application US/08765108
;  GENERAL INFORMATION:
;    APPLICANT:  Massachusetts Institute of Technology
;    TITLE OF INVENTION:  Class BI and CI Scavenger Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/765,108
;      FILING DATE:  27-MAR-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07721
;      FILING DATE:  19-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  MIT6620
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2032 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc. feature
;      LOCATION:  40..1926
;      OTHER INFORMATION:  /Function = "Nucleotides 40 through
;      OTHER INFORMATION:  1926 encode the amino acid sequence for the Drosophila
;      OTHER INFORMATION:  Melanogaster Scavenger Receptor Class CI."
US-08-765-108-5

Alignment Scores:
Pred. No.:                152             Length:          2032
Score:                    66.00           Matches:         22
Percent Similarity:       41.0%           Conservative:    12
Best Local Similarity:    26.5%           Mismatches:      19
Query Match:              14.7%           Indels:          30
DB:                       4               Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x US-08-765-108-5 (1-2032)

Qy         20 ArgArgProGluThrIleSer---------GluProLysThrTyrValAspLeuThrAsn 36
              :::::::|||   |||   |||          :::|||   |||              |||
Db       1384 AAAAAACCGACAACAACTTCAACAACGCCGAAGCCAACAACAACGACTTCAACCACACCA 1443

Qy         37 GluGluThrThrAspSerThrThrSerLysIleSer---------------------- 48
              :::   ||||||   ||||||||||||     |||
Db       1444 AAGTCTACAACT---TCTACAACGTCTACAACTTCAACAACACCAACGACAACAACTACA 1500

Qy         49 ------------------------------------ProSerGluAspThrGlnGln 55
                                                   |||:::      |||::::::
Db       1501 ATAAATGTGTTTACAACAAAGAAAACAACAATAATGATCCCTACTTCCAGTACCGAAAAG 1560

Qy         56 GluAsnGlySerMetPheSerLeu--------------IleThrTrpAsnIleAspGly 70
              |||   :::   ::::::               ||||||||||||:::|||
Db       1561 ACTACAGGCATCATCACCACCATGAAGACACGCAAGCGCATCACTTGGAACGTTGATCCT 1620

Qy         71 LeuAspLeu 73
              |||:::
Db       1621 CAGGACATC 1629

RESULT 66
US-08-765-108-5
; Sequence 5, Application US/08765108
;  GENERAL INFORMATION:
;    APPLICANT:  Massachusetts Institute of Technology
;    TITLE OF INVENTION:  Class BI and CI Scavenger Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/765,108
;      FILING DATE:  27-MAR-1997
```

```
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07721
;      FILING DATE:  19-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  MIT6620
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2032 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc. feature
;      LOCATION:  40..1926
;      OTHER INFORMATION:  /Function = "Nucleotides 40 through
;      OTHER INFORMATION:  1926 encode the amino acid sequence for the Drosophila
;      OTHER INFORMATION:  Melanogaster Scavenger Receptor Class CI."
US-08-765-108-5

Alignment Scores:
Pred. No.:                152             Length:         2032
Score:                    66.00           Matches:        22
Percent Similarity:       41.0%           Conservative:   12
Best Local Similarity:    26.5%           Mismatches:     19
Query Match:              14.7%           Indels:         30
DB:                       4               Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-08-765-108-5 (1-2032)

Qy          20 ArgArgProGluThrIleSer---------GluProLysThrTyrValAspLeuThrAsn 36
               ::::::|||   |||   |||         :::|||   |||             |||
Db        1384 AAAAAACCGACAACAACTTCAACAACGCCGAAGCCAACAACAACGACTTCAACCACACCA 1443

Qy          37 GluGluThrThrAspSerThrThrSerLysIleSer---------------------- 48
               :::   ||||||   ||||||||||||      |||
Db        1444 AAGTCTACAACT---TCTACAACGTCTACAACTTCAACAACACCAACGACAACAACTACA 1500

Qy          49 ------------------------------------ProSerGluAspThrGlnGln 55
                                                   |||:::      |||::::::
Db        1501 ATAAATGTGTTTACAACAAAGAAAACAACAATAATGATCCCTACTTCCAGTACCGAAAAG 1560

Qy          56 GluAsnGlySerMetPheSerLeu--------------IleThrTrpAsnIleAspGly 70
                 |||   :::   ::::::               ||||||||||||:::|||
Db        1561 ACTACAGGCATCATCACCACCATGAAGACACGCAAGCGCATCACTTGGAACGTTGATCCT 1620

Qy          71 LeuAspLeu 73
                 |||:::
Db        1621 CAGGACATC 1629

RESULT 67
US-08-765-108-5
; Sequence 5, Application US/08765108
;  GENERAL INFORMATION:
;    APPLICANT:  Massachusetts Institute of Technology
;    TITLE OF INVENTION:  Class BI and CI Scavenger Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center; 1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/765,108
;      FILING DATE:  27-MAR-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07721
;      FILING DATE:  19-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  MIT6620
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2032 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc. feature
;      LOCATION:  40..1926
;      OTHER INFORMATION:  /Function = "Nucleotides 40 through
;      OTHER INFORMATION:  1926 encode the amino acid sequence
;      OTHER INFORMATION:  for the Drosophila Melanogaster
;      OTHER INFORMATION:  Scavenger  Receptor Class CI."
US-08-765-108-5

Alignment Scores:
Pred. No.:                152             Length:         2032
Score:                    66.00           Matches:        22
Percent Similarity:       41.0%           Conservative:   12
Best Local Similarity:    26.5%           Mismatches:     19
Query Match:              14.7%           Indels:         30
DB:                       4               Gaps:           4

US-10-757-745-2_COPY_54_140 (1-87) x US-08-765-108-5 (1-2032)

Qy          20 ArgArgProGluThrIleSer---------GluProLysThrTyrValAspLeuThrAsn 36
               ::::::|||   |||   |||         :::|||   |||             |||
Db        1384 AAAAAACCGACAACAACTTCAACAACGCCGAAGCCAACAACAACGACTTCAACCACACCA 1443

Qy          37 GluGluThrThrAspSerThrThrSerLysIleSer---------------------- 48
               :::   ||||||   ||||||||||||      |||
Db        1444 AAGTCTACAACT---TCTACAACGTCTACAACTTCAACAACACCAACGACAACAACTACA 1500

Qy          49 ------------------------------------ProSerGluAspThrGlnGln 55
                                                   |||:::      |||::::::
Db        1501 ATAAATGTGTTTACAACAAAGAAAACAACAATAATGATCCCTACTTCCAGTACCGAAAAG 1560

Qy          56 GluAsnGlySerMetPheSerLeu--------------IleThrTrpAsnIleAspGly 70
                 |||   :::   ::::::               ||||||||||||:::|||
Db        1561 ACTACAGGCATCATCACCACCATGAAGACACGCAAGCGCATCACTTGGAACGTTGATCCT 1620

Qy          71 LeuAspLeu 73
                 |||:::
Db        1621 CAGGACATC 1629

RESULT 68
US-11-266-748A-56767/c
; Sequence 56767, Application US/11266748A
; GENERAL INFORMATION:
;  APPLICANT: Harkin, Paul
;  APPLICANT:  Johnston, Patrick
;  APPLICANT:  Mulligan, Karl
```

```
;  TITLE OF INVENTION: Transcriptome Microarray Technology and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 55815-0102 (319189)
;  CURRENT APPLICATION NUMBER: US/11/266,748A
;  CURRENT FILING DATE:  2005-11-03
;  PRIOR APPLICATION NUMBER: EP 04105479.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105482.6
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105483.4
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105507.0
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105485.9
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: EP 04105484.2
;  PRIOR FILING DATE: 2004-11-03
;  PRIOR APPLICATION NUMBER: US 60/662,276
;  PRIOR FILING DATE: 2005-03-14
;  PRIOR APPLICATION NUMBER: US 60/700,293
;  PRIOR FILING DATE: 2005-07-18
;  NUMBER OF SEQ ID NOS: 483996
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56767
;   LENGTH: 4775
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-11-266-748A-56767

Alignment Scores:
Pred. No.:                506             Length:         4775
Score:                    66.00           Matches:        20
Percent Similarity:       45.1%           Conservative:   21
Best Local Similarity:    22.0%           Mismatches:     28
Query Match:              14.7%           Indels:         22
DB:                       9               Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-11-266-748A-56767 (1-4775)

Qy          6 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle 25
              :::|||::::::   |||   |||:::   |||
Db        681 CATTCTCATTATTCTCCTAAAGTAAAATTTTCTCATTTTCTGTATCTGACGAGCTGCGGC 622

Qy         26 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 45
              ::::::||||||   |||::::::::::::  ::::::       ||||||   :::|||
Db        621 AATGACCCAAAAGCATATTTGAATATATCCTGTGATAAGTACTGCGACAGCAATTCCAGC 562

Qy         46 LysIleSerPro---------------SerGluAspThrGlnGlnGlu------------ 56
              ::::::                    :::   |||   ||||||
Db        561 ATAGTAACTGGCAAATTTGATCATTTCGCTCTCGATGTTCAGCAACCCACAACGTGTTCC 502

Qy         57 -------------------------------------AsnGlySerMetPheSerLeu 63
                                                    ||||||   :::   :::
Db        501 ATTTGTCATGTTCTGGTTGAGGGAACTGTTAGTCCATACAATGGTGTTATTCACACATGC 442

Qy         64 IleThrTrpAsnIleAspGlyLeuAspLeuAsn 74
              :::|||||||::::::   |||   ||||||
Db        441 TTTTCCTGGAATCTGGAGTTCTTGTAACTCAAC 409

RESULT 69
US-60-742-219-3003/c
; Sequence 3003, Application US/60742219
; GENERAL INFORMATION:
;  APPLICANT: Belouchi, Abdelmajid
;  APPLICANT:  Raelson, John Verner
;  APPLICANT:  Bradley, Walter Edward
;  APPLICANT:  Paquin, Bruno
;  APPLICANT:  Nguyen-Huu, Quynh
;  APPLICANT:  Croteau, Pascal
;  APPLICANT:  Allard, Rene
;  APPLICANT:  Little, Randall David
;  APPLICANT:  Keith, Tim
;  APPLICANT:  Cousineau, Johanne
;  APPLICANT:  Eerdewegh, Paul Van
;  APPLICANT:  Segal, Jonathan
;  TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
;  FILE REFERENCE: 306522-2000
;  CURRENT APPLICATION NUMBER: US/60/742,219
;  CURRENT FILING DATE:  2005-12-05
;  NUMBER OF SEQ ID NOS: 7303
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3003
;   LENGTH: 4775
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-742-219-3003

Alignment Scores:
Pred. No.:                506             Length:         4775
Score:                    66.00           Matches:        20
Percent Similarity:       45.1%           Conservative:   21
Best Local Similarity:    22.0%           Mismatches:     28
Query Match:              14.7%           Indels:         22
DB:                       11              Gaps:           2

US-10-757-745-2_COPY_54_140 (1-87) x US-60-742-219-3003 (1-4775)

Qy          6 AsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle 25
              :::|||::::::   |||   |||:::   |||
Db        681 CATTCTCATTATTCTCCTAAAGTAAAATTTTCTCATTTTCTGTATCTGACGAGCTGCGGC 622

Qy         26 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 45
              ::::::||||||   |||::::::::::::  ::::::       ||||||   :::|||
Db        621 AATGACCCAAAAGCATATTTGAATATATCCTGTGATAAGTACTGCGACAGCAATTCCAGC 562

Qy         46 LysIleSerPro---------------SerGluAspThrGlnGlnGlu------------ 56
              ::::::                    :::   |||   ||||||
Db        561 ATAGTAACTGGCAAATTTGATCATTTCGCTCTCGATGTTCAGCAACCCACAACGTGTTCC 502

Qy         57 -------------------------------------AsnGlySerMetPheSerLeu 63
                                                    ||||||   :::   :::
Db        501 ATTTGTCATGTTCTGGTTGAGGGAACTGTTAGTCCATACAATGGTGTTATTCACACATGC 442

Qy         64 IleThrTrpAsnIleAspGlyLeuAspLeuAsn 74
              :::|||||||::::::   |||   ||||||
Db        441 TTTTCCTGGAATCTGGAGTTCTTGTAACTCAAC 409

RESULT 70
US-11-331-032-22187/c
; Sequence 22187, Application US/11331032
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Xiao, Jinhua
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.361 - 38-21(51462)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,032
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/666,355
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US 60/155,005
;  PRIOR FILING DATE: 1999-09-21
;  NUMBER OF SEQ ID NOS: 25978
; SEQ ID NO 22187
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3272-009-P1-K1-B12
US-11-331-032-22187

Alignment Scores:
Pred. No.:                18.3            Length:         401
Score:                    65.50           Matches:        19
Percent Similarity:       44.3%           Conservative:   12
```

```
Best Local Similarity:  27.1%          Mismatches:     30
Query Match:            14.6%          Indels:         9
DB:                     7              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-032-22187 (1-401)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              ||||||   :::|||       |||
Db        250 CTCAACCATTTCTTTCATCAACCC-------------------------AAACGCATC 218

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              :::||||||   :::   |||      :::   :::   |||||||||:::|||   |||
Db        217 CTCTCGGAGTTTCGTCGATACCCTTTCATATACCACCTCGAAACAACCGAAAGTCTCACA 158

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              |||   ::::::|||                                ||||||         :::
Db        157 AGCCTCTTGAACCCTCTTGCACAACCGGTTCCATCCTTCGGTTCCTCGCTTCAAATCCGA 98

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsn 74
                      :::::::|||   :::||||||
Db         97 TGGACGGAACTCAATGGCTGGGAACTCAAT 68

RESULT 71
US-11-331-032-21617/c
; Sequence 21617, Application US/11331032
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Xiao, Jinhua
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.361 - 38-21(51462)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,032
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/666,355
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US 60/155,005
;  PRIOR FILING DATE: 1999-09-21
;  NUMBER OF SEQ ID NOS: 25978
; SEQ ID NO 21617
;   LENGTH: 408
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3272-002-P1-K1-B11
US-11-331-032-21617

Alignment Scores:
Pred. No.:              18.7           Length:         408
Score:                  65.50          Matches:        19
Percent Similarity:     44.3%          Conservative:   12
Best Local Similarity:  27.1%          Mismatches:     30
Query Match:            14.6%          Indels:         9
DB:                     7              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-032-21617 (1-408)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              ||||||   :::|||       |||
Db        257 CTCAACCATTTCTTTCATCAACCC-------------------------AAACGCATC 225

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              :::||||||   :::   |||      :::   :::   |||||||||:::|||   |||
Db        224 CTCTCGGAGTTTCGTCGATACCCTTTCATATACCACCTCGAAACAACCGAAAGTCTCACA 165

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              |||   ::::::|||                                ||||||         :::
Db        164 AGCCTCTTGAACCCTCTTGCACAACCGGTTCCATCCTTCGGTTCCTCGCTTCAAATCCGA 105

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsn 74
                      :::::::|||   :::||||||
Db        104 TGGACGGAACTCAATGGCTGGGAACTCAAT 75

RESULT 72
US-11-331-032-11044/c
; Sequence 11044, Application US/11331032
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Xiao, Jinhua
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 16517.361 - 38-21(51462)C/US
;  CURRENT APPLICATION NUMBER: US/11/331,032
;  CURRENT FILING DATE:  2006-01-13
;  PRIOR APPLICATION NUMBER: US 09/666,355
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US 60/155,005
;  PRIOR FILING DATE: 1999-09-21
;  NUMBER OF SEQ ID NOS: 25978
; SEQ ID NO 11044
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   OTHER INFORMATION: Clone ID: uC-gsflmaxxa067f10b1
US-11-331-032-11044

Alignment Scores:
Pred. No.:              23.9           Length:         485
Score:                  65.50          Matches:        19
Percent Similarity:     44.3%          Conservative:   12
Best Local Similarity:  27.1%          Mismatches:     30
Query Match:            14.6%          Indels:         9
DB:                     7              Gaps:           1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-331-032-11044 (1-485)

Qy          5 LeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThr 24
              ||||||   :::|||       |||
Db        254 CTCAACCATTTCTTTCATCAACCC-------------------------AAACGCATC 222

Qy         25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
              :::||||||   :::   |||      :::   :::   |||||||||:::|||   |||
Db        221 CTCTCGGAGTTTCGTCGATACCCTTTCATATACCACCTCGAAACAACCGAAAGTCTCACA 162

Qy         45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
              |||   ::::::|||                                ||||||         :::
Db        161 AGCCTCTTGAACCCTCTTGCACAACCGGTTCCATCCTTCGGTTCCTCGCTTCAAATCCGA 102

Qy         65 ThrTrpAsnIleAspGlyLeuAspLeuAsn 74
                      :::::::|||   :::||||||
Db        101 TGGACGGAACTCAATGGCTGGGAACTCAAT 72

RESULT 73
US-10-917-503B-620
; Sequence 620, Application US/10917503B
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
```

```
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 620
;   LENGTH: 855
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (826)..(826)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (848)..(848)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (853)..(853)
;   OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
US-10-917-503B-620

Alignment Scores:
Pred. No.:                53              Length:          855
Score:                    65.50           Matches:         18
Percent Similarity:       56.2%           Conservative:    23
Best Local Similarity:    24.7%           Mismatches:      23
Query Match:              14.6%           Indels:          9
DB:                       6               Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x US-10-917-503B-620 (1-855)

Qy          12 ProValGluGluSerAlaLeuGluArgArgPro---GluThrIleSerGluProLys--- 29
               |||:::          :::|||||||   :::|||       :::::::::||||||
Db         354 CCATTGGCAAATGAATCCTTGGAGCACAAACCTGTATCCAGTTTAGCAGAACCTGACTTG 413

Qy          30 ---ThrTyrValAspLeuThr------AsnGluGluThrThrAspSerThrThrSerLys 46
                  ::::::|||              |||:::      |||:::   |||   |||
Db         414 ATCAACTTTATGGACTTCCCAAAACATAACCAGATCATAACTGAAGAAACAGGCTCTGCA 473

Qy          47 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
               :::   ||||||::::::  ::::::   :::|||   :::   ::::::
Db         474 GTTGAACCAAGTGATGAAATAAAGAGAGCCAGTGGAGATGTCCAAACTATGAAAATTTCA 533

Qy          67 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
               ::::::                |||:::||||||:::|||
Db         534 TCTGTGCCT------------AATAGTTTATCAAAGCGA 560

RESULT 74
US-10-917-503B-10820
; Sequence 10820, Application US/10917503B
; GENERAL INFORMATION:
;  APPLICANT: OTA, TOSHIO
;  APPLICANT:  ISOGAI, TAKAO
;  APPLICANT:  NISHIKAWA, TETSUO
;  APPLICANT:  HAYASHI, KOJI
;  APPLICANT:  SAITO, KAORU
;  APPLICANT:  YAMAMOTO, JUNICHI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  OTSUKI, TETSUJI
;  TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
;  FILE REFERENCE: 084335/0123
;  CURRENT APPLICATION NUMBER: US/10/917,503B
;  CURRENT FILING DATE:  2004-08-13
;  PRIOR APPLICATION NUMBER: US/09/629,469
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: JP 1999-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 1999-300253
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  NUMBER OF SEQ ID NOS: 19025
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10820
;   LENGTH: 2082
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1394)
US-10-917-503B-10820

Alignment Scores:
Pred. No.:                185             Length:          2082
Score:                    65.50           Matches:         18
Percent Similarity:       56.2%           Conservative:    23
Best Local Similarity:    24.7%           Mismatches:      23
Query Match:              14.6%           Indels:          9
DB:                       6               Gaps:            4

US-10-757-745-2_COPY_54_140 (1-87) x US-10-917-503B-10820 (1-2082)

Qy          12 ProValGluGluSerAlaLeuGluArgArgPro---GluThrIleSerGluProLys--- 29
               |||:::          :::|||||||   :::|||       :::::::::||||||
Db         354 CCATTGGCAAATGAATCCTTGGAGCACAAACCTGTATCCAGTTTAGCAGAACCTGACTTG 413

Qy          30 ---ThrTyrValAspLeuThr------AsnGluGluThrThrAspSerThrThrSerLys 46
                  ::::::|||              |||:::      |||:::   |||   |||
Db         414 ATCAACTTTATGGACTTCCCAAAACATAACCAGATCATAACTGAAGAAACAGGCTCTGCA 473

Qy          47 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
               :::   ||||||::::::  ::::::   :::|||   :::   ::::::
Db         474 GTTGAACCAAGTGATGAAATAAAGAGAGCCAGTGGAGATGTCCAAACTATGAAAATTTCA 533

Qy          67 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
               ::::::                |||:::||||||:::|||
Db         534 TCTGTGCCT------------AATAGTTTATCAAAGCGA 560

RESULT 75
US-10-558-119-19
; Sequence 19, Application US/10558119
; GENERAL INFORMATION:
;  APPLICANT: MEINKE, ANDREAS
;  APPLICANT:  NAGY, ESZTER
;  APPLICANT:  HANNER, MARKUS
;  APPLICANT:  GELBMANN, DIETER
;  TITLE OF INVENTION: ENTEROCOCCUS ANTIGENS
;  FILE REFERENCE: SONN:083US
;  CURRENT APPLICATION NUMBER: US/10/558,119
;  CURRENT FILING DATE:  2005-11-23
;  PRIOR APPLICATION NUMBER: PCT/EP2004/005664
;  PRIOR FILING DATE: 2004-05-26
;  PRIOR APPLICATION NUMBER: EP03450137.9
```

```
;  PRIOR FILING DATE: 2003-05-30
;  NUMBER OF SEQ ID NOS: 478
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
;   LENGTH: 2211
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
US-10-558-119-19

Alignment Scores:
Pred. No.:              202          Length:          2211
Score:                  65.50        Matches:         22
Percent Similarity:     47.2%        Conservative:    12
Best Local Similarity:  30.6%        Mismatches:      26
Query Match:            14.6%        Indels:          12
DB:                     6            Gaps:            2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-558-119-19 (1-2211)

Qy         10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29
              :::   |||      |||:::   |||          |||   |||   :::   |||
Db        193 CAAGCACCTACTACGGAAACTGGCTTAGTAGAGACACCAACAACAGAAACTACGCCAGGA 252

Qy         30 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 49
              |||             |||   :::   ||||||||||||:::|||:::      |||
Db        253 ACT-----------ACGGAACAACCGACAACGGATTCGTCAACAACGACTGAATCGACA 300

Qy         50 SerGluAspThrGlnGln--------------------Glu-AsnGlySerMetPheSe 62
              :::|||   :::::::::                    ||| |||      :::
Db        301 ACTGAATCATCAAAAGAAACACCAACAACACCAAGTACCGAGCAACCAACAGCTGATTCA 360

Qy         62 rLeuIleThrTrpAsnIleAspGlyLeuAspLeu 73
               |||      ||||||   :::   |||   |||
Db        361 ACTACACCTGTGGAATCAGGAACGACTGATTCTT 394


Search completed: March 11, 2006, 00:07:35
Job time : 369.552 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:38:30 ; Search time 878.391 Seconds
(without alignments)
1722.336 Million cell updates/sec

Title: US-10-757-745-2_COPY_121_347
Perfect score: 1192
Sequence: 1 IDGLDLNNLSERARGVCSYL.....HIIPRSLDLLGLEKLDCGRF 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170325_3993/app_query.fasta_1
-DB=N_Geneseq -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1589_@runat_09032006_170325_3993 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1192	100.0	1898	4	AAH15146	Aah15146 Human cDN
2	1192	100.0	1920	3	AAZ47118	Aaz47118 Human CD4
3	1192	100.0	1936	10	ADD19013	Add19013 Human dis
4	1192	100.0	1936	13	ADP25361	Adp25361 PRO polyp
5	1192	100.0	1940	14	ADX06356	Adx06356 Cyclin-de
6	1192	100.0	1948	4	AAI58997	Aai58997 Human pol
7	1192	100.0	1948	5	ADQ99219	Adq99219 DNA encod
8	1192	100.0	1948	9	ADB48979	Adb48979 Novel hum
9	1192	100.0	1958	4	AAI60783	Aai60783 Human pol
10	1192	100.0	2499	2	AAX28153	Aax28153 Topoisome
11	1192	100.0	3152	9	ADA10970	Ada10970 Human cDN
12	1184	99.3	1296	3	AAC98160	Aac98160 Human col
13	1116	93.6	1088	5	AAS86254	Aas86254 DNA encod
14	1102	92.4	674	11	ADT95548	Adt95548 Colon can
15	1102	92.4	674	11	ADX42030	Adx42030 Human cDN
c 16	1090	91.4	1079	2	AAX84209	Aax84209 DNA encod
c 17	1090	91.4	1079	3	AAC79438	Aac79438 cDNA sequ
c 18	1090	91.4	1079	6	ABK28982	Abk28982 Human bre
19	1069	89.7	644	11	ADT95551	Adt95551 Colon can
20	1069	89.7	644	11	ADX42033	Adx42033 Human cDN
21	1054	88.4	633	11	ADT95565	Adt95565 Colon can
22	1054	88.4	633	11	ADX42047	Adx42047 Human cDN
23	1031	86.5	625	11	ADT95029	Adt95029 Colon can
24	1031	86.5	625	11	ADX41511	Adx41511 Human cDN
25	1020	85.6	625	11	ADT95512	Adt95512 Colon can
26	1020	85.6	625	11	ADX41994	Adx41994 Human cDN
27	945	79.3	1312	3	AAZ47119	Aaz47119 Mouse CD4
28	902	75.7	553	11	ADT95464	Adt95464 Colon can
29	902	75.7	553	11	ADX41946	Adx41946 Human cDN
c 30	783	65.7	774	10	ADD34375	Add34375 Mouse mit
31	651	54.6	401	11	ADT95844	Adt95844 Colon can
32	651	54.6	401	11	ADX42326	Adx42326 Human cDN
33	608	51.0	602	10	ADD34376	Add34376 Mouse mit
34	568	47.7	752	4	AAH08073	Aah08073 Human cDN
35	465	39.0	1227	5	AAS86255	Aas86255 DNA encod
36	298	25.0	176	6	ABK27684	Abk27684 Human col
c 37	298	25.0	176	6	ABV95959	Abv95959 Human pan
c 38	298	25.0	179	6	ABL36461	Abl36461 Human col
39	214	18.0	483	2	AAX40590	Aax40590 Human sec
40	201	16.9	1566	3	AAC47239	Aac47239 Arabidops
41	201	16.9	1606	3	AAC41089	Aac41089 Arabidops
42	182.5	15.3	444	4	AAI15311	Aai15311 Probe #52
43	182.5	15.3	691	4	AAI24484	Aai24484 Probe #14
44	158.5	13.3	525	13	ACN59762	Acn59762 Cotton gy
c 45	134	11.2	268	6	ABT07150	Abt07150 Human ova
c 46	134	11.2	268	8	ABX73028	Abx73028 Human ova
c 47	101.5	8.5	1877	6	ABL90668	Abl90668 Human pol
48	101.5	8.5	8446	6	ADG79370	Adg79370 Human sec
49	98.5	8.3	1644	13	ADS45761	Ads45761 Bacterial
50	97.5	8.2	2928	12	ADQ97808	Adq97808 Mouse can
51	97.5	8.2	9737	14	ADZ62001	Adz62001 Murine Cc
52	97.5	8.2	47040	12	ADQ97807	Adq97807 Mouse can
c 53	97.5	8.2	90043	11	ACN44608	Acn44608 Mouse gen
54	97	8.1	1692	8	ACA34492	Aca34492 Prokaryot
55	97	8.1	110000	2	AAT42063_16	Continuation (17 o
c 56	95	8.0	5386	13	ADT05432	Adt05432 Haemophil
c 57	95	8.0	349980	13	ADT05649	Adt05649 Haemophil
58	94.5	7.9	1650	8	ACA42939	Aca42939 Prokaryot
59	94.5	7.9	1767	12	ADE43923	Ade43923 Human DP1
60	94.5	7.9	1769	12	ADQ97811	Adq97811 Human can
61	94.5	7.9	1971	14	ADZ49510	Adz49510 Insulin s
62	94.5	7.9	49355	12	ADQ97810	Adq97810 Human can
63	94.5	7.9	110000	14	AEB39175_05	Continuation (6 of
64	94.5	7.9	110000	14	AEB42401_05	Continuation (6 of
65	94.5	7.9	243335	14	AEB42735	Aeb42735 L. pneumo
66	94.5	7.9	295644	14	AEB35721	Aeb35721 L. pneumo
67	94.5	7.9	298667	14	AEB39173	Aeb39173 L. pneumo
68	94	7.9	3285	6	ABL65231	Abl65231 Lung canc
69	94	7.9	3285	6	ABN97250	Abn97250 Gene #374
70	94	7.9	3285	10	ADE85050	Ade85050 Farnesyl
71	94	7.9	3397	6	ABL61829	Abl61829 Colon ade
72	94	7.9	3397	6	ABN95906	Abn95906 Gene #240
73	94	7.9	3397	10	ADE85000	Ade85000 Farnesyl
74	94	7.9	3423	13	ADR24971	Adr24971 Breast ca
75	94	7.9	3423	14	ADY19709	Ady19709 DNA encod

ALIGNMENTS

RESULT 1
AAH15146
ID AAH15146 standard; cDNA; 1898 BP.
XX
AC AAH15146;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:13209.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 13209; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 1898 BP; 590 A; 327 C; 426 G; 555 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.54e-135	Length:	1898
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0

DB: 4 Gaps: 0

US-10-757-745-2_COPY_121_347 (1-227) x AAH15146 (1-1898)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTT 1041
```

RESULT 2
AAZ47118
ID AAZ47118 standard; cDNA; 1920 BP.
XX
AC AAZ47118;
XX
DT 15-MAR-2000 (first entry)
XX
DE Human CD40 receptor associated protein gene.
XX
KW Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW human; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW graft versus host disease; autoimmune disease; ds.
XX
OS Homo sapiens.
XX

PN WO9955859-A2.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1999; 99WO-EP003025.
XX
PR 29-APR-1998; 98EP-00201392.
XX
PA (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Pype SMC, Remacle JEFJG, Huylebroeck DFE;
XX
DR WPI; 2000-062029/05.
DR P-PSDB; AAY56019.
XX
PT Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT diseases and for improvement of anti-tumor treatments.
XX
PS Claim 9; Page 37-39; 48pp; English.
XX
CC This sequence represents the gene encoding human CD40 receptor associated
CC protein (CRAP). CRAP is a functional protein capable of interacting with
CC the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC I, where the protein has no homology to TNF receptor associated factor
CC (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC diseases, and for the improvement of anti-tumour diseases. Diseases which
CC may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC systemic lupus erythematosus, graft rejection, graft versus host disease,
CC allergy, and autoimmune disease. The proteins can be used to sensitize
CC tumour cells to anti-tumour treatments and to screen for compounds which
CC interfere with the interaction of the proteins with other protein
CC components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ Sequence 1920 BP; 599 A; 327 C; 435 G; 557 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	1.56e-135	Length:	1920
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x AAZ47118 (1-1920)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuGluLeuGluGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTT 1060
```

RESULT 3
ADD19013
ID ADD19013 standard; DNA; 1936 BP.
XX
AC ADD19013;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human disease related protein DNA sequence SeqID502.
XX
KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW antiarteriosclerotic; vulnerary; gene therapy;
KW hypoxia-regulated condition; tumourigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003018621-A2.
XX
PD 06-MAR-2003.
XX
PF 23-AUG-2002; 2002WO-GB003892.
XX
PR 23-AUG-2001; 2001GB-00020558.
PR 05-OCT-2001; 2001GB-00024037.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
DR WPI; 2003-290046/28.
DR P-PSDB; ADD19012.
XX
PT New substantially purified polypeptide, useful for diagnosing or treating
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT wound healing.
XX
PS Claim 27; SEQ ID NO 502; 424pp; English.
XX
CC This invention relates to novel human genes and gene product which are
CC implicated in certain disease states. Compounds which modulate the
CC proteins of the invention may have cytostatic, antiinflammatory,
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The
CC sequences of the invention may be useful for gene therapy. The invention
CC may be useful for diagnosing or treating a hypoxia-regulated condition,

CC such as tumourigenesis, angiogenesis, apoptosis, inflammation,
CC erythropoiesis, or the biological response to hypoxia conditions
CC including processes such as glycolysis, gluconeogenesis, glucose
CC transportation, catecholamine synthesis, iron transport or nitric oxide
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC inflammatory conditions or wound healing. The present sequence is that of
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.58e-135	Length:	1936
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x ADD19013 (1-1936)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 436

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 496

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 556

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 616

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 676

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 736

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 796

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 856

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 916

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 976

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1036

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1037 AAACTGGACTGTGGTAGATTT 1057
```

RESULT 4
ADP25361
ID ADP25361 standard; cDNA; 1936 BP.
XX
AC ADP25361;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide encoding cDNA SEQ ID NO:475.
XX
KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-419628/39.
DR P-PSDB; ADP25362.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 1; SEQ ID NO 475; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence encodes a PRO protein of
CC the invention.
XX
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.58e-135	Length:	1936
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	13	Gaps:	0

```
US-10-757-745-2_COPY_121_347 (1-227) x ADP25361 (1-1936)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 436

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 496

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 556

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 616

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 676

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 736

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 796

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 856

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 916

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 976

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1036

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1037 AAACTGGACTGTGGTAGATTT 1057

RESULT 5
ADX06356
ID   ADX06356 standard; DNA; 1940 BP.
XX
AC   ADX06356;
XX
DT   21-APR-2005  (first entry)
XX
DE   Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 921.
XX
KW   cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2005012875-A2.
XX
PD   10-FEB-2005.
XX
PF   29-JUL-2004; 2004WO-US024424.
XX
PR   29-JUL-2003; 2003US-0490890P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Li M,  Rupnow BA,  Webster KR,  Jackson DG,  Wong TW;
XX
DR   WPI; 2005-163068/17.
DR   P-PSDB; ADX06357.
XX
PT   Biomarkers useful for predicting or determining the response of a mammal
PT   to a cancer treatment comprising administration of a modulator of cyclin-
PT   dependent kinase activity.
XX
PS   Claim 5; SEQ ID NO 921; 141pp; English.
XX
CC   This invention describes a novel method of predicting or determining
CC   whether a mammal will respond or is responding to an anti-cancer agent
CC   that modulates cyclin-dependent kinase (cdk) activity. The method
CC   comprises measuring the level of one or more biomarkers selected from
CC   2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC   NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC   invention is utilized in a kit for determining or predicting whether
CC   patient would be susceptible or resistant to treatment by an agent
CC   modulating cdk activity. The invention also describes a method for
CC   utilizing individualized genetic profiles for treating diseases and
CC   disorders based on patient's response and molecular level, specialized
CC   microarrays comprising the biomarkers described, antibodies directed
CC   against the biomarkers and a cell culture model to identify biomarkers.
CC   The cdk modulator is preferably N-5-[[5-(1,1-Dimethylethyl)-2-
CC   oxazolyl]methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC   tartaric acid salt. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC   sequence encodes a biomarker used in the method of the invention.
XX
SQ   Sequence 1940 BP; 613 A; 330 C; 439 G; 558 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.59e-135       Length:          1940
Score:                    1192.00         Matches:         227
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       14              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x ADX06356 (1-1940)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805
```

```
Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066
```

RESULT 6
AAI58997
ID AAI58997 standard; cDNA; 1948 BP.
XX
AC AAI58997;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1200.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM39841.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 1200; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.6e-135	Length:	1948
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x AAI58997 (1-1948)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079
```

RESULT 7

```
ADQ99219
ID   ADQ99219 standard; cDNA; 1948 BP.
XX
AC   ADQ99219;
XX
DT   23-SEP-2004  (first entry)
XX
DE   DNA encoding human GPCR-like protein seqid 889.
XX
KW   ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
KW   antidiabetic; GPCR-like protein; ophthalmic disorder;
KW   neurological disorder; immunological disorder; nephritic disorder;
KW   hormonal dysfunction; cancer; atherosclerosis; diabetes;
KW   molecular weight marker; food supplement; human; ss.
XX
OS   Homo sapiens.
XX
PN   US6569662-B1.
XX
PD   27-MAY-2003.
XX
PF   19-JUL-2000; 2000US-00620312.
XX
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Zhou P,  Drmanac RT;
XX
DR   WPI; 2001-442255/47.
XX
PT   New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT   useful for treating diseases of ophthalmic, neurological, immunological
PT   and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT   and diabetes.
XX
PS   Example 2; SEQ ID NO 889; 92pp; English.
XX
CC   The invention describes an isolated polynucleotide (I) comprising a fully
CC   defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC   1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC   given in the specification, its translated or protein coding portion, its
CC   extracellular portion or its active domain. The GPCR-like polypeptides
CC   and polynucleotides are useful for the treatment of diseases of
CC   ophthalmic, neurological, immunological and nephritic systems. They may
CC   also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC   diabetes. The antibodies are useful for detecting or quantitating the
CC   polypeptide in tissue. The polypeptides can also be used as molecular
CC   weight markers and as a food supplement. This sequence represents a human
CC   polynucleotide of the invention.
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.6e-135      Length:          1948
Score:                  1192.00       Matches:         227
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     5             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x ADQ99219 (1-1948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079

RESULT 8
ADB48979
ID   ADB48979 standard; cDNA; 1948 BP.
XX
AC   ADB48979;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Novel human cDNA SEQ ID NO 889.
XX
KW   ss; cancer; neurodegenerative disease; human.
XX
OS   Homo sapiens.
XX
PN   US2003104529-A1.
XX
PD   05-JUN-2003.
XX
PF   04-JAN-2002; 2002US-00037270.
XX
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   19-JUL-2000; 2000US-00620312.
XX
PA   (ZHOU/) ZHOU P.
PA   (TANG/) TANG Y T.
PA   (LIUC/) LIU C.
PA   (ASUN/) ASUNDI V.
PA   (DRMA/) DRMANAC R T.
XX
PI   Zhou P,  Tang YT,  Liu C,  Asundi V,  Drmanac RT;
XX
DR   WPI; 2003-678194/64.
XX
```

```
PT   New polynucleotide, useful for treating diseases e.g., cancer or
PT   neurodegenerative diseases.
XX
PS   Claim 1; SEQ ID NO 889; 99pp; English.
XX
CC   The invention relates to a polynucleotide comprising a sequence given in
CC   the specification, or its mature protein-coding portion, or its
CC   complement. The polynucleotide is useful for treating diseases e.g.,
CC   cancer or neurodegenerative diseases and many others listed in the
CC   specification. The present sequence represents a novel human cDNA. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=20030104529.
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.6e-135        Length:         1948
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADB48979 (1-1948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079
```

```
RESULT 9
AAI60783
ID   AAI60783 standard; cDNA; 1958 BP.
XX
AC   AAI60783;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4772.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J,  Zhao QA;
PI   Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41627.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders such
PT   as central nervous system injuries.
XX
PS   Claim 1; SEQ ID NO 4772; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC   encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders. Note: The sequence data for this patent did not form
CC   part of the printed specification
XX
SQ   Sequence 1958 BP; 613 A; 339 C; 442 G; 564 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.61e-135       Length:         1958
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
```

```
Query Match:          100.0%      Indels:        0
DB:                   4           Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x AAI60783 (1-1958)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 470

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 530

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 590

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 650

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 710

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 770

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 830

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 890

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 950

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1010

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1070

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||
Db       1071 AAACTGGACTGTGGTAGATTT 1091

RESULT 10
AAX28153
ID   AAX28153 standard; DNA; 2499 BP.
XX
AC   AAX28153;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Topoisomerase II binding protein 1 coding sequence.
XX
KW   Topoisomerase II binding protein; TopBP; anticancer agent; ds.
XX
OS   Homo sapiens.
XX
PN   JP11075856-A.
XX
PD   23-MAR-1999.
XX
PF   17-SEP-1997;   97JP-00251544.
XX
PR   17-SEP-1997;   97JP-00251544.
XX
PA   (TSUR/) TSURUO T.
PA   (CHUS ) CHUGAI PHARM CO LTD.
XX
DR   WPI; 1999-257704/22.
DR   P-PSDB; AAY03182.
XX
PT   New Topoisomerase II- binding protein - useful as an anticancer agent.
XX
PS   Disclosure; Page 18-19; 28pp; Japanese.
XX
CC   This sequence encodes the topoisomerase II binding protein (TopBP) of the
CC   invention. The TopBP protein is useful as an anticancer agent. TopBP can
CC   be used as the target molecule for anticancer agent
XX
SQ   Sequence 2499 BP; 720 A; 505 C; 587 G; 687 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.29e-135      Length:        2499
Score:                    1192.00        Matches:       227
Percent Similarity:       100.0%         Conservative:  0
Best Local Similarity:    100.0%         Mismatches:    0
Query Match:              100.0%         Indels:        0
DB:                       2              Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x AAX28153 (1-2499)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 901

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 961

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1021

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1081

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1141

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1201

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1261

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1321

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1381

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1441

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1501

Qy        221 LysLeuAspCysGlyArgPhe 227
```

```
                    ||||||||||||||||||||
Db       1502 AAACTGGACTGTGGTAGATTT 1522

RESULT 11
ADA10970
ID   ADA10970 standard; cDNA; 3152 BP.
XX
AC   ADA10970;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human cDNA differentially expressed in colon cancer #60.
XX
KW   ss; differential expression; colon cancer; cancer; human.
XX
OS   Homo sapiens.
XX
PN   US2002160382-A1.
XX
PD   31-OCT-2002.
XX
PF   11-OCT-2001; 2001US-00981353.
XX
PR   11-OCT-2000; 2000US-0239841P.
XX
PA   (LASE/) LASEK A W.
PA   (JONE/) JONES D A.
XX
PI   Lasek AW,  Jones DA;
XX
DR   WPI; 2003-265756/26.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   colon disorder, useful for diagnosing, treating, staging or monitoring
PT   treatment for colon cancers.
XX
PS   Claim 1; SEQ ID NO 88; 231pp; English.
XX
CC   The invention relates to a combination comprising cDNAs that are
CC   differentially expressed in colon disorder. The methods and compositions
CC   of the present invention are useful for diagnosing, treating, staging or
CC   monitoring treatment for colon cancer. They are also useful in high
CC   throughput methods for using cDNAs to detect differential expression of
CC   nucleic acids in a sample, screening molecules or compounds to identify a
CC   ligand which specifically binds a cDNA and using a protein to screen
CC   molecules or compounds to identify at least one ligand which specifically
CC   binds the protein. The present sequence represents a human cDNA
CC   differentially expressed in colon cancer.
XX
SQ   Sequence 3152 BP; 875 A; 643 C; 707 G; 926 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              3.2e-135        Length:         3152
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADA10970 (1-3152)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 1671

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1672 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 1731

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1732 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1791

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1792 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1851

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1852 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1911

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1912 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1971

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1972 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 2031

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2032 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 2091

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2092 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 2151

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2152 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 2211

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2212 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 2271

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       2272 AAACTGGACTGTGGTAGATTT 2292

RESULT 12
AAC98160
ID   AAC98160 standard; cDNA; 1296 BP.
XX
AC   AAC98160;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human colon cancer antigen nucleotide sequence SEQ ID NO:170.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW   immunomodulatory; muscular; gynaecological; gastrointestinal;
KW   nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW   neural disorder; immune system disorder; muscular disorder;
KW   reproductive disorder; gastrointestinal disorder; renal disorder;
KW   infectious disease; cardiovascular disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055351-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005883.
XX
PR   12-MAR-1999;   99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587534/55.
DR   P-PSDB; AAB53403.
XX
PT   Colon cancer associated gene sequences, referred to as colon cancer
PT   antigens, useful for the treatment, prevention, and diagnosis of colon
```

```
PT   disorders such as colon cancer.
XX
PS   Claim 1; Page 597; 2104pp; English.
XX
CC   AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC   called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC   human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC   neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC   vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC   can be used in gene therapy. The colon cancer antigen polynucleotides,
CC   proteins and antibodies to the proteins are useful for the prevention,
CC   treatment and diagnosis of colon disorders, such as colon cancer. The
CC   polynucleotides may be used in diagnostics and research, such as for
CC   chromosome identification, and as hybridisation probes. The proteins may
CC   also be used to prevent diseases such as neural disorders, immune system
CC   disorders, muscular disorders, reproductive disorders, gastrointestinal
CC   disorders, wounds, renal disorders, infectious diseases, and
CC   cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC   sequences used in the exemplification of the present invention
XX
SQ   Sequence 1296 BP; 376 A; 250 C; 326 G; 333 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.:              8.49e-135      Length:          1296
Score:                  1184.00        Matches:         226
Percent Similarity:     99.6%          Conservative:    0
Best Local Similarity:  99.6%          Mismatches:      1
Query Match:            99.3%          Indels:          0
DB:                     3              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x AAC98160 (1-1296)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 443

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 503

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACA 563

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 623

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTT 683

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 743

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 803

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 863

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 GTGGATGTCTGGGAGTTTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAA 923

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        924 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 983

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        984 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1043

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1044 AAACTGGACTGTGGTAGATTT 1064

RESULT 13
AAS86254
ID   AAS86254 standard; cDNA; 1088 BP.
XX
AC   AAS86254;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #22058.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG22067.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 22058; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1088 BP; 345 A; 182 C; 236 G; 325 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.46e-126      Length:          1088
Score:                  1116.00        Matches:         212
Percent Similarity:     99.1%          Conservative:    1
```

```
Best Local Similarity:  98.6%          Mismatches:    2
Query Match:            93.6%          Indels:        0
DB:                     5              Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x AAS86254 (1-1088)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy        213 SerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 645

RESULT 14
ADT95548
ID   ADT95548 standard; cDNA; 674 BP.
XX
AC   ADT95548;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1067.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1067; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, which both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.:              3.82e-125      Length:        674
Score:                  1102.00        Matches:       213
Percent Similarity:     98.6%          Conservative:  0
Best Local Similarity:  98.6%          Mismatches:    2
Query Match:            92.4%          Indels:        1
DB:                     11             Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x ADT95548 (1-674)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 77

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        258 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        133 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProAr 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy        212 gSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||    ||||||||   |||||||||
Db        618 AAGTTTGGACCTTCTTGGATTANAAAAACTGGACTGNGGTAGATTT 663

RESULT 15
ADX42030
ID   ADX42030 standard; cDNA; 674 BP.
XX
AC   ADX42030;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1067.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1067; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.:                3.82e-125      Length:          674
Score:                    1102.00        Matches:         213
Percent Similarity:       98.6%          Conservative:    0
Best Local Similarity:    98.6%          Mismatches:      2
Query Match:              92.4%          Indels:          1
DB:                       11             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x ADX42030 (1-674)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 77

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        133 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProAr 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy        212 gSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||    ||||||||   |||||||||
Db        618 AAGTTTGGACCTTCTTGGATTANAAAAACTGGACTGNGGTAGATTT 663

RESULT 16
AAX84209/c
ID   AAX84209 standard; cDNA; 1079 BP.
XX
AC   AAX84209;
XX
DT   08-SEP-1999  (first entry)
XX
DE   DNA encoding human breast tumour protein immunogenic fragment.
XX
KW   Breast tumour protein; immunogenic fragment; vaccine; detection;
KW   breast cancer development; therapy; ss.
XX
```

OS Homo sapiens.
XX
PN WO9933869-A2.
XX
PD 08-JUL-1999.
XX
PF 22-DEC-1998; 98WO-US027416.
XX
PR 24-DEC-1997; 97US-00998253.
PR 24-DEC-1997; 97US-00998255.
PR 17-JUL-1998; 98US-00118554.
PR 17-JUL-1998; 98US-00118627.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J;
XX
DR WPI; 1999-405486/34.
XX
PT New breast tumor protein genes used, in vaccines for immunotherapy, or
PT for diagnosis of breast cancer.
XX
PS Claim 3; Page 60; 70pp; English.
XX
CC This sequence encodes a human breast tumour protein immunogenic fragment
CC of the invention. The polypeptides or nucleic acids encoding them are
CC useful in vaccines and pharmaceutical compositions for manufacture of
CC medicaments for inhibiting the development of breast cancer in a patient.
CC They can also be used to treat breast cancer. Antibodies against these
CC polypeptides can be used to detect and monitor progression of breast
CC cancer in patients. Primers and probes derived from the polynucleotides
CC encoding the breast proteins are useful for detection of breast cancer.
CC Peripheral blood cells from a patient incubated in the presence of at
CC least one polypeptide, such that T cells proliferate, are useful in
CC manufacture of a medicament for treating breast cancer in a patient.
CC Antigen presenting cells incubated in the presence of at least one
CC polypeptide are also useful for treating breast cancer
XX
SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.24e-123	Length:	1079
Score:	1090.00	Matches:	211
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	91.4%	Indels:	1
DB:	2	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x AAX84209 (1-1079)

```
Qy        16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy        36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy        56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy        76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy        96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy       116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy       136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy       156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy       176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy       196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
             |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db       537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy       216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
             ||||||||||||||||||||||||||||||||||||
Db       478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443
```

RESULT 17
AAC79438/c
ID AAC79438 standard; cDNA; 1079 BP.
XX
AC AAC79438;
XX
DT 07-FEB-2001 (first entry)
XX
DE cDNA sequence of human breast tumour clone 1015D11.
XX
KW Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;
KW vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200061756-A2.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US009688.
XX
PR 09-APR-1999; 99US-00288950.
PR 02-JUL-1999; 99US-00346327.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J, Dillon DC;
XX
DR WPI; 2000-638568/61.
XX
PT A novel isolated polypeptide comprising an immunogenic portion of a
PT breast cancer protein useful in the detection and treatment of breast
PT cancer.
XX
PS Claim 4; Page 77; 95pp; English.
XX
CC The present sequence was isolated from a breast tumour cDNA library. It
CC is provided in a specification relating to compounds for immunotherapy
CC and diagnosis of breast cancer. Breast tumour antigens and the
CC polynucleotides that encode them may be used in the production of a
CC pharmaceutical composition to be used in the treatment of breast cancer.
CC Proliferated T cells and incubated antigen presenting cells are also
CC required. The polypeptides and polynucleotides may also be used to
CC produce a vaccine
XX
SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.24e-123	Length:	1079
Score:	1090.00	Matches:	211
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1

```
Query Match:            91.4%       Indels:        1
DB:                     3           Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x AAC79438 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 18
ABK28982/c
ID   ABK28982 standard; cDNA; 1079 BP.
XX
AC   ABK28982;
XX
DT   23-APR-2002  (first entry)
XX
DE   Human breast tumour polypeptide cDNA clone #11.
XX
KW   Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
KW   immunostimulant.
XX
OS   Homo sapiens.
XX
PN   WO200198339-A2.
XX
PD   27-DEC-2001.
XX
PF   12-JUN-2001; 2001WO-US019032.
XX
PR   22-JUN-2000; 2000US-00602877.
PR   12-OCT-2000; 2000US-00687507.
PR   06-FEB-2001; 2001US-00778381.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J,  Dillon DC,  Retter MW,  Harlocker SL;
XX
DR   WPI; 2002-147792/19.
XX
PT   Polynucleotides encoding breast tumor polypeptides, useful for treating
PT   breast cancer or stimulating an immune response.
XX
PS   Claim 1; Page 131; 150pp; English.
XX
CC   The invention relates to polynucleotides encoding breast tumour
CC   polypeptides. The sequences are useful for treating cancer, preferably
CC   breast cancer, in a patient or for stimulating an immune response. The
CC   polynucleotides and polypeptides are also useful in the diagnosis and
CC   monitoring of breast cancer. A method for detecting the presence of a
CC   cancer in a patient, comprises obtaining a biological sample from the
CC   patient, contacting the biological sample with a binding agent that binds
CC   to a breast tumour polypeptide, detecting in the sample an amount of
CC   polypeptide that binds to the binding agent, and comparing the amount of
CC   polypeptide to a predetermined cut-off value, therefore determining the
CC   presence of a cancer in the patient. Sequences ABK28920-ABK29025
CC   represent cDNA clones encoding human breast tumour polypeptides of the
CC   invention
XX
SQ   Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.24e-123       Length:        1079
Score:                  1090.00         Matches:       211
Percent Similarity:     99.5%           Conservative:  0
Best Local Similarity:  99.5%           Mismatches:    1
Query Match:            91.4%           Indels:        1
DB:                     6               Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x ABK28982 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538
```

```
Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 19
ADT95551
ID   ADT95551 standard; cDNA; 644 BP.
XX
AC   ADT95551;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1070.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1070; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               4.01e-121      Length:          644
Score:                   1069.00        Matches:         203
Percent Similarity:      100.0%         Conservative:    1
Best Local Similarity:   99.5%          Mismatches:      0
Query Match:             89.7%          Indels:          0
DB:                      11             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x ADT95551 (1-644)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy        213 SerLeuAspLeu 216
              ||||||||||||
Db        632 AGTTTGGACCTT 643

RESULT 20
ADX42033
ID   ADX42033 standard; cDNA; 644 BP.
XX
AC   ADX42033;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1070.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
```

PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1070; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.01e-121	Length:	644
Score:	1069.00	Matches:	203
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	99.5%	Mismatches:	0
Query Match:	89.7%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x ADX42033 (1-644)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db        332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy        213 SerLeuAspLeu 216
              ||||||||||||
Db        632 AGTTTGGACCTT 643
```

RESULT 21
ADT95565
ID ADT95565 standard; cDNA; 633 BP.
XX
AC ADT95565;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #1084.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1084; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the

```
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              2.71e-119       Length:         633
Score:                  1054.00         Matches:        200
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            88.4%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADT95565 (1-633)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 633

RESULT 22
ADX42047
ID   ADX42047 standard; cDNA; 633 BP.
XX
AC   ADX42047;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1084.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1084; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              2.71e-119       Length:         633
Score:                  1054.00         Matches:        200
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            88.4%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADX42047 (1-633)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333
```

```
Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 633
```

RESULT 23
ADT95029
ID ADT95029 standard; cDNA; 625 BP.
XX
AC ADT95029;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #548.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 548; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	1.77e-116	Length:	625
Score:	1031.00	Matches:	198
Percent Similarity:	99.0%	Conservative:	0
Best Local Similarity:	99.0%	Mismatches:	1
Query Match:	86.5%	Indels:	1
DB:	11	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x ADT95029 (1-625)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy         53 -GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 72
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy         72 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy         92 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        112 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy        132 oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 623
```

RESULT 24
ADX41511
ID ADX41511 standard; cDNA; 625 BP.
XX
AC ADX41511;
XX
DT 21-APR-2005 (first entry)

```
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 548.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 548; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              1.77e-116       Length:         625
Score:                  1031.00         Matches:        198
Percent Similarity:     99.0%           Conservative:   0
Best Local Similarity:  99.0%           Mismatches:     1
Query Match:            86.5%           Indels:         1
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADX41511 (1-625)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy         53 -GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 72
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy         72 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy         92 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        112 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy        132 oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 623

RESULT 25
ADT95512
ID   ADT95512 standard; cDNA; 625 BP.
XX
AC   ADT95512;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1031.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1031; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
```

```
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              3.96e-115      Length:         625
Score:                  1020.00        Matches:        194
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            85.6%          Indels:         0
DB:                     11             Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADT95512 (1-625)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 581

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 206
              ||||||||||||||||||||||||||||||||||||||||||
Db        582 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAG 623

RESULT 26
ADX41994
ID   ADX41994 standard; cDNA; 625 BP.
XX
AC   ADX41994;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1031.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI   Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1031; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              3.96e-115      Length:         625
Score:                  1020.00        Matches:        194
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            85.6%          Indels:         0
DB:                     11             Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADX41994 (1-625)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281
```

```
Qy          93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy         113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy         133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy         153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy         173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         522 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 581

Qy         193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 206
               |||||||||||||||||||||||||||||||||||||||||
Db         582 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAG 623
```

RESULT 27
AAZ47119
ID AAZ47119 standard; cDNA; 1312 BP.
XX
AC AAZ47119;
XX
DT 15-MAR-2000 (first entry)
XX
DE Mouse CD40 receptor associated protein gene.
XX
KW Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW mouse; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW graft versus host disease; autoimmune disease; ds.
XX
OS Mus musculus.
XX
PN WO9955859-A2.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1999; 99WO-EP003025.
XX
PR 29-APR-1998; 98EP-00201392.
XX
PA (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Pype SMC, Remacle JEFJG, Huylebroeck DFE;
XX
DR WPI; 2000-062029/05.
DR P-PSDB; AAY56020.
XX
PT Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT diseases and for improvement of anti-tumor treatments.
XX
PS Claim 10; Page 41-43; 48pp; English.
XX
CC This sequence represents the gene encoding mouse CD40 receptor associated
CC protein (CRAP). CRAP is a functional protein capable of interacting with
CC the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC I, where the protein has no homology to TNF receptor associated factor
CC (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC diseases, and for the improvement of anti-tumour diseases. Diseases which
CC may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC systemic lupus erythematosus, graft rejection, graft versus host disease,
CC allergy, and autoimmune disease. The proteins can be used to sensitize
CC tumour cells to anti-tumour treatments and to screen for compounds which
CC interfere with the interaction of the proteins with other protein
CC components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ Sequence 1312 BP; 359 A; 279 C; 352 G; 322 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.85e-105	Length:	1312
Score:	945.00	Matches:	179
Percent Similarity:	89.0%	Conservative:	23
Best Local Similarity:	78.9%	Mismatches:	23
Query Match:	79.3%	Indels:	2
DB:	3	Gaps:	1

US-10-757-745-2_COPY_121_347 (1-227) x AAZ47119 (1-1312)

```
Qy           1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
               |||||||||||||||      ||||||   ||||||||||||||||||||||||   |||
Db         512 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 571

Qy          21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
               ||||||||||||||||||||:::|||||||||||||||||||||||||||   :::|||
Db         572 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTAC 631

Qy          41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
               ||||||||||||:::::::::|||   ||||||||||||:::||||||||||||||||||
Db         632 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 691

Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               ||||||:::|||||||||   |||||||||   ||||||||||||||||||||||||||||
Db         692 GCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 751

Qy          81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
               :::|||||||||||||||||||||||||||||:::|||:::::   |||||||||
Db         752 AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 811

Qy         101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
               ||||||||||||||||||||||||||||||   |||:::|||||||:::   |||
Db         812 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 871

Qy         121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
               ||||||   ||||||   ||||||||||||||||||:::|||   ||||||||||||||||
Db         872 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 931

Qy         141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
               ||||||||||||||||||:::|||||||   :::|||||||||||||||:::|||:::
Db         932 GGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTT 991

Qy         161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
               |||   |||||||||||||||||||||||||||||||||||||||||||||||||:::
Db         992 TTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAA 1051

Qy         181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
                  |||:::||||||   |||   ||||||   |||   ||||||||||||||||||||
Db        1052 GCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTC 1111

Qy         201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
               |||      |||||||||||||||:::||||||:::||||||||||||:::||||||||||
Db        1112 AGA------GCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 1165

Qy         221 LysLeuAspCysGlyArgPhe 227
               |||||||||||||||||||||
Db        1166 AAACTGGACTGTGGTAGATTT 1186
```

RESULT 28
ADT95464
ID ADT95464 standard; cDNA; 553 BP.
XX

```
AC   ADT95464;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #983.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 983; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              1.01e-100      Length:         553
Score:                  902.00         Matches:        171
Percent Similarity:     98.3%          Conservative:   0
Best Local Similarity:  98.3%          Mismatches:     3
Query Match:            75.7%          Indels:         0
DB:                     11             Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADT95464 (1-553)

Qy          9 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 28
              |||           ||||||||||||||||||||||||||||||||||||||||||||||
Db         30 CTCTANATGCATGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATA 89

Qy         29 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 48
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 TTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTAT 149

Qy         49 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 GAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 209

Qy         69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 GTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTT 269

Qy         89 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 TTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAG 329

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 AGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATG 389

Qy        129 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 CAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGA 449

Qy        149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 GAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 509

Qy        169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 182
              |||||||||||||||||||||||||||||||||||||||||
Db        510 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 551

RESULT 29
ADX41946
ID   ADX41946 standard; cDNA; 553 BP.
XX
AC   ADX41946;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 983.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
```

PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 983; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	1.01e-100	Length:	553
Score:	902.00	Matches:	171
Percent Similarity:	98.3%	Conservative:	0
Best Local Similarity:	98.3%	Mismatches:	3
Query Match:	75.7%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x ADX41946 (1-553)

```
Qy          9 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 28
              |||          ||||||||||||||||||||||||||||||||||||||||||||||||
Db         30 CTCTANATGCATGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATA 89

Qy         29 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 48
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 TTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTAT 149

Qy         49 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 GAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 209

Qy         69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 GTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTT 269

Qy         89 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 TTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAG 329

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 AGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATG 389

Qy        129 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 CAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGA 449

Qy        149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 GAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 509

Qy        169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 182
              |||||||||||||||||||||||||||||||||||||||||
Db        510 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 551
```

RESULT 30
ADD34375/c
ID ADD34375 standard; DNA; 774 BP.
XX
AC ADD34375;
XX
DT 15-JAN-2004 (first entry)
XX
DE Mouse mitochondrial DNA sequence SEQ ID NO:2153.
XX
KW ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW mitochondrial disease; oxidative phosphorylation dysfunction;
KW oxidative stress; apoptosis; aging.
XX
OS Mus musculus.
XX
PN WO2003020220-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027886.
XX
PR 30-AUG-2001; 2001US-0316323P.
PR 31-AUG-2001; 2001CA-02356540.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Wallace DC, Levy S, Kerstann K, Procaccio V;
XX
DR WPI; 2003-300821/29.
XX
PT Array containing probes for genes involved in mitochondrial biology,
PT useful for determining mitochondrial biology gene expression profiles for
PT use in diagnosing pathologies and identifying biochemical pathways.
XX
PS Claim 2; SEQ ID NO 2153; 201pp; English.
XX
CC The invention relates to a novel array comprising at least two isolated
CC nucleotide molecules, each molecule having a sequence capable of uniquely
CC hybridising to a nucleic acid molecule which is an expression product of
CC a gene involved in mitochondrial biology. The array comprises two or more
CC isolated nucleic acid molecules or spots, each molecule having a sequence
CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC of the invention is useful for determining an expression profile of a
CC mouse or human sample containing nucleic acid, by contacting the array
CC with the sample under conditions allowing selective hybridisation, and
CC measuring hybridisation of nucleic acid in the sample to the array to
CC produce an expression profile. The array is also useful for determining
CC an expression profile of a first labelled sample containing nucleic acid
CC relative to a second, differently labelled sample containing nucleic
CC acid. The second sample is a reference or a standard. An array is useful
CC for determining an expression profile diagnostic of an energy-metabolism-
CC related physiological condition. An array of the invention is useful for
CC determining mitochondrial biology gene expression profiles of organisms,
CC such as human, mice and closely related species, tissue and organs of
CC such organisms, which are useful for determining expression profiles
CC diagnostic of energy metabolism-related physiological conditions,
CC diagnosing such physiological conditions, identifying biochemical
CC pathways, genes, and mutations involved in such physiological conditions,
CC identifying therapeutic agents useful for preventing and/or treating such
CC physiological conditions, evaluating and/or monitoring the efficacy of
CC such therapies, and creating and identifying animal models of human
CC energy metabolism-related physiological conditions. An array is also
CC useful for defining expression signatures or profiles for mitochondrial
CC diseases, as well as distinguishing clinical disorders that result from
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC apoptosis and aging. An array of the invention contains probes of genes
CC not previously recognised to participate in mitochondrial biology. The
CC sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC clones used to make the probes of the invention. Some sequences are not
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC 1906, 2408 and 2643.
XX
SQ Sequence 774 BP; 218 A; 169 C; 135 G; 252 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.57e-86	Length:	774
Score:	783.00	Matches:	155
Percent Similarity:	88.0%	Conservative:	21
Best Local Similarity:	77.5%	Mismatches:	22
Query Match:	65.7%	Indels:	4
DB:	10	Gaps:	1

```
US-10-757-745-2_COPY_121_347 (1-227) x ADD34375 (1-774)

Qy          28 IlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsn 47
               :::|||||||||||||   ||||||||||||   :::|||||||||||||||||   |||
Db         768 GTATTTCTACAGGAAGT-ATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGC-AGCCAG 711

Qy          48 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 67
               |||   ||||||||||||:::||||||||||||||||||||||||:::|||||||||
Db         710 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 651

Qy          68 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 87
               ||||||||   ||||||||||||||||||||||||||:::|||||||||||||||:::
Db         650 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAGC 591

Qy          88 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 107
               ||||||||||||:::|||::::::   |||||||||   ||||||||||||||||||||
Db         590 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 531

Qy         108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
               ||||||||||||   |||:::||||||||:::   ||||||||   ||||||   |||
Db         530 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 471

Qy         128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 147
               ||||||||||||||:::|||   ||||||||||||||||||||||||||||||||||||
Db         470 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 411

Qy         148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
               :::||||||   :::|||||||||||||||:::|||:::   |||   ||||||||||||
Db         410 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 351

Qy         168 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 187
               ||||||||||||||||||||||||||||||||||:::   |||:::||||||   |||
Db         350 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 291

Qy         188 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 207
               ||||||   |||   ||||||||||||||||||||||||      ||||||||||||
Db         290 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGG 237

Qy         208 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
               |||:::|||||||:::|||||||||||||:::||||||||||||||||||||||||||||
Db         236 CACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTT 177

RESULT 31
ADT95844
ID   ADT95844 standard; cDNA; 401 BP.
XX
AC   ADT95844;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1363.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1363; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 401 BP; 132 A; 67 C; 89 G; 113 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.02e-70        Length:         401
Score:                  651.00          Matches:        127
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            54.6%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADT95844 (1-401)

Qy          13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          19 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 78

Qy          33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          79 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 138

Qy          53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         139 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 198

Qy          73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         199 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258

Qy          93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         259 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 318

Qy         113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         319 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 378

Qy         133 GluSerAlaThrValIlePhe 139
               |||||||||||||||||||||
```

```
Db        379 GAGTCAGCTACAGTTATATTT 399

RESULT 32
ADX42326
ID   ADX42326 standard; cDNA; 401 BP.
XX
AC   ADX42326;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1363.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1363; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 401 BP; 132 A; 67 C; 89 G; 113 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.02e-70        Length:         401
Score:                  651.00          Matches:        127
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            54.6%           Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADX42326 (1-401)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 78

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 138

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 198

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 318

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 378

Qy        133 GluSerAlaThrValIlePhe 139
              |||||||||||||||||||||
Db        379 GAGTCAGCTACAGTTATATTT 399

RESULT 33
ADD34376
ID   ADD34376 standard; DNA; 602 BP.
XX
AC   ADD34376;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Mouse mitochondrial DNA sequence SEQ ID NO:2154.
XX
KW   ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW   mitochondrial disease; oxidative phosphorylation dysfunction;
KW   oxidative stress; apoptosis; aging.
XX
OS   Mus musculus.
XX
PN   WO2003020220-A2.
XX
PD   13-MAR-2003.
XX
PF   30-AUG-2002; 2002WO-US027886.
XX
PR   30-AUG-2001; 2001US-0316323P.
PR   31-AUG-2001; 2001CA-02356540.
XX
PA   (UYEM-) UNIV EMORY.
XX
PI   Wallace DC,  Levy S,  Kerstann K,  Procaccio V;
XX
DR   WPI; 2003-300821/29.
XX
PT   Array containing probes for genes involved in mitochondrial biology,
PT   useful for determining mitochondrial biology gene expression profiles for
PT   use in diagnosing pathologies and identifying biochemical pathways.
XX
PS   Claim 2; SEQ ID NO 2154; 201pp; English.
XX
CC   The invention relates to a novel array comprising at least two isolated
CC   nucleotide molecules, each molecule having a sequence capable of uniquely
CC   hybridising to a nucleic acid molecule which is an expression product of
CC   a gene involved in mitochondrial biology. The array comprises two or more
CC   isolated nucleic acid molecules or spots, each molecule having a sequence
CC   chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC   of the invention is useful for determining an expression profile of a
CC   mouse or human sample containing nucleic acid, by contacting the array
CC   with the sample under conditions allowing selective hybridisation, and
CC   measuring hybridisation of nucleic acid in the sample to the array to
CC   produce an expression profile. The array is also useful for determining
CC   an expression profile of a first labelled sample containing nucleic acid
CC   relative to a second, differently labelled sample containing nucleic
CC   acid. The second sample is a reference or a standard. An array is useful
CC   for determining an expression profile diagnostic of an energy-metabolism-
CC   related physiological condition. An array of the invention is useful for
```

```
CC   determining mitochondrial biology gene expression profiles of organisms,
CC   such as human, mice and closely related species, tissue and organs of
CC   such organisms, which are useful for determining expression profiles
CC   diagnostic of energy metabolism-related physiological conditions,
CC   diagnosing such physiological conditions, identifying biochemical
CC   pathways, genes, and mutations involved in such physiological conditions,
CC   identifying therapeutic agents useful for preventing and/or treating such
CC   physiological conditions, evaluating and/or monitoring the efficacy of
CC   such therapies, and creating and identifying animal models of human
CC   energy metabolism-related physiological conditions. An array is also
CC   useful for defining expression signatures or profiles for mitochondrial
CC   diseases, as well as distinguishing clinical disorders that result from
CC   oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC   apoptosis and aging. An array of the invention contains probes of genes
CC   not previously recognised to participate in mitochondrial biology. The
CC   sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC   clones used to make the probes of the invention. Some sequences are not
CC   present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC   1906, 2408 and 2643.
XX
SQ   Sequence 602 BP; 194 A; 119 C; 134 G; 155 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.37e-64      Length:         602
Score:                  608.00        Matches:        118
Percent Similarity:     89.9%         Conservative:   16
Best Local Similarity:  79.2%         Mismatches:     15
Query Match:            51.0%         Indels:         0
DB:                     10            Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ADD34376 (1-602)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||      ||||||  ||||||||||||||||||||||||    |||
Db        156 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 215

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||:::|||||||||||||||||||||||||||   :::|||
Db        216 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTAC 275

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::::::::|||   ||||||||||||:::|||||||||||||||||||
Db        276 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 335

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||||||:::|||||||||    |||||||||   ||||||||||||||||||||||||||
Db        336 GCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 395

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::|||||||||||||||:::||||||||||||:::|||::::::   |||||||||
Db        396 AATACCAAAATGATGAGAAGCCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 455

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||   |||:::|||||||||:::   |||
Db        456 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 515

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||  ||||||  ||||||||||||||||||:::|||   ||||||||||||||||
Db        516 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 575

Qy        141 GlyAspThrAsnLeuArgAspArgGlu 149
              |||||||||||||||||||:::|||
Db        576 GGAGATACAAATTTAAGAGATCAAGAA 602

RESULT 34
AAH08073
ID   AAH08073 standard; cDNA; 752 BP.
XX
AC   AAH08073;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:4908.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 1; SEQ ID NO 4908; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 752 BP; 225 A; 148 C; 200 G; 176 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:              1.53e-59      Length:         752
Score:                  568.00        Matches:        119
Percent Similarity:     90.9%         Conservative:   1
Best Local Similarity:  90.2%         Mismatches:     9
Query Match:            47.7%         Indels:         3
DB:                     4             Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AAH08073 (1-752)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480
```

```
Qy          41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy          81 SerThrLysMetMetArg-AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
               |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         601 AGTACCAAAATGATGAGAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 660

Qy         100 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 120
               ||||||| |||||| ||| :::||| |||||| ||||||||||||
Db         661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy         120 lnLeuLysMetValLeu-LysLysMetGln 129
               ||  |||||| ||| ||||||||
Db         721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750

RESULT 35
AAS86255
ID   AAS86255 standard; cDNA; 1227 BP.
XX
AC   AAS86255;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #22059.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG22068.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 22059; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1227 BP; 311 A; 239 C; 360 G; 317 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.36e-46      Length:          1227
Score:                  465.00        Matches:         115
Percent Similarity:     54.8%         Conservative:    11
Best Local Similarity:  50.0%         Mismatches:      22
Query Match:            39.0%         Indels:          82
DB:                     5             Gaps:            5

US-10-757-745-2_COPY_121_347 (1-227) x AAS86255 (1-1227)

Qy           2 AspGlyLeuAsp---LeuAsnAsnLeu-SerGlu-ArgAlaArgGlyValCysSerTyrL 20
               ||||||:::      |||||||||||| |||||| ||||||||||||||||||||||||
Db         444 GATGGGATTAGGATCTTAAACAATCTGTTCAGAGGAGGGCTCGAGGGGTGTGTTCCTACT 503

Qy          20 euAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerT 40
               ||||||||            :::::::|||         |||                 ||||
Db         504 TAGCTTTGTAAGTATTATCACTTCTATTTAATAGTAATGTG---------------TCTT 548

Qy          40 yrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheT 60
               ||:::
Db         549 ATGTA---------------------------------------------------------- 553

Qy          60 hrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheP 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         553 ---------------------------------------------------------- 553

Qy          80 roSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluL 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         553 ---------------------------------------------------------- 553

Qy         100 euCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnG 120
                                                          |||||||||||||
Db         554 ------------------------------------------TATACGGAACGAATGAATC 572

Qy         120 lnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheA 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         573 AGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTG 632

Qy         140 laGlyAspThrAsnLeuArgAspArgGluValThr---ArgCysGlyGlyLeuProAsnA 159
               |||||||||||||||||||||||||||||:::   :::         :::      :
Db         633 CAGGAGATACAAATCTAAGGGATCGAGAGGTGAGTGACAAAGTCAGTTTGATGTTTGAGG 692

Qy         159 snIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspT 179
               ::  :::   :::|||     :::     |||     |||||||||||||||||||
Db         693 ACAGATTATTGCTATGGAGACTAATTTAGGTTCCT---TTTGCCAGTATACATGGGATA 749

Qy         179 hrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIleP 199
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         750 CACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATAT 809

Qy         199 hePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyL 219
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         810 TTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGAT 869

Qy         219 euGluLysLeuAspCysGlyArgPhe 227
               |||||||||||||||||||||||||
Db         870 TAGAAAAACTGGACTGTGGTAGATTT 895

RESULT 36
ABK27684
ID   ABK27684 standard; cDNA; 176 BP.
XX
AC   ABK27684;
```

```
XX
DT   09-APR-2002  (first entry)
XX
DE   Human colon cancer expressed sequence tag, Seq ID no 121.
XX
KW   Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
KW   expressed sequence tag.
XX
OS   Homo sapiens.
XX
PN   WO200196390-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US018577.
XX
PR   09-JUN-2000; 2000US-0210821P.
PR   18-DEC-2000; 2000US-0256571P.
PR   10-MAY-2001; 2001US-0290240P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Hepler WT,  Clapper JD,  Wang A,  Secrist H;
XX
DR   WPI; 2002-139708/18.
XX
PT   Novel isolated polynucleotide encoding a polypeptide comprising a portion
PT   of colon tumor protein, useful for detection, diagnosis and therapy of
PT   human colon cancer.
XX
PS   Claim 1; Page 182; 220pp; English.
XX
CC   The invention relates to an isolated polynucleotide (I) encoding a
CC   polypeptide (II) comprising at least a portion of a colon tumour protein.
CC   (I), (II) and antibody (III) to (II) are useful for determining the
CC   presence of a cancer in a patient. (I), (II) or antigen presenting cells
CC   expressing (I) is useful for stimulating and/or expanding T cells
CC   specific for a tumour protein, by contacting T cells with (I), (II) or
CC   antigen-presenting cells that express (I), under conditions and for a
CC   time sufficient to permit the stimulation and/or expansion of T cells.
CC   (I), (II), or antigen presenting cells that express (II) are useful for
CC   treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
CC   isolated from a patient with (I), (II) or antigen presenting cells that
CC   express (II), such that T cells proliferate, and administering to the
CC   patient an effective amount of the proliferated T cells, thus inhibiting
CC   the development of a cancer in the patient. (I) or (II) is useful in
CC   vaccines and pharmaceutical compositions for prevention and treatment of
CC   colon malignancies and for the diagnosis and monitoring of such cancers.
CC   (I), (II) or (III) is useful for detection, diagnosis and/or therapy of
CC   human colon cancer. (I) is useful as a probe or primer for nucleic acid
CC   hybridisation, and in the design and preparation of ribozyme molecules
CC   for inhibiting expression of (II) in tumour cells. ABK27564-ABK27807
CC   represent novel human colon cancer coding sequences and primers of the
CC   invention
XX
SQ   Sequence 176 BP; 68 A; 27 C; 31 G; 50 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.56e-27      Length:          176
Score:                  298.00        Matches:         58
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            25.0%         Indels:          0
DB:                     6             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x ABK27684 (1-176)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 62

Qy         44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 122

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176

RESULT 37
ABV95959/c
ID   ABV95959 standard; cDNA; 176 BP.
XX
AC   ABV95959;
XX
DT   14-JAN-2003  (first entry)
XX
DE   Human pancreatic cancer expressed cDNA SEQ ID NO 1367.
XX
KW   Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW   cytostatic; tumour; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200260317-A2.
XX
PD   08-AUG-2002.
XX
PF   30-JAN-2002; 2002WO-US002781.
XX
PR   30-JAN-2001; 2001US-0265305P.
PR   31-JAN-2001; 2001US-0265682P.
PR   09-FEB-2001; 2001US-0267568P.
PR   21-MAR-2001; 2001US-0278651P.
PR   28-APR-2001; 2001US-0287112P.
PR   16-MAY-2001; 2001US-0291631P.
PR   12-JUL-2001; 2001US-0305484P.
PR   20-AUG-2001; 2001US-0313999P.
PR   27-NOV-2001; 2001US-0333626P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Benson DR,  Kalos MD,  Lodes MJ,  Persing DH,  Hepler WT,  Jiang Y;
XX
DR   WPI; 2002-627435/67.
XX
PT   New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT   diagnosing, preventing and/or treating cancer, particularly pancreatic
PT   cancer.
XX
PS   Claim 1; SEQ ID NO 1367; 300pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (I) comprising: (a)
CC   any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
CC   complements of (a); (c) sequences consisting of at least 20 contiguous
CC   residues of (a); (d) sequences that hybridize to (a), under moderately
CC   stringent conditions; (e) sequences having at least 75% or 90% identity
CC   to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
CC   ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer
CC   in a patient and compositions comprising polypeptides, polynucleotides,
CC   antibodies, fusion proteins, T cell populations and antigen presenting
CC   cells expressing the polypeptide are useful in treating pancreatic cancer
CC   and stimulating an immune response. The polynucleotides can be used as
CC   probes or primers for nucleic acid hybridisation, in the design and
CC   preparation of ribozyme molecules for inhibiting expression of the tumour
CC   polypeptides and proteins in the tumour cells, in vaccines and for gene
CC   therapy. Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 176 BP; 50 A; 31 C; 27 G; 68 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.56e-27      Length:          176
Score:                  298.00        Matches:         58
Percent Similarity:     100.0%        Conservative:    0
```

```
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            25.0%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ABV95959 (1-176)

Qy          24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy          44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy          64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 38
ABL36461/c
ID   ABL36461 standard; cDNA; 179 BP.
XX
AC   ABL36461;
XX
DT   08-APR-2002  (first entry)
XX
DE   Human colon tumour antigen polynucleotide SEQ ID NO:50.
XX
KW   Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW   colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200196388-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US018557.
XX
PR   09-JUN-2000; 2000US-0210899P.
PR   20-FEB-2001; 2001US-0270216P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Harlocker SL,  Secrist H;
XX
DR   WPI; 2002-114514/15.
XX
PT   Novel isolated colon tumor polynucleotide differentially expressed in
PT   colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT   useful for inhibiting development of cancer in patient.
XX
PS   Claim 1; SEQ ID NO 50; 105pp; English.
XX
CC   ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC   which were isolated from human colon tumour and colon metastatic tumour
CC   cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC   production. (I) can be used for stimulating and/or expanding T cells
CC   specific for a tumour protein on contact with the T cells. They are also
CC   useful for inhibiting the development of cancer in a patient. (I) can be
CC   used as probes or primers for nucleic acid hybridisation, for preparing
CC   mutant species primers, or primers for use in genetic constructions. (I)
CC   can be used in the diagnosis of a colon tumour
XX
SQ   Sequence 179 BP; 51 A; 31 C; 27 G; 68 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              2.63e-27        Length:         179
Score:                  298.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            25.0%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ABL36461 (1-179)

Qy          24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy          44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy          64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 39
AAX40590
ID   AAX40590 standard; cDNA; 483 BP.
XX
AC   AAX40590;
XX
DT   18-JUN-1999  (first entry)
XX
DE   Human secreted protein 5' EST SEQ ID No: 190.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS   Homo sapiens.
XX
PN   WO9906550-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB001232.
XX
PR   01-AUG-1997;   97US-00905144.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Lacroix B;
XX
DR   WPI; 1999-153780/13.
DR   P-PSDB; AAY11868.
XX
PT   New isolated prostate-derived nucleic acids - used to develop products
PT   which may have cytokine, immune regulatory, haematopoiesis regulating,
PT   anti-inflammatory or tumour inhibition activity.
XX
PS   Claim 1; Page 298; 675pp; English.
XX
CC   AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
CC   human secreted proteins expressed in prostate, and encode the proteins
CC   given in AAY11716 to AAY11993 respectively. The proteins given represent
CC   the signal peptide and an N-terminal fragment of a secreted protein. The
CC   nucleic acid sequences can be used for producing secreted human gene
CC   products. They can also be used to develop products for diagnosis and
CC   therapy. The proteins obtained may have cytokine activity, cell
CC   proliferation and differentiation activity, haematopoiesis regulating
CC   activity, tissue growth regulating activity, reproductive hormone
CC   regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC   thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, tumour inhibition activity or other activities. The products
CC   can be used in forensic, gene therapy and chromosome mapping procedures.
CC   The sequences can also be used for obtaining corresponding promoter
CC   sequences. The nucleic acids encoding the signal peptides can be used for
CC   directing extracellular secretion of a polypeptide or the insertion of a
CC   polypeptide into a membrane, or importing a polypeptide into a cell
XX
```

```
SQ   Sequence 483 BP; 123 A; 111 C; 139 G; 110 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.24e-16        Length:         483
Score:                  214.00          Matches:        41
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            18.0%           Indels:         0
DB:                     2               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AAX40590 (1-483)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 Leu 41
              |||
Db        481 CTA 483

RESULT 40
AAC47239
ID   AAC47239 standard; DNA; 1566 BP.
XX
AC   AAC47239;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 53077.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
PR   27-JUL-1999;   99US-0145918P.
```

PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	4.83e-14	Length:	1566
Score:	201.00	Matches:	69
Percent Similarity:	46.3%	Conservative:	32
Best Local Similarity:	31.7%	Mismatches:	85
Query Match:	16.9%	Indels:	32
DB:	3	Gaps:	9

US-10-757-745-2_COPY_121_347 (1-227) x AAC47239 (1-1566)

```
Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
              :::   |||:::||||||           |||  |||   :::         :::
Db        550 GAAGACCTTGAGTTGAAT---------CTAAGGATGAGAGCGATTGGTCATCTTATTCAG 600

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
              |||:::||||||  :::|||      |||||||||   |||      |||
Db        601 CTTCATTCTCCTCATCTCATATGTTTTCAGGAAGTCACTCCTGAGATATATGATATTTTC 660

Qy         42 LysLysArg-----------------SerSerAsnTyrGluIleIleThrGlyHisGlu 55
              :::|||                   |||   :::      :::
Db        661 CGCAAGTCGAATTGGTGGAAAGCGTATTCTTGTTCTGTCTCAGTTGATGTGGCAGTTTCG 720

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
            ||||||:::      ::::::         |||:::::   :::|||||
Db        721 AGAGGCTACTACTGCATGCTGTTG---------AGCAAGTTGGGAGTGAAATCTTTCAGC 771

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
                  |||   ::::::  |||   |||  |||          |||  |||
Db        772 AGCAAAAGCTTTGGGAACTCGATAATGGGAAGAGAACTATCTATCGCAGAAGTTGAAGTT 831

Qy         96 SerGlyAsnGlu---LeuCysLeuMetThrSerHisLeuGluSer--------------- 109
              |||   :::   |||        ||||||||||||||||||
Db        832 CCAGGGAGAAAGCCACTTGTTTTTGCAACTAGCCATCTTGAGAGTCCCTGTCCAGGTCCT 891

Qy        110 ---------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
                      :::     :::   ||||||:::   |||  |||      ::::::
Db        892 CCAAAATGGGACCAAATGTTCAGTCGAGAACGCGTTGAACAAGCCAAGGAGGCAATTGAG 951

Qy        127 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 146
               ::::::          :::|||  |||||||||  ||||||  |||
Db        952 ATTCTCAGA---------CCTAATGCCAATGTGATATTCGGGGGAGACATGAACTGGTGC 1002

Qy        147 AspArgGluValThrArgCysGlyGly---------LeuProAsnAsnIleValAspVal 163
              |||           :::      |||         ||||||:::      |||||||||
Db       1003 GAC-----------AAGCTAGATGGGAAGTTTCCTTTGCCAGACAAATGGGTGGATGTT 1050

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
              ||||||   |||  ||||||           :::|||:::|||||:::   |||
Db       1051 TGGGAGGTTTTG---AAACCAGGTGATTTGGGTTTTACATATGACACAAAAGCAAACCCA 1107

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
               |||          |||  :::  |||   |||||||||      |||
Db       1108 ATGTTGTCCGGTAACCGGGCATTGCAGAAACGGTTAGACCGGATTTTGTGTCGT 1161
```

RESULT 41
AAC41089
ID AAC41089 standard; DNA; 1606 BP.
XX
AC AAC41089;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 30607.
XX
KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.

```
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

Alignment Scores:
Pred. No.:                5.01e-14        Length:          1606
Score:                    201.00          Matches:         69
Percent Similarity:       46.3%           Conservative:    32
Best Local Similarity:    31.7%           Mismatches:      85
Query Match:              16.9%           Indels:          32
DB:                       3               Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x AAC41089 (1-1606)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
              :::   |||:::||||||           |||   |||   :::         :::
Db        539 GAAGACCTTGAGTTGAAT---------CTAAGGATGAGAGCGATTGGTCATCTTATTCAG 589

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
              |||:::||||||   :::|||       |||||||||   |||       |||
Db        590 CTTCATTCTCCTCATCTCATATGTTTTCAGGAAGTCACTCCTGAGATATATGATATTTTC 649

Qy         42 LysLysArg-----------------SerSerAsnTyrGluIleIleThrGlyHisGlu 55
              :::|||                 |||   :::       :::
Db        650 CGCAAGTCGAATTGGTGGAAAGCGTATTCTTGTTCTGTCTCAGTTGATGTGGCAGTTTCG 709

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||:::       ::::::          |||::::::   :::||||||
Db        710 AGAGGCTACTACTGCATGCTGTTG---------AGCAAGTTGGGAGTGAAATCTTTCAGC 760

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
                    |||   ::::::   |||   |||   |||             |||   |||
Db        761 AGCAAAAGCTTTGGGAACTCGATAATGGGAAGAGAACTATCTATCGCAGAAGTTGAAGTT 820

Qy         96 SerGlyAsnGlu---LeuCysLeuMetThrSerHisLeuGluSer--------------- 109
              |||   :::   |||           ||||||||||||||||
Db        821 CCAGGGAGAAAGCCACTTGTTTTTGCAACTAGCCATCTTGAGAGTCCCTGTCCAGGTCCT 880

Qy        110 ---------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
                      :::       :::   ||||||:::   |||   |||       ::::::
Db        881 CCAAAATGGGACCAAATGTTCAGTCGAGAACGCGTTGAACAAGCCAAGGAGGCAATTGAG 940

Qy        127 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 146
              ::::::           :::|||   |||||||||   ||||||   |||
Db        941 ATTCTCAGA---------CCTAATGCCAATGTGATATTCGGGGGAGACATGAACTGGTGC 991

Qy        147 AspArgGluValThrArgCysGlyGly---------LeuProAsnAsnIleValAspVal 163
              |||       :::       |||          ||||||:::          |||||||||
Db        992 GAC------------AAGCTAGATGGGAAGTTTCCTTTGCCAGACAAATGGGTGGATGTT 1039

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
              ||||||   |||   ||||||              :::|||:::||||||:::   |||
Db       1040 TGGGAGGTTTTG---AAACCAGGTGATTTGGGTTTTACATATGACACAAAAGCAAACCCA 1096

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                 |||           |||   :::   |||   |||||||||      |||
Db       1097 ATGTTGTCCGGTAACCGGGCATTGCAGAAACGGTTAGACCGGATTTTGTGTCGT 1150

RESULT 42
AAI15311
ID   AAI15311 standard; DNA; 444 BP.
XX
AC   AAI15311;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #5244 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 5244; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
```

```
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 444 BP; 138 A; 60 C; 80 G; 166 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.46e-12        Length:         444
Score:                  182.50          Matches:        37
Percent Similarity:     74.5%           Conservative:   4
Best Local Similarity:  67.3%           Mismatches:     11
Query Match:            15.3%           Indels:         3
DB:                     4               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x AAI15311 (1-444)

Qy          15 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 31
               |||:::|||      |||                |||||||||||||||||||||||||
Db         221 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 280

Qy          32 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 51
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         281 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 340

Qy          52 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 66
               ||||||:::        |||   |||   :::     :::|||
Db         341 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 385

RESULT 43
AAI24484
ID   AAI24484 standard; DNA; 691 BP.
XX
AC   AAI24484;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #14417 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 14417; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 691 BP; 230 A; 81 C; 112 G; 268 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.76e-12        Length:         691
Score:                  182.50          Matches:        37
Percent Similarity:     74.5%           Conservative:   4
Best Local Similarity:  67.3%           Mismatches:     11
Query Match:            15.3%           Indels:         3
DB:                     4               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x AAI24484 (1-691)

Qy          15 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 31
               |||:::|||      |||                |||||||||||||||||||||||||
Db          83 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 142

Qy          32 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 51
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         143 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 202

Qy          52 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 66
               ||||||:::        |||   |||   :::     :::|||
Db         203 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 247

RESULT 44
ACN59762
ID   ACN59762 standard; cDNA; 525 BP.
XX
AC   ACN59762;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Cotton gynoecium tissue EST Clone ID: LIB3829-034-Q1-K6-G12, SEQ:14543.
XX
KW   Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoecium;
KW   variety Nucotton33B; library LIB3829; molecular tag; molecular marker;
KW   genetic mapping; molecular mapping; seed germination; plant growth;
KW   plant quality; plant yield; plant breeding; tissue printing; ss.
XX
OS   Gossypium hirsutum.
XX
PN   US2004123340-A1.
XX
PD   24-JUN-2004.
XX
PF   12-DEC-2001; 2001US-00021323.
XX
PR   14-DEC-2000; 2000US-0255619P.
XX
PA   (DEIK/) DEIKMAN J.
PA   (FENG/) FENG P C C.
PA   (FINC/) FINCHER K L.
PA   (ZIEG/) ZIEGLER T E.
XX
PI   Deikman J,  Feng PCC,  Fincher KL,  Ziegler TE;
XX
DR   WPI; 2004-479808/45.
XX
PT   New isolated nucleic acid molecule that encodes a plant protein or its
PT   fragment, useful for isolating a variety of agronomically significant
PT   genes associated with plant growth, quality or yield, and as molecular
PT   tags to map genes.
XX
PS   Claim 1; SEQ ID NO 14543; 34pp; English.
XX
CC   The invention relates to 17880 cotton expressed sequence tags (ESTs;
```

```
CC   ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC   from primed or non-primed seeds from variety DP50B, mature seeds from
CC   variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC   tissue, developing fibres, carpel walls and septa from variety
CC   Nucotton33B. The invention also relates to substantially purified
CC   proteins or their fragments encoded by nucleic acid molecules of the
CC   invention, and to transformed plants having a nucleic acid construct
CC   comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC   molecular tags to isolate genetic regions, to isolate genes, to map
CC   genes, to determine gene function and to determining whether genes are
CC   members of a particular gene family. The nucleic acid molecules may be
CC   used for isolating a variety of agronomically significant genes
CC   associated with plant growth, quality, yield, and could also serve as
CC   links in metabolic and catabolic pathways. The nucleic acid molecules are
CC   also useful for identifying genes important in initiating and maintaining
CC   seed germination or that may be used to mitigate stresses encountered
CC   during seed germination. The ESTs additionally enable the acquisition of
CC   promoters and cis-regulatory elements which will be useful to express
CC   agronomically significant genes in these tissues and/or other tissues,
CC   and also permits the acquisition of molecular markers useful in breeding
CC   schemes, genetic and molecular mapping, and in cloning of agronomically
CC   significant genes. The nucleic acid molecules are further useful for
CC   detecting the expression level or pattern of a protein or mRNA and for
CC   detecting the presence or quantity of a protein by tissue printing. The
CC   present sequence represents a specifically claimed EST isolated from a
CC   cotton variety Nucotton33B gynoecium tissue cDNA library (LIB3829). The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the US
CC   patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX
SQ   Sequence 525 BP; 141 A; 88 C; 134 G; 162 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.64e-09        Length:         525
Score:                  158.50          Matches:        53
Percent Similarity:     46.1%           Conservative:   29
Best Local Similarity:  29.8%           Mismatches:     63
Query Match:            13.3%           Indels:         33
DB:                     13              Gaps:           7

US-10-757-745-2_COPY_121_347 (1-227) x ACN59762 (1-525)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              :::|||   :::   :::         :::   |||:::||||||||||:::|||
Db         19 AAGAGGATGAAAGCTATTGGTGACCTTATACAACTTCATTCCCCTGATATCATTTGTTTT 78

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||   |||     |||                 |||   ||||||:::
Db         79 CAAGAGGTTACTCCGATGATTTATGACATTTTT------CGTGGGTCTAATTGG------ 126

Qy         51 IleThrGlyHisGluGluGlyTyrPheThrAlaIle---------------------- 62
                        :::||||||      :::|||
Db        127 ------------TGGAAAGGGTACCGGTGCTCTATCTCGGACAACACGGCGAGTTTGAGA 174

Qy         63 ---------MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 79
                      ::::::  |||:::::::  :::|||||             ||||||
Db        175 GCTTACTTTTGTATACAGTTAAGCAAATTGCCTGTGAAATCCTTTCGTCGTGAACCTTTT 234

Qy         80 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 99
                 ::::::   |||    |||   |||             |||   |||||||||   |||
Db        235 GATAATTCGATAATGGGGCGAGAACTTTGCATGACTGAGGTTGAAGTTTCTGGGGGTGAA 294

Qy        100 ------LeuCysLeuMetThrSerHisLeuGluSer----------------------- 109
                    |||   :::   |||:::||||||||||||
Db        295 ACAACGCTTGTTGTTGCTACTACCCATCTCGAGAGTCCGTGTCCTGCTCCGCCGAAATGG 354

Qy        110 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 129
                 :::   ::::::   ||||||:::   |||  |||:::  :::   :::::
Db        355 GATCAAATGTATAGTAAGGAACGTGTAGAACAAGCGAAAGTAGCGATCAATATTCTCGAG 414

Qy        130 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 149
                 |||           |||||||||   ||||||   |||      |||:::
Db        415 AATAACCCG---------AATGTTATTTTTGGGGGTGATATGAATTGGGATGACAAATTA 465
```

```
Qy        150 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                  |||         |||   :::     :::|||   |||       |||
Db        466 GATGGTCGATTT---CCTTTAACTGATGGATGGATTGATGCTTGGAACGAGTTA 516

RESULT 45
ABT07150/c
ID   ABT07150 standard; cDNA; 268 BP.
XX
AC   ABT07150;
XX
DT   07-NOV-2002  (first entry)
XX
DE   Human ovarian cancer associated coding sequence SEQ ID NO: 312.
XX
KW   Human; ovarian cancer; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2002076715-A1.
XX
PD   20-JUN-2002.
XX
PF   06-JUN-2001; 2001US-00876889.
XX
PR   23-SEP-1998;   98US-00159320.
PR   08-FEB-1999;   99US-00246429.
PR   16-SEP-1999;   99US-00397787.
XX
PA   (BENS/) BENSON D R.
PA   (LODE/) LODES M J.
PA   (MITC/) MITCHAM J L.
PA   (KING/) KING G E.
XX
PI   Benson DR,  Lodes MJ,  Mitcham JL,  King GE;
XX
DR   WPI; 2002-598720/64.
XX
PT   Composition for detecting and treating ovarian cancer, comprises a
PT   specific polypeptide, polynucleotide, T cell population, or antigen
PT   presenting cell.
XX
PS   Example 1; Page 116; 188pp; English.
XX
CC   The present invention relates to a method of detecting the presence of
CC   ovarian cancer in a patient, involving detecting ovarian cancer
CC   associated polynucleotides. The method is not only used to detect the
CC   presence of cancer, preferably ovarian cancer in a patient, but also is
CC   used to stimulate and/or expand T cells specific for an ovarian tumour
CC   protein. The sequences can be used in vaccines used to treat cancer. The
CC   present sequence is an ovarian cancer associated coding sequence
XX
SQ   Sequence 268 BP; 87 A; 62 C; 43 G; 76 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.3e-07         Length:         268
Score:                  134.00          Matches:        26
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            11.2%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x ABT07150 (1-268)

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAA 208

Qy        222 LeuAspCysGlyArgPhe 227
              ||||||||||||||||||
Db        207 CTGGACTGTGGTAGATTT 190

RESULT 46
```

ABX73028/c
ID ABX73028 standard; cDNA; 268 BP.
XX
AC ABX73028;
XX
DT 14-MAR-2003 (first entry)
XX
DE Human ovarian carcinoma antigen partial cDNA sequence #294.
XX
KW Human; cancer detection; ovarian carcinoma antigen; ovarian cancer;
KW tumour antigen; tumour; OV2; OV3; OV6; OV9; OV10; OV12; OV14; OV17; OV18;
KW OV23; OV24; OV27; OV41; OV54; OV57; gene; ss.
XX
OS Homo sapiens.
XX
PN US6468758-B1.
XX
PD 22-OCT-2002.
XX
PF 16-SEP-1999; 99US-00397787.
XX
PR 23-SEP-1998; 98US-00159320.
PR 08-FEB-1999; 99US-00246429.
XX
PA (CORI-) CORIXA CORP.
XX
PI Benson DR, Lodes MJ, Mitcham JL, King GE;
XX
DR WPI; 2003-147101/14.
XX
PT Determining presence or absence of cancer in patient by contacting
PT patient sample with oligonucleotide that hybridizes to polynucleotide
PT encoding ovarian carcinoma antigen, and detecting amount of hybridization
PT in sample.
XX
PS Example 1; Col 223-224; 152pp; English.
XX
CC The present invention relates to compositions and methods for determining
CC presence or absence of cancer in patient. The method comprising
CC contacting a biological sample with an oligonucleotide that hybridises to
CC partial polynucleotide sequence encoding for human ovarian carcinoma
CC antigen, or its complement. The method and compositions are useful for
CC the therapy and diagnosis of ovarian cancer. The method can be used to
CC identify tumour antigens that are secreted from ovarian carcinoma and/or
CC other tumours. Effective cancer detection is achieved using the method of
CC the invention. ABX72720-ABX73050 represent partial cDNA sequences
CC encoding human ovarian carcinoma antigens
XX
SQ Sequence 268 BP; 87 A; 62 C; 43 G; 76 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.3e-07	Length:	268
Score:	134.00	Matches:	26
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	11.2%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x ABX73028 (1-268)

```
Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAA 208

Qy        222 LeuAspCysGlyArgPhe 227
              ||||||||||||||||||
Db        207 CTGGACTGTGGTAGATTT 190
```

RESULT 47
ABL90668/c
ID ABL90668 standard; cDNA; 1877 BP.
XX
AC ABL90668;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1230.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR P-PSDB; ABB90259.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 4; SEQ ID NO 1230; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1877 BP; 426 A; 523 C; 485 G; 436 T; 0 U; 7 Other;

Alignment Scores:

Pred. No.:	0.102	Length:	1877
Score:	101.50	Matches:	48
Percent Similarity:	32.3%	Conservative:	28
Best Local Similarity:	20.4%	Mismatches:	68
Query Match:	8.5%	Indels:	91
DB:	6	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x ABL90668 (1-1877)

```
Qy         24 SerProAspValIlePheLeuGlnGluVal--------------IleProProTyrTyr 38
              |||   |||::::::   ||||||||||||                :::   |||   :::
Db       1663 AGCGCCGATATTGTTTGCCTGCAGGAAGTGATGGGCGCGCACGAAGTTCATCCGCTGCAT 1604

Qy         39 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly-------------- 53
                          :::|||:::||||||   :::
Db       1603 GTGGAAAACTGGCCCGATACCTCGCACTACGAGTTTCTCGCCGACACTATGTGGAGCGAT 1544
```

```
Qy          54 ------------------------HisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
                                       :::   ||||||:::           :::|||
Db        1543 TTTGCCTACGGTCGCAATGCCGTATACCCGGAAGGGCATCACGGCAACGCCGTACTG--- 1487

Qy          66 LysSerArgValLysLeuLysSerGlnGlu------------------------------ 75
                 ||||||     ::::::     |||
Db        1486 ---TCGCGTTATCCCATTGAACATTATGAGAATCGCGATGTTTCGGTCGATGGTGCGGAA 1430

Qy          76 ------------------------IleIleProPheProSerThrLysMetMetArgAsn 87
                                       |||:::|||   |||  |||          :::
Db        1429 AAGCGCGGCGTGCTCTACTGCCGCATTGTGCCG---CCGATGACCGGAAAAGCGATTCAT 1373

Qy          88 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 107
               ::::::|||||||||:::   :::
Db        1372 GTGATGTGCGTACATCTGGGCCTG---------------------------------- 1349

Qy         108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
                                   |||   |||      ||||||   |||:::   :::
Db        1348 --------------CGTGAGGCGCACCGTCAGGCGCAGCTTGCGATGCTCGCCGAATGG 1304

Qy         128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 147
               :::   |||   |||:::       |||:::   ||||||||   |||
Db        1303 GTGAATGAGCTACCGGACGGCGAACCGGTATTGGTGGCGGGTGATTTCAAT--------- 1253

Qy         148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167

Db        1253 ------------------------------------------------------------ 1253

Qy         168 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 187
                                       |||      :::   |||      |||   :::
Db        1252 ------------------------GACTGGCGGCAAAAAGCTAATCATCCGTTAAAAGTG 1217

Qy         188 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 207
                 |||          |||  ||||||  ||||||
Db        1216 CAGGCCGGA------------CTGGATGAGATTTTTACCCGCGCCCACGGACGCCCGGCG 1169

Qy         208 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 222
                    |||           ||||||   |||::::::::
Db        1168 CGCACGTTTCCGGTGCAATTTCCTCTACTACGACTGGACAGGATC 1124

RESULT 48
ADG79370
ID   ADG79370 standard; cDNA; 8446 BP.
XX
AC   ADG79370;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human secreted protein cDNA of the invention SEQ ID NO:176.
XX
KW   ss; cytostatic; vasotropic; haemostatic; cardiovascular;
KW   gastrointestinal; immunomodulator; inotropic; cerebroprotective;
KW   neuroprotective; nephrotropic; antiinflammatory; antibacterial; virucide;
KW   gynaecological; antidiabetic; gene therapy; vaccine; cancer;
KW   blood disorder; immune disorder; infection; inflammatory disorder;
KW   type II diabetes; gene; human; secreted protein.
XX
OS   Homo sapiens.
XX
PN   WO200268638-A1.
XX
PD   06-SEP-2002.
XX
PF   21-FEB-2002; 2002WO-US005064.
XX
PR   23-FEB-2001; 2001US-0270658P.
PR   12-JUL-2001; 2001US-0304444P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Komatsoulis GA,  Birse CE,  Choi GH,  Olsen HS,  Ni J;
PI   Bell A;
XX
DR   WPI; 2002-750418/81.
DR   P-PSDB; ADG79556.
XX
PT   New isolated polypeptide and encoding polynucleotide useful for
PT   diagnosing, preventing, treating and/or ameliorating diseases such as
PT   cancer, blood disorders, infections, inflammatory and immune disorders
PT   and type II diabetes.
XX
PS   Disclosure; SEQ ID NO 176; 936pp; English.
XX
CC   The invention relates to a novel isolated polypeptide. A protein of the
CC   invention has cytostatic, vasotropic, haemostatic, cardiovascular,
CC   gastrointestinal, immunomodulator, inotropic, cerebroprotective,
CC   neuroprotective, nephrotropic, antiinflammatory, antibacterial, virucide,
CC   gynaecological, and antidiabetic activity. A polynucleotide of the
CC   invention may have a use in gene therapy, and as a vaccine. The methods
CC   and compositions of the invention are useful for diagnosing, preventing,
CC   treating and/or ameliorating diseases such as cancer (neural,
CC   reproductive, gastrointestinal, endocrine, renal, CNS and respiratory
CC   neoplasias), blood disorders, immune disorders, infections, inflammatory
CC   disorders and type II diabetes. They can also be used in chromosome
CC   identification, screening assays and molecular weight markers. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 8446 BP; 1967 A; 2203 C; 2249 G; 2018 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.:                0.896          Length:          8446
Score:                    101.50         Matches:         48
Percent Similarity:       32.3%          Conservative:    28
Best Local Similarity:    20.4%          Mismatches:      68
Query Match:              8.5%           Indels:          91
DB:                       6              Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x ADG79370 (1-8446)

Qy          24 SerProAspValIlePheLeuGlnGluVal---------------IleProProTyrTyr 38
               |||   |||::::::   ||||||||||||               :::   |||   :::
Db        4770 AGCGCCGATATTGTTTGCCTGCAGGAAGTGATGGGCGCGCACGAAGTTCATCCGCTGCAT 4829

Qy          39 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly--------------- 53
                                     :::|||:::||||||   :::
Db        4830 GTGGAAAACTGGCCCGATACCTCGCACTACGAGTTTCTCGCCGACACTATGTGGAGCGAT 4889

Qy          54 ------------------------HisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
                                       :::   ||||||:::           :::|||
Db        4890 TTTGCCTACGGTCGCAATGCCGTATACCCGGAAGGGCATCACGGCAACGCCGTACTG--- 4946

Qy          66 LysSerArgValLysLeuLysSerGlnGlu------------------------------ 75
                 ||||||     ::::::     |||
Db        4947 ---TCGCGTTATCCCATTGAACATTATGAGAATCGCGATGTTTCGGTCGATGGTGCGGAA 5003

Qy          76 ------------------------IleIleProPheProSerThrLysMetMetArgAsn 87
                                       |||:::|||   |||  |||          :::
Db        5004 AAGCGCGGCGTGCTCTACTGCCGCATTGTGCCG---CCGATGACCGGAAAAGCGATTCAT 5060

Qy          88 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 107
               ::::::|||||||||:::   :::
Db        5061 GTGATGTGCGTACATCTGGGCCTG---------------------------------- 5084

Qy         108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
                                   |||   |||      ||||||   |||:::   :::
Db        5085 --------------CGTGAGGCGCACCGTCAGGCGCAGCTTGCGATGCTCGCCGAATGG 5129

Qy         128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 147
               :::   |||   |||:::       |||:::   ||||||||   |||
Db        5130 GTGAATGAGCTACCGGACGGCGAACCGGTATTGGTGGCGGGTGATTTCAAT--------- 5180

Qy         148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167

Db        5180 ------------------------------------------------------------ 5180

Qy         168 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 187
```

```
Db       5181 ------------------------GACTGGCGGCAAAAAGCTAATCATCCGTTAAAAGTG 5216

Qy        188 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 207

Db       5217 CAGGCCGGA------------CTGGATGAGATTTTTACCCGCGCCCACGGACGCCCGGCG 5264

Qy        208 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 222

Db       5265 CGCACGTTTCCGGTGCAATTTCCTCTACTACGACTGGACAGGATC 5309
```

RESULT 49
ADS45761
ID ADS45761 standard; cDNA; 1644 BP.
XX
AC ADS45761;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #504.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 24191; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1644 BP; 463 A; 339 C; 388 G; 454 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.197	Length:	1644
Score:	98.50	Matches:	57
Percent Similarity:	40.9%	Conservative:	49
Best Local Similarity:	22.0%	Mismatches:	90
Query Match:	8.3%	Indels:	63
DB:	13	Gaps:	14

US-10-757-745-2_COPY_121_347 (1-227) x ADS45761 (1-1644)

```
Qy          3 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 22

Db        253 GGTGAAAAGATTAACCGTACTGAAGATCGCGCT---------GTATTGCATATTGCTTTG 303

Qy         23 ------------TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 38

Db        304 CGTAACCGCAGTAACTCTCCGATTAAGGTTGATGAGGAAGATGTTATGCCGCAGGTCAAT 363

Qy         39 SerTyrLeu---LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGly 57

Db        364 GCGGTACTGGCAAAAATGAAATCATTCAGTGAGCGCATTATTAGTGGGGAGTGGAAAGGT 423

Qy         58 TyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle 77

Db        424 TATACCGGTAAAGCTATTACTGATGTGGTAAACATCGGCATTGGTGGTTCAGATCTTGGC 483

Qy         78 ProPheProSerThr---------LysMetMetArgAsnLeuLeuCysValHisVal--- 93

Db        484 CCATATATGGTGACAGAGGCACTGAAAGCCTATAAGAACCACCTTAATATGCACTTTGTA 543

Qy         94 ---AsnValSerGlyAsnGluLeuCys-------------------------------- 101

Db        544 TCCAATGTGGATGGCACTCATATCGCGGAAACACTGAAGTCTCTGGATCCGGAAACCACG 603

Qy        102 --------------------LeuMetThrSerHisLeuGluSerThrArgGlyHisAla 114

Db        604 CTGTTCCTGATTGCATCGAAGACTTTTACCACGCAGGAAACGATGACTAACGCACATTCT 663

Qy        115 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 134

Db        664 GCC---CGTGATTGGTTCTTGCAAAGTGCAGTTGATGAAGTACATGTTGCCAAACATTTT 720

Qy        135 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 154

Db        721 GCGGCACTT------------TCAACGAAT------GAACAGGAAGTCAGCAAGTTCGGT 762

Qy        155 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 174

Db        763 ATTGATCCGCAAAACATGTTTGAGTTCTGGGACTGGGTCGGCGGC------------CGT 810

Qy        175 TyrThr---TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeu 193

Db        811 TATTCTCTTTGG--------------TCTGCGATTGGTTTATCCATTGTCCTGTCCATT 855

Qy        194 ArgPheAspArgIle---------------------------------------PhePhe 200

Db        856 GGATATGATAATTTCGAGCAATTGCTGAATGGTGCCCATGCTATGGATAAGCACTTTGCA 915

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeu 219

Db        916 CAAACTGCTCCTGAACAGAAT------ATTCCGGTTATTCTGGCGCTGATTGGCATC 966
```

RESULT 50
ADQ97808

```
ID   ADQ97808 standard; DNA; 2928 BP.
XX
AC   ADQ97808;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Mouse cancer associated sequence MR11-001, SEQ ID 785.
XX
KW   Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX
OS   Mus musculus.
XX
PN   WO2004060304-A2.
XX
PD   22-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US041389.
XX
PR   27-DEC-2002; 2002US-00330773.
XX
PA   (SAGR-) SAGRES DISCOVERY INC.
XX
PI   Morris DW,  Malandro MS;
XX
DR   WPI; 2004-543781/52.
XX
PT   New isolated cancer associated nucleic acids comprising at least 10
PT   contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT   cancers such as leukemia and lymphoma.
XX
PS   Claim 1; SEQ ID NO 785; 199pp; English.
XX
CC   The present invention relates to cancer associated sequences (ADQ97025-
CC   ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC   treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic formate directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2928 BP; 679 A; 840 C; 704 G; 685 T; 0 U; 20 Other;

Alignment Scores:
Pred. No.:              0.601          Length:          2928
Score:                  97.50          Matches:         51
Percent Similarity:     37.1%          Conservative:    40
Best Local Similarity:  20.8%          Mismatches:      111
Query Match:            8.2%           Indels:          43
DB:                     12             Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x ADQ97808 (1-2928)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              ||||||          :::         :::      |||   ||||||::::::   |||
Db       1041 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 1100

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||       |||:::       :::         |||
Db       1101 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 1148

Qy         51 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 59
                 |||::::::     :::|||
Db       1149 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 1205

Qy         60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                  |||:::         ::::::|||   ||||||
Db       1206 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 1265

Qy         73 SerGlnGluIleIleProProProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
              |||      |||        :::     :::       |||   :::   :::
Db       1266 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 1325

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                       ||||||   :::   |||:::      :::||||||:::

Db       1326 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 1385

Qy         109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
               :::      |||           |||      |||        :::|||:::   :::
Db       1386 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 1445

Qy         129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
               |||        :::   |||   :::      |||      ||||||   |||
Db       1446 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 1505

Qy         147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                  |||                                  |||:::         :::
Db       1506 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 1565

Qy         165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
               :::   |||         |||:::   :::           ||||||
Db       1566 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 1613

Qy         182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                         :::      :::      |||:::      |||   |||::::::
Db       1614 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 1667

Qy         202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                  |||                  :::      :::||||||||||||      |||:::
Db       1668 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 1718

Qy         222 LeuAspCysGlyArg 226
               :::      |||
Db       1719 ATTGGGCCCAACCGG 1733

RESULT 51
ADZ62001
ID   ADZ62001 standard; DNA; 9737 BP.
XX
AC   ADZ62001;
XX
DT   30-JUN-2005  (first entry)
XX
DE   Murine Ccr4 gene, SEQ ID 1725.
XX
KW   Drug screening; Peroxisome Proliferator-Activated Receptor gamma;
KW   PPAR gamma; gene; ds.
XX
OS   Mus musculus.
XX
PN   US2005084872-A1.
XX
PD   21-APR-2005.
XX
PF   23-JAN-2004; 2004US-00764420.
XX
PR   24-JAN-2003; 2003US-0442797P.
PR   30-MAY-2003; 2003US-0474413P.
XX
PA   (LUMP/) LUM P Y.
PA   (TANY/) TAN Y.
PA   (DAIH/) DAI H.
PA   (MUIS/) MUISE E S.
PA   (BERG/) BERGER J P.
PA   (THOM/) THOMPSON J R.
XX
PI   Lum PY,  Tan Y,  Dai H,  Muise ES,  Berger JP,  Thompson JR;
XX
DR   WPI; 2005-313963/32.
DR   GENBANK; U70139.
XX
PT   Determining whether agent has defined biological activity, by making
PT   comparison e.g. comparing efficacy value, toxicity value and classifier
PT   value of agent to reference values, and using results to determine
PT   whether agent has activity.
XX
PS   Example 3; SEQ ID NO 1725; 51pp; English.
XX
```

```
CC   The present invention relates to a methods for screening biologically
CC   active agents, such as candidate drugs, to determine whether the agent
CC   possesses a defined biological activity. The methods involve making one
CC   or more comparisons chosen from comparing efficacy value of agent to
CC   reference efficacy value, comparing a toxicity value of the agent to
CC   reference toxicity value, comparing a classifier value of the agent to
CC   reference classifier value, and using the comparison result(s) to
CC   determine whether the agent possesses the defined biological activity.
CC   Also claimed is a population of oligonucleotide probes (I) specific for
CC   measuring the expression levels of members of a classifier population of
CC   genes or a toxicity-related population of genes. (I) are useful for
CC   measuring the expression levels of genes that are useful for identifying
CC   agonists or partial agonists of Peroxisome Proliferator-Activated
CC   Receptor (PPAR) gamma. ADZ62007-ADZ62272 are oligonucleotide probes which
CC   are useful in the method of the invention to measure the expression
CC   pattern of the classifier-related population of mouse genes ADZ61748-
CC   ADZ62006. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from USPTO at seqdata.uspto.gov/sequence.html?DocID=20050084872.
XX
SQ   Sequence 9737 BP; 2569 A; 2228 C; 2243 G; 2686 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.:              3.41          Length:          9737
Score:                  97.50         Matches:         51
Percent Similarity:     37.1%         Conservative:    40
Best Local Similarity:  20.8%         Mismatches:      111
Query Match:            8.2%          Indels:          43
DB:                     14            Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x ADZ62001 (1-9737)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
Db       7856 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 7915

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
Db       7916 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 7963

Qy         51 IleThrGlyHisGluGluGlyTyrPhe---------------------------------- 59
Db       7964 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 8020

Qy         60 ----------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
Db       8021 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 8080

Qy         73 SerGlnGluIleIleProProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
Db       8081 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 8140

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
Db       8141 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 8200

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
Db       8201 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 8260

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
Db       8261 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 8320

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
Db       8321 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 8380

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
Db       8381 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 8428

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201

Db       8429 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 8482

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
Db       8483 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 8533

Qy        222 LeuAspCysGlyArg 226
Db       8534 ATTGGGCCCAACCGG 8548

RESULT 52
ADQ97807
ID   ADQ97807 standard; DNA; 47040 BP.
XX
AC   ADQ97807;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Mouse cancer associated sequence MD11-001, SEQ ID 784.
XX
KW   Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX
OS   Mus musculus.
XX
PN   WO2004060304-A2.
XX
PD   22-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US041389.
XX
PR   27-DEC-2002; 2002US-00330773.
XX
PA   (SAGR-) SAGRES DISCOVERY INC.
XX
PI   Morris DW,  Malandro MS;
XX
DR   WPI; 2004-543781/52.
XX
PT   New isolated cancer associated nucleic acids comprising at least 10
PT   contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT   cancers such as leukemia and lymphoma.
XX
PS   Claim 1; SEQ ID NO 784; 199pp; English.
XX
CC   The present invention relates to cancer associated sequences (ADQ97025-
CC   ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC   treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 47040 BP; 12069 A; 10729 C; 10594 G; 12790 T; 0 U; 858 Other;

Alignment Scores:
Pred. No.:              33.2          Length:          47040
Score:                  97.50         Matches:         51
Percent Similarity:     37.1%         Conservative:    40
Best Local Similarity:  20.8%         Mismatches:      111
Query Match:            8.2%          Indels:          43
DB:                     12            Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x ADQ97807 (1-47040)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
Db      35153 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 35212

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
Db      35213 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 35260

Qy         51 IleThrGlyHisGluGluGlyTyrPhe---------------------------------- 59
```

```
Db      35261 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 35317

Qy         60 ----------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72

Db      35318 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 35377

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91

Db      35378 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 35437

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108

Db      35438 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 35497

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128

Db      35498 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 35557

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146

Db      35558 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 35617

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164

Db      35618 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 35677

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181

Db      35678 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 35725

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201

Db      35726 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 35779

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuGlyLeuGluLys 221

Db      35780 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 35830

Qy        222 LeuAspCysGlyArg 226

Db      35831 ATTGGGCCCAACCGG 35845

RESULT 53
ACN44608/c
ID   ACN44608 standard; DNA; 90043 BP.
XX
AC   ACN44608;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Mouse genomic sequence mCG22175.
XX
KW   Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
OS   Mus musculus.
XX
PN   WO2003073826-A2.
XX
PD   12-SEP-2003.
XX
PF   28-FEB-2003; 2003WO-US006235.
XX
PR   01-MAR-2002; 2002US-00087192.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-328604/31.
XX
PT   Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT   comprises a nucleotide sequence.
XX
PS   Claim 1; SEQ ID NO 1141; 0pp; English.
XX
CC   The present invention relates to novel DNA and protein sequences which
CC   are associated with carcinomas. The sequences are useful for: (i) for
CC   screening drug candidates; (ii) for screening of bioactive agent capable
CC   of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC   a bioactive agent capable of modulating the activity of CAP; (iv) for
CC   evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC   carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC   carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC   (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC   determining Carcinoma Associated (CA) gene copy number. In addition, the
CC   CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC   carcinoma including lymphoma. The present sequence is one such CA coding
CC   sequence. Note: This patent is an equivalent to basic patent
CC   US2002182586A1, for which no sequence data was published
XX
SQ   Sequence 90043 BP; 23656 A; 18143 C; 20242 G; 26180 T; 0 U; 1822 Other;

Alignment Scores:
Pred. No.:                84.7          Length:          90043
Score:                    97.50         Matches:         51
Percent Similarity:       37.1%         Conservative:    40
Best Local Similarity:    20.8%         Mismatches:      111
Query Match:              8.2%          Indels:          43
DB:                       11            Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x ACN44608 (1-90043)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30

Db      85949 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 85890

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50

Db      85889 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 85842

Qy         51 IleThrGlyHisGluGluGlyTyrPhe---------------------------------- 59

Db      85841 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 85785

Qy         60 ----------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72

Db      85784 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 85725

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91

Db      85724 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 85665

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108

Db      85664 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 85605

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128

Db      85604 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 85545

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146

Db      85544 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 85485

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164

Db      85484 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 85425

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181

Db      85424 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 85377

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201

Db      85376 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 85323
```

```
Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
              |||                  :::      :::||||||||||||      |||:::
Db      85322 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 85272

Qy        222 LeuAspCysGlyArg 226
              :::      |||
Db      85271 ATTGGGCCCAACCGG 85257
```

RESULT 54
ACA34492
ID ACA34492 standard; DNA; 1692 BP.
XX
AC ACA34492;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #16149.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Haemophilus influenzae.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU30622.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 22362; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1692 BP; 578 A; 319 C; 301 G; 494 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.313	Length:	1692
Score:	97.00	Matches:	51
Percent Similarity:	41.0%	Conservative:	43
Best Local Similarity:	22.3%	Mismatches:	89
Query Match:	8.1%	Indels:	46
DB:	8	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x ACA34492 (1-1692)

```
Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||   ::::::     ::::::|||:::|||         :::   |||   |||
Db        361 ACGCCAGTGCTTGTTGATGGCAAAGATGTTATGCCTGAAGTCAATGCTGTGCTAGCTAAA 420

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
              :::                |||||::|||      :::||||||
Db        421 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGTAAAGCC 480

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::              :::  :::        ::::::   |||:::      |||
Db        481 ATTACGGATGTTGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 540

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
              :::   :::|||         |||   |||   ::::|||         ||||||   |||
Db        541 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 600

Qy         98 AsnGlu----------------------------------------LeuCysLeuMetThr 104
                                                          |||   |||:::
Db        601 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 660

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              |||          :::|||:::                  :::             :::
Db        661 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCTCAAAGTGCGCGTGATTGGTTA 720

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||      ::::::          |||          |||      ||||||
Db        721 CTAAAAGCAGCGAAAGATGAAAGTGCAGTTGCAAAACATTTTGCAGCATTATCAACCAAT 780

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              :::|||       |||   :::   |||          ||||||:::   :::   |||
Db        781 GCTAAAGAT------GTAGAAAAATTTGGCATTGATACCAATAATATGTTTGAATTTTGG 834

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
              :::::::::|||          :::|||:::   |||             |||
Db        835 GATTGGGTTGGCGGT------------CGTTACTCTTTATGG--------------TCA 867

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe----------------------- 195
              :::|||::::::   |||      :::   |||
Db        868 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAACTTTGAAGCGTTATTAAAT 927

Qy        196 ---------------AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210
                          |||:::   |||              |||:::         |||
Db        928 GGTGCGCATGAAATGGATAAACATTTCCTCTCTACTCCAATCGAACAAAAT------ATC 981

Qy        211 ProArgSerLeuAspLeuLeuGlyLeu 219
              |||   :::|||   |||:::||||||
Db        982 CCAACCACTTTAGCATTAGTTGGTTTA 1008
```

RESULT 55
AAT42063_16
Continuation (17 of 19) of AAT42063 from base 1600001 (Haemophilus influenzae complete g
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

WP	Fragment Name	Begin	End
WP	AAT42063_00	1	110000
WP	AAT42063_01	100001	210000
WP	AAT42063_02	200001	310000
WP	AAT42063_03	300001	410000
WP	AAT42063_04	400001	510000
WP	AAT42063_05	500001	610000
WP	AAT42063_06	600001	710000
WP	AAT42063_07	700001	810000
WP	AAT42063_08	800001	910000
WP	AAT42063_09	900001	1010000
WP	AAT42063_10	1000001	1110000
WP	AAT42063_11	1100001	1210000
WP	AAT42063_12	1200001	1310000
WP	AAT42063_13	1300001	1410000
WP	AAT42063_14	1400001	1510000
WP	AAT42063_15	1500001	1610000
WP	AAT42063_16	1600001	1710000
WP	AAT42063_17	1700001	1810000
WP	AAT42063_18	1800001	1830121

Alignment Scores:

Pred. No.:	130	Length:	110000
Score:	97.00	Matches:	51
Percent Similarity:	41.0%	Conservative:	43
Best Local Similarity:	22.3%	Mismatches:	89
Query Match:	8.1%	Indels:	46
DB:	2	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x AAT42063_16 (1-110000)

```
Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||   ::::::     ::::::|||:::|||         :::   |||   |||
Db      44110 ACGCCAGTGCTTGTTGATGGCAAAGATGTTATGCCTGAAGTCAATGCTGTGCTAGCTAAA 44169

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
              :::                 ||||||:::|||      :::||||||
Db      44170 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGTAAAGCC 44229

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::          :::   :::        ::::::   |||:::      |||
Db      44230 ATTACGGATGTTGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 44289

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
              :::   :::|||         |||   |||   :::|||         ||||||   |||
Db      44290 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 44349

Qy         98 AsnGlu---------------------------------------LeuCysLeuMetThr 104
                                                         |||   |||:::
Db      44350 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 44409

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              |||         :::|||:::                :::                :::
Db      44410 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCTCAAAGTGCGCGTGATTGGTTA 44469

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||      ::::::          |||          |||      ||||||
Db      44470 CTAAAAGCAGCGAAAGATGAAAGTGCAGTTGCAAAACATTTTGCAGCATTATCAACCAAT 44529

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              :::|||       |||   :::   |||         ||||||:::   :::   |||
Db      44530 GCTAAAGAT------GTAGAAAAATTTGGCATTGATACCAATAATATGTTTGAATTTTGG 44583

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
              :::::::::|||          :::|||:::   |||                  |||
Db      44584 GATTGGGTTGGCGGT-----------CGTTACTCTTTATGG--------------TCA 44616

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe---------------------- 195
              :::|||::::::   |||     :::   |||
Db      44617 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAACTTTGAAGCGTTATTAAAT 44676

Qy        196 ---------------AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210
                           |||:::   |||               |||:::       |||
```

```
Db      44677 GGTGCGCATGAAATGGATAAACATTTCCTCTCTACTCCAATCGAACAAAAT------ATC 44730

Qy        211 ProArgSerLeuAspLeuLeuGlyLeu 219
              |||   :::|||   |||:::||||||
Db      44731 CCAACCACTTTAGCATTAGTTGGTTTA 44757
```

RESULT 56
ADT05432/c
ID ADT05432 standard; DNA; 5386 BP.
XX
AC ADT05432;
XX
DT 02-DEC-2004 (first entry)
XX
DE Haemophilus influenzae (NTHi) contig DNA sequence - SEQ ID 468.
XX
KW middle ear bacterial infection; nasopharynx bacterial infection; ds;
KW contig.
XX
OS Haemophilus influenzae.
XX
PN WO2004078949-A2.
XX
PD 16-SEP-2004.
XX
PF 05-MAR-2004; 2004WO-US007001.
XX
PR 06-MAR-2003; 2003US-0453134P.
XX
PA (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI Bakaletz LO, Munson RS, Dyer DW;
XX
DR WPI; 2004-662422/64.
XX
PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.
XX
PS Example 1; SEQ ID NO 468; 88pp; English.
XX
CC The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC nucleic acid represents an NTHi contig sequence of the invention.
XX
SQ Sequence 5386 BP; 1592 A; 1037 C; 958 G; 1799 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.94	Length:	5386
Score:	95.00	Matches:	51
Percent Similarity:	40.6%	Conservative:	42
Best Local Similarity:	22.3%	Mismatches:	90
Query Match:	8.0%	Indels:	46
DB:	13	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x ADT05432 (1-5386)

```
Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||   ::::::     ::::::|||:::|||         :::   |||   |||
Db       2429 ACGCCAGTACTTGTTGATGGTAAAGATGTTATGCCTGAAGTCAATGCTGTGTTAGCTAAA 2370

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
              :::                 ||||||:::|||      :::||||||
Db       2369 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGCAAAGCC 2310

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::          :::   :::        ::::::   |||:::      |||
Db       2309 ATTACGGATGTAGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 2250

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
              :::   :::|||         |||   |||   :::|||         ||||||   |||
```

```
Db       2249 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 2190

Qy         98 AsnGlu----------------------------------------LeuCysLeuMetThr 104
                                                           |||   |||:::
Db       2189 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 2130

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
               |||         :::|||:::                 :::         :::
Db       2129 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCACAAAGTGCGCGTGATTGGTTA 2070

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
               ||||||      ::::::          |||          |||      ||||||
Db       2069 CTAAAAGCAGCGAAAGATGAAAGTGCGGTTGCAAAACATTTTGCAGCATTATCAACCAAT 2010

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
               :::|||      |||  :::  |||         ||||||:::  :::   |||
Db       2009 GCTAAAGAT------GTAGAAAAATTTGGTATTGATACCAATAACATGTTTGAATTTTGG 1956

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
               :::::::::|||                 :::|||:::  |||                |||
Db       1955 GATTGGGTTGGCGGT------------CGTTACTCTTTATGG--------------TCA 1923

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAla 203
               :::|||::::::  |||     :::   |||:::         |||   |||
Db       1922 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAAC------TTTGAAGCGTTA 1869

Qy        204 AlaGluGluGlyHis----------------------------------------IleIle 210
                          |||                                        |||
Db       1868 TTAAATGGCGCGCATGAAATGGATAAACATTTCCGCACTACTCCAATTGAAAAAAATATT 1809

Qy        211 ProArgSerLeuAspLeuLeuGlyLeu 219
               |||   :::|||   |||:::||||||
Db       1808 CCAACTACTTTAGCCTTAGTTGGTTTA 1782
```

RESULT 57
ADT05649/c
ID ADT05649 standard; DNA; 349980 BP.
XX
AC ADT05649;
XX
DT 02-DEC-2004 (first entry)
XX
DE Haemophilus influenzae (NTHi) DNA sequence - SEQ ID 685.
XX
KW middle ear bacterial infection; nasopharynx bacterial infection; ds.
XX
OS Haemophilus influenzae.
XX
PN WO2004078949-A2.
XX
PD 16-SEP-2004.
XX
PF 05-MAR-2004; 2004WO-US007001.
XX
PR 06-MAR-2003; 2003US-0453134P.
XX
PA (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI Bakaletz LO, Munson RS, Dyer DW;
XX
DR WPI; 2004-662422/64.
XX
PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.
XX
PS Claim 1; SEQ ID NO 685; 88pp; English.
XX
CC The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC nucleic acid represents an NTHi DNA sequence of the invention.

XX
SQ Sequence 349980 BP; 105127 A; 63538 C; 70035 G; 111276 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.:	1.22e+03	Length:	349980
Score:	95.00	Matches:	51
Percent Similarity:	40.6%	Conservative:	42
Best Local Similarity:	22.3%	Mismatches:	90
Query Match:	8.0%	Indels:	46
DB:	13	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x ADT05649 (1-349980)

```
Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
               :::|||   ::::::      ::::::|||:::|||          :::   |||   |||
Db     272652 ACGCCAGTACTTGTTGATGGTAAAGATGTTATGCCTGAAGTCAATGCTGTGTTAGCTAAA 272593

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                 :::                 ||||||:::|||      :::||||||
Db     272592 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGCAAAGCC 272533

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
               :::           :::   :::         ::::::   |||:::      |||
Db     272532 ATTACGGATGTAGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 272473

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
               :::   :::|||         |||   |||   :::|||         ||||||   |||
Db     272472 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 272413

Qy         98 AsnGlu----------------------------------------LeuCysLeuMetThr 104
                                                          |||   |||:::
Db     272412 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 272353

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
               |||         :::|||:::                 :::         :::
Db     272352 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCACAAAGTGCGCGTGATTGGTTA 272293

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
               ||||||      ::::::          |||          |||      ||||||
Db     272292 CTAAAAGCAGCGAAAGATGAAAGTGCGGTTGCAAAACATTTTGCAGCATTATCAACCAAT 272233

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
               :::|||      |||  :::  |||         ||||||:::  :::   |||
Db     272232 GCTAAAGAT------GTAGAAAAATTTGGTATTGATACCAATAACATGTTTGAATTTTGG 272179

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
               :::::::::|||                 :::|||:::  |||                |||
Db     272178 GATTGGGTTGGCGGT------------CGTTACTCTTTATGG--------------TCA 272146

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAla 203
               :::|||::::::  |||     :::   |||:::         |||   |||
Db     272145 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAAC------TTTGAAGCGTTA 272092

Qy        204 AlaGluGluGlyHis----------------------------------------IleIle 210
                          |||                                        |||
Db     272091 TTAAATGGCGCGCATGAAATGGATAAACATTTCCGCACTACTCCAATTGAAAAAAATATT 272032

Qy        211 ProArgSerLeuAspLeuLeuGlyLeu 219
               |||   :::|||   |||:::||||||
Db     272031 CCAACTACTTTAGCCTTAGTTGGTTTA 272005
```

RESULT 58
ACA42939
ID ACA42939 standard; DNA; 1650 BP.
XX
AC ACA42939;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #24596.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.

```
XX
OS   Pasteurella multocida.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU39069.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 30809; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1650 BP; 526 A; 307 C; 351 G; 466 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.613          Length:          1650
Score:                    94.50          Matches:         56
Percent Similarity:       38.8%          Conservative:    44
Best Local Similarity:    21.7%          Mismatches:      97
Query Match:              7.9%           Indels:          61
DB:                       8              Gaps:            13

US-10-757-745-2_COPY_121_347 (1-227) x ACA42939 (1-1650)
```

```
Qy          3 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 22
              |||      :::|||         :::||||||                 :::   ||||||
Db        253 GGGGCAAAAATCAATAAAACGGAAGATCGCGCC---------GTATTACATACCGCACTT 303

Qy         23 -----------TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 38
                         ||||||   ::::::      ::::::|||:::|||
Db        304 CGTAATCGTTCTAATTCGCCTGTGTTGGTTGATGGGAAAGATGTGATGCCAGAAGTGAAT 363

Qy         39 SerTyrLeu---LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGly 57
              :::   |||   |||  :::      :::   :::|||:::|||        :::|||
Db        364 GCGGTTTAGCCAAAATGAAAGATTTTTGCCATCGTGTGATTTCTGGTGAATGGAAAGGT 423

Qy         58 TyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle 77
              |||           :::              :::   :::         ::::::
Db        424 TATACTGGTAAAGCGATTACAGATGTGGTGAACATCGGAATTGGTGGGTCTGATTTAGGA 483

Qy         78 ProProSerThrLysMetMetArg---------AsnLeuLeuCysValHisVal--- 93
              |||:::      |||:::   :::|||         |||   |||   :::|||
Db        484 CCTTATATGGTAACAGAAGCCTTACGTCCTTATAAAAATCATTTGAATCTACATTTCGTG 543

Qy         94 ---AsnValSerGlyAsnGluLeuCys---------------------------------- 101
                 ||||||   |||     :::
Db        544 TCTAACGTAGATGGTACACATATTGCGGAGACCTTAAAGAAAGTTAATCCAGAAACCACG 603

Qy        102 --------------------LeuMetThrSerHisLeuGluSerThrArgGlyHisAla 114
                                  |||:::          |||      |||:::
Db        604 CTATTTTTGGTTGCGTCAAAAACCTTTACCACACAAGAAACTATGACCAATGCGCATTCT 663

Qy        115 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 134
              |||                  |||          ::::::
Db        664 GCGCGT--------------AATTGGTTCTTAGCCACGGCAAAAGACGAAAGTCATGTG 708

Qy        135 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 154
              |||          |||      ||||||      :::   |||          |||
Db        709 GCAAAACACTTTGCAGCACTGTCCACAAAC------AGTAAAGCCGTTGCTGAATTTGGT 762

Qy        155 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 174
                  ||||||:::   :::   |||:::::::::|||                :::
Db        763 ATTGATACGAATAATATGTTTGAATTTTGGGATTGGGTAGGGGGT-----------CGT 810

Qy        175 TyrThr---TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeu 193
              |||:::   |||           |||   :::|||::::::   |||      :::
Db        811 TATTCATTATGG--------------TCAGCGATTGGTTTATCTATTGCCCTTTCAATC 855

Qy        194 ArgPheAspArgIle-------------------------------------PhePheArg 201
                 |||:::                                          ||||||
Db        856 GGTTTTGAGCATTTTGAAGCGCTTTTAGCGGGTGCACACGAAATGGATAAACATTTCCGT 915

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeu 219
              |||      |||   :::   ||||||   :::|||   |||:::||||||
Db        916 ACTGCCCCAATTGAGCAAAAT---ATTCCGACCACATTAGCGTTAATTGGTTTG 966
```

```
RESULT 59
ADE43923
ID   ADE43923 standard; cDNA; 1767 BP.
XX
AC   ADE43923;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Human DP160 encoding cDNA SEQ ID NO:33.
XX
KW   DP119; DP444; DP810; DP685; WE474; DP160; RA977; RA770; antidiabetic;
KW   anorectic; neuroprotective; gene therapy; pancreatic disorder; diabetes;
KW   neurodegenerative disorder; hyperglycaemia; impaired glucose tolerance;
KW   obesity; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             194. .1489
FT                   /*tag=  a
FT                   /product= "DP160"
```

```
XX
PN   WO2003099318-A2.
XX
PD   04-DEC-2003.
XX
PF   30-MAY-2003; 2003WO-EP005700.
XX
PR   29-MAY-2002; 2002EP-00011963.
PR   17-SEP-2002; 2002EP-00020829.
XX
PA   (DEVE-) DEVELOGEN ENTWICKLUNGSBIOLOGISCHE FORSCH.
XX
PI   Dohrmann C,  Austen M;
XX
DR   WPI; 2004-035045/03.
DR   P-PSDB; ADE43924.
XX
PT   Use of DP119, DP444, DP810, DP685, WE474, DP160, RA977 or RA770 nucleic
PT   acid molecules or polypeptides, or their effectors or modulators, for
PT   manufacturing an agent for the diagnosis, prevention or treatment of
PT   pancreatic disorders.
XX
PS   Claim 3; SEQ ID NO 33; 125pp; English.
XX
CC   The present invention describes a nucleic acid molecule (I) selected from
CC   DP119, DP444, DP810, DP685, WE474, DP160, RA977 and RA770, or its encoded
CC   polypeptide, or a fragment or variant of the nucleic acid or polypeptide,
CC   or an effector/modulator of the nucleic acid or polypeptide. (I) is used
CC   in manufacturing a pharmaceutical agent, in monitoring or controlling the
CC   function of a gene or gene product influenced or modified by a DP119,
CC   DP444, DP810, DP685, WE474, DP160, RA977 or RA770 polypeptide, or in
CC   identifying substances that interact with the polypeptide. (I) has
CC   antidiabetic, anorectic and neuroprotective activities, and can be used
CC   in gene therapy. (I) is useful for diagnostic or therapeutic
CC   applications, or for manufacturing an agent for diagnosis, monitoring,
CC   prevention or treatment of pancreatic disorders, such as diabetes and
CC   related disorders, as well as neurodegenerative disorders, and other
CC   diseases. It may also be used for detecting and/or verifying, for the
CC   treatment, alleviation and/or prevention, of a pancreatic dysfunction
CC   (e.g. diabetes, hyperglycaemia or impaired glucose tolerance), and
CC   related disorders including obesity, and neurodegenerative disorders, and
CC   others, in cells, cell masses, organs and/or subjects. It is also used
CC   for regenerating or for promoting the differentiation and/or function of
CC   beta-cells in vitro and/or in vivo. (I) may also be used in preparing a
CC   non-human animal which over- or under-expresses the DP119, DP444, DP810,
CC   DP685, WE474, DP160, RA977 or RA770 gene product. In addition, the
CC   nucleic acid molecule, polypeptide or an effector/modulator of the
CC   nucleic acid or polypeptide, is used in monitoring or controlling the
CC   function of a gene or gene product influenced or modified by a DP119,
CC   DP444, DP810, DP685, WE474, DP160, RA977 or RA770 polypeptide, or in
CC   identifying substances that interact with the polypeptide. The present
CC   sequence is used in the exemplification of the present invention.
XX
SQ   Sequence 1767 BP; 375 A; 538 C; 444 G; 410 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0.676          Length:         1767
Score:                  94.50          Matches:        50
Percent Similarity:     37.1%          Conservative:   41
Best Local Similarity:  20.4%          Mismatches:     111
Query Match:            7.9%           Indels:         43
DB:                     12             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x ADE43923 (1-1767)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              ||||||         :::         :::      |||   ||||||::::::   |||
Db        713 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 772

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||         |||:::         :::         |||
Db        773 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 820

Qy         51 IleThrGlyHisGluGluGlyTyrPhe---------------------------------- 59
                 |||::::::      :::|||
Db        821 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 877

Qy         60 ----------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                    |||:::         ::::::|||   ||||||
Db        878 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 937

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
              :::         |||         :::      :::      |||   :::   :::
Db        938 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 997

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                      ||||||   :::   |||:::      :::|||||:::
Db        998 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 1057

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              :::      |||         |||      |||         :::|||:::   :::
Db       1058 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 1117

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
              |||         :::   |||   :::      |||      ||||||   |||
Db       1118 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 1177

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                 |||                                  |||:::         :::
Db       1178 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 1237

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
              :::   |||         |||:::   :::             ||||||
Db       1238 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 1285

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                    :::      :::      |||:::         |||   |||::::::
Db       1286 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 1339

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                 |||                  :::      :::||||||||||||      |||:::
Db       1340 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 1390

Qy        222 LeuAspCysGlyArg 226
              :::         |||
Db       1391 ATTGGACCCAACAGG 1405

RESULT 60
ADQ97811
ID   ADQ97811 standard; DNA; 1769 BP.
XX
AC   ADQ97811;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Human cancer associated sequence HR11-001, SEQ ID 788.
XX
KW   Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004060304-A2.
XX
PD   22-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US041389.
XX
PR   27-DEC-2002; 2002US-00330773.
XX
PA   (SAGR-) SAGRES DISCOVERY INC.
XX
PI   Morris DW,  Malandro MS;
XX
DR   WPI; 2004-543781/52.
XX
PT   New isolated cancer associated nucleic acids comprising at least 10
```

```
PT   contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT   cancers such as leukemia and lymphoma.
XX
PS   Claim 1; SEQ ID NO 788; 199pp; English.
XX
CC   The present invention relates to cancer associated sequences (ADQ97025-
CC   ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC   treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic formate directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1769 BP; 376 A; 536 C; 444 G; 413 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.677          Length:         1769
Score:                    94.50          Matches:        50
Percent Similarity:       37.1%          Conservative:   41
Best Local Similarity:    20.4%          Mismatches:     111
Query Match:              7.9%           Indels:         43
DB:                       12             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x ADQ97811 (1-1769)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              ||||||      :::         :::          |||   ||||||::::::    |||
Db        713 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 772

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||          |||:::          :::          |||
Db        773 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 820

Qy         51 IleThrGlyHisGluGluGlyTyrPhe---------------------------------- 59
                    |||::::::      :::|||
Db        821 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 877

Qy         60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                  |||:::          ::::::|||   ||||||
Db        878 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 937

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
              :::         |||          :::      :::       |||   :::   :::
Db        938 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 997

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                      ||||||   :::   |||:::      :::|||||||:::
Db        998 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 1057

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              :::   |||         |||         |||          :::|||:::   :::
Db       1058 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 1117

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
              |||         :::   |||   :::      |||      ||||||   |||
Db       1118 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 1177

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                 |||                                  |||:::          :::
Db       1178 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 1237

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
              :::   |||         |||:::   :::          ||||||
Db       1238 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 1285

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                    :::       :::       |||:::          |||   |||::::::
Db       1286 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 1339

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                 |||          :::       :::||||||||||||      |||:::
Db       1340 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 1390

Qy        222 LeuAspCysGlyArg 226
```

```
                  :::          |||
Db       1391 ATTGGACCCAACAGG 1405

RESULT 61
ADZ49510
ID   ADZ49510 standard; DNA; 1971 BP.
XX
AC   ADZ49510;
XX
DT   30-JUN-2005  (first entry)
XX
DE   Insulin signaling pathway related gene, SEQ ID 839.
XX
KW   Antidiabetic; Gene Therapy; Non-insulin dependent diabetes;
KW   insulin resistance; gene; ds.
XX
OS   Homo sapiens.
XX
PN   US2005085436-A1.
XX
PD   21-APR-2005.
XX
PF   08-JUL-2004; 2004US-00887553.
XX
PR   08-JUL-2003; 2003US-0485883P.
XX
PA   (LIHH/) LI H.
PA   (MAJJ/) MA J.
XX
PI   Li H,  Ma J;
XX
DR   WPI; 2005-305194/31.
XX
PT   Treating, preventing or ameliorating pathological conditions associated
PT   with dysregulation of the insulin signaling pathway (ISP) comprises
PT   administering to a subject an amount of a modulator of any of the
PT   proteins regulated by ISP.
XX
PS   Disclosure; SEQ ID NO 839; 70pp; English.
XX
CC   The present invention relates to a method for treating, preventing or
CC   ameliorating pathological conditions associated with dysregulation of the
CC   insulin signaling pathway (ISP). The method comprises administering to a
CC   subject a modulator for ISP-regulated proteins or a pharmaceutical
CC   composition comprising the described modulator. The method is useful for
CC   treating, preventing or ameliorating pathological conditions associated
CC   with dysregulation of the ISP such as Type II diabetes or Type A syndrome
CC   of insulin resistance. The present sequence is a human homolog of a
CC   Drosophila gene regulated by ISP. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20050085436.
XX
SQ   Sequence 1971 BP; 455 A; 574 C; 481 G; 461 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0.792          Length:         1971
Score:                    94.50          Matches:        50
Percent Similarity:       37.1%          Conservative:   41
Best Local Similarity:    20.4%          Mismatches:     111
Query Match:              7.9%           Indels:         43
DB:                       14             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x ADZ49510 (1-1971)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              ||||||      :::         :::          |||   ||||||::::::    |||
Db        713 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 772

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||          |||:::          :::          |||
Db        773 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 820
```

```
Qy        51 IleThrGlyHisGluGluGlyTyrPhe------------------------------ 59
             |||:::::::      :::|||
Db       821 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 877

Qy        60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                               |||:::         :::::::|||   ||||||
Db       878 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 937

Qy        73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
             :::        |||          :::       :::       |||   :::   :::
Db       938 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 997

Qy        92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                      ||||||   :::   |||:::       :::||||||:::
Db       998 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 1057

Qy       109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
             :::        |||          |||       |||         :::|||:::   :::
Db      1058 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 1117

Qy       129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
             |||          :::   |||   :::       |||       ||||||   |||
Db      1118 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 1177

Qy       147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                 |||                                  |||:::         :::
Db      1178 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 1237

Qy       165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
             :::   |||         |||:::   :::          ||||||
Db      1238 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 1285

Qy       182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                  :::       :::       |||:::          |||   |||::::::
Db      1286 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 1339

Qy       202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                 |||          :::      :::|||||||||||||       |||:::
Db      1340 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 1390

Qy       222 LeuAspCysGlyArg 226
             :::       |||
Db      1391 ATTGGACCCAACAGG 1405
```

RESULT 62
ADQ97810
ID ADQ97810 standard; DNA; 49355 BP.
XX
AC ADQ97810;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human cancer associated sequence HD11-001, SEQ ID 787.
XX
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
OS Homo sapiens.
XX
PN WO2004060304-A2.
XX
PD 22-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US041389.
XX
PR 27-DEC-2002; 2002US-00330773.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2004-543781/52.
XX
PT New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX
PS Claim 1; SEQ ID NO 787; 199pp; English.
XX
CC The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic formate directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 49355 BP; 12474 A; 10173 C; 10092 G; 13597 T; 0 U; 3019 Other;

Alignment Scores:

Pred. No.:	83	Length:	49355
Score:	94.50	Matches:	50
Percent Similarity:	37.1%	Conservative:	41
Best Local Similarity:	20.4%	Mismatches:	111
Query Match:	7.9%	Indels:	43
DB:	12	Gaps:	9

US-10-757-745-2_COPY_121_347 (1-227) x ADQ97810 (1-49355)

```
Qy        11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
             ||||||         :::       :::      |||   ||||||::::::   |||
Db     38299 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 38358

Qy        31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
             |||||||||         |||:::       :::         |||
Db     38359 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 38406

Qy        51 IleThrGlyHisGluGluGlyTyrPhe------------------------------ 59
                ||||::::::      :::|||
Db     38407 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 38463

Qy        60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                  |||:::         :::::::|||   ||||||
Db     38464 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 38523

Qy        73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
             :::        |||          :::       :::       |||   :::   :::
Db     38524 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 38583

Qy        92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                      ||||||   :::   |||:::       :::||||||:::
Db     38584 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 38643

Qy       109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
             :::        |||          |||       |||         :::|||:::   :::
Db     38644 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 38703

Qy       129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
             |||          :::   |||   :::       |||       ||||||   |||
Db     38704 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 38763

Qy       147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                 |||                                  |||:::         :::
Db     38764 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 38823

Qy       165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
             :::   |||         |||:::   :::          ||||||
Db     38824 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 38871

Qy       182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                  :::       :::       |||:::          |||   |||::::::
Db     38872 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 38925

Qy       202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                 |||          :::      :::|||||||||||||       |||:::
Db     38926 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 38976

Qy       222 LeuAspCysGlyArg 226
```

```
              :::          |||
Db      38977 ATTGGACCCAACAGG 38991

RESULT 63
AEB39175_05
Continuation (6 of 35) of AEB39175 from base 500001 (L. pneumophila DNA SEQ ID NO 3507.
WP Sequence split into 35 fragments  LOCUS AEB39175 Accession Aeb39175
WP      Fragment Name       Begin      End
WP      AEB39175_00             1    110000
WP      AEB39175_01        100001    210000
WP      AEB39175_02        200001    310000
WP      AEB39175_03        300001    410000
WP      AEB39175_04        400001    510000
WP      AEB39175_05        500001    610000
WP      AEB39175_06        600001    710000
WP      AEB39175_07        700001    810000
WP      AEB39175_08        800001    910000
WP      AEB39175_09        900001   1010000
WP      AEB39175_10       1000001   1110000
WP      AEB39175_11       1100001   1210000
WP      AEB39175_12       1200001   1310000
WP      AEB39175_13       1300001   1410000
WP      AEB39175_14       1400001   1510000
WP      AEB39175_15       1500001   1610000
WP      AEB39175_16       1600001   1710000
WP      AEB39175_17       1700001   1810000
WP      AEB39175_18       1800001   1910000
WP      AEB39175_19       1900001   2010000
WP      AEB39175_20       2000001   2110000
WP      AEB39175_21       2100001   2210000
WP      AEB39175_22       2200001   2310000
WP      AEB39175_23       2300001   2410000
WP      AEB39175_24       2400001   2510000
WP      AEB39175_25       2500001   2610000
WP      AEB39175_26       2600001   2710000
WP      AEB39175_27       2700001   2810000
WP      AEB39175_28       2800001   2910000
WP      AEB39175_29       2900001   3010000
WP      AEB39175_30       3000001   3110000
WP      AEB39175_31       3100001   3210000
WP      AEB39175_32       3200001   3310000
WP      AEB39175_33       3300001   3410000
WP      AEB39175_34       3400001   3503610

Alignment Scores:
Pred. No.:              264          Length:         110000
Score:                  94.50        Matches:        39
Percent Similarity:     43.1%        Conservative:   27
Best Local Similarity:  25.5%        Mismatches:     46
Query Match:            7.9%         Indels:         41
DB:                     14           Gaps:           7

US-10-757-745-2_COPY_121_347 (1-227) x AEB39175_05 (1-110000)

Qy         20 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal------------------ 33
              :::|||      :::||||||  :::|||||||||||||||
Db      83023 ATCGCTGGATTAAATCCGGACTTTGTTTTTTTACAAGAAGTGCAAGGTAAACATAGAAAA 83082

Qy         33 ------------------------------------------------------------ 33

Db      83083 CGCGAAAAAAGAATTGTCAGTTGGCCTGATTCCCCTCAATGTGAGTACATTGCTGAGAAT 83142

Qy         34 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 53
              |||   |||:::|||   |||   |||            |||           :::|||
Db      83143 ATATGGCCACATTACGTTTATGCAAAG------------AATGCAGTATATCAATCTGGC 83190

Qy         54 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 73
              |||     ::      |||||:::      |||:::          :::|||
Db      83191 CAC---------CACGGAAATGCTATTTTA---------AGCAAATACCCGTTCGAGAGT 83232

Qy         74 GlnGluIleIleProPheProSer---ThrLysMetMetArgAsnLeuLeuCysValHis 92
                 |||   |||       :::      :::      |||   :::|||
Db      83233 TTTGAAAGTATTAATCTTTCAAACATGAACAGAGCATCAAGAGGTATACTTCATACCCAA 83292

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              :::   :::       |||   :::   |||:::       ||||||         :::
Db      83293 TTAAAATTGGATAATAATATTATTCATCTTTTGTGTGTTCATTTAGGCCTATTTAAG--- 83349

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu---Ala 131
                 |||||||||      |||   |||   ::::::::::::::::::::
Db      83350 ------GCAGAGCGTACTGAGCAATGTAAGGCATTGATACAACGTATTAAAGATGTTGTG 83403

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
                 |||:::::::      :::|||   |||||||||   |||
Db      83404 CCTAAAAATGAGCCTATGATAATGGCAGGAGATTTTAAT 83442

RESULT 64
AEB42401_05
Continuation (6 of 34) of AEB42401 from base 500001 (L. pneumophila DNA SEQ ID NO 6733
WP Sequence split into 34 fragments  LOCUS AEB42401 Accession Aeb42401
WP      Fragment Name       Begin      End
WP      AEB42401_00             1    110000
WP      AEB42401_01        100001    210000
WP      AEB42401_02        200001    310000
WP      AEB42401_03        300001    410000
WP      AEB42401_04        400001    510000
WP      AEB42401_05        500001    610000
WP      AEB42401_06        600001    710000
WP      AEB42401_07        700001    810000
WP      AEB42401_08        800001    910000
WP      AEB42401_09        900001   1010000
WP      AEB42401_10       1000001   1110000
WP      AEB42401_11       1100001   1210000
WP      AEB42401_12       1200001   1310000
WP      AEB42401_13       1300001   1410000
WP      AEB42401_14       1400001   1510000
WP      AEB42401_15       1500001   1610000
WP      AEB42401_16       1600001   1710000
WP      AEB42401_17       1700001   1810000
WP      AEB42401_18       1800001   1910000
WP      AEB42401_19       1900001   2010000
WP      AEB42401_20       2000001   2110000
WP      AEB42401_21       2100001   2210000
WP      AEB42401_22       2200001   2310000
WP      AEB42401_23       2300001   2410000
WP      AEB42401_24       2400001   2510000
WP      AEB42401_25       2500001   2610000
WP      AEB42401_26       2600001   2710000
WP      AEB42401_27       2700001   2810000
WP      AEB42401_28       2800001   2910000
WP      AEB42401_29       2900001   3010000
WP      AEB42401_30       3000001   3110000
WP      AEB42401_31       3100001   3210000
WP      AEB42401_32       3200001   3310000
WP      AEB42401_33       3300001   3345687

Alignment Scores:
Pred. No.:              264          Length:         110000
Score:                  94.50        Matches:        39
Percent Similarity:     43.1%        Conservative:   27
Best Local Similarity:  25.5%        Mismatches:     46
Query Match:            7.9%         Indels:         41
DB:                     14           Gaps:           7

US-10-757-745-2_COPY_121_347 (1-227) x AEB42401_05 (1-110000)

Qy         20 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal------------------ 33
              :::|||      :::||||||  :::|||||||||||||||
Db      59089 ATCGCTGGATTAAATCCGGACTTTGTTTTTTTACAAGAAGTGCAAGGTAAACATAGAAAA 59148

Qy         33 ------------------------------------------------------------ 33

Db      59149 CGCGAAAAAAGAATTGTCAGTTGGCCTGATTCCCCTCAATGTGAATACATTGCTGAGAAT 59208

Qy         34 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 53
              |||   |||:::|||   |||   |||            |||           :::|||
```

```
Db      59209 ATATGGCCACATTACGTTTATGCAAAG-----------AATGCAGTATATCAATCTGGC 59256

Qy         54 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 73
              |||         :::      ||||||:::         |||:::         :::|||
Db      59257 CAC---------CACGGAAATGCTATTTTA---------AGCAAATACCCGTTCGAGAGT 59298

Qy         74 GlnGluIleIleProPheProSer---ThrLysMetMetArgAsnLeuLeuCysValHis 92
                 |||   |||         :::      :::      |||   :::|||
Db      59299 TTTGAAAGCATTAATCTTTCAAACATGAACAGAGCATCAAGAGGTATACTTCATACCCAA 59358

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              :::   :::      |||   :::   |||:::      ||||||          :::
Db      59359 TTAAAATTGGATAATAATATTATTCATCTTTTGTGTGTTCATTTAGGCCTATTCAAG--- 59415

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu---Ala 131
                 |||||||||      |||   |||   :::::::::::::::::::
Db      59416 ------GCAGAGCGTACTGAGCAATGTAAGGCATTGATACAACGTATTAAAGATGTTGTG 59469

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              |||::::::      :::|||   |||||||||   |||
Db      59470 CCTAAAAATGAGCCTTTGATAATGGCAGGAGACTTTAAT 59508
```

RESULT 65
AEB42735
ID AEB42735 standard; DNA; 243335 BP.
XX
AC AEB42735;
XX
DT 08-SEP-2005 (first entry)
XX
DE L. pneumophila DNA SEQ ID NO 7067.
XX
KW detection; infection; Antibacterial; Vaccine; ds; gene.
XX
OS Legionella pneumophila.
XX
PN WO2005049642-A2.
XX
PD 02-JUN-2005.
XX
PF 23-SEP-2004; 2004WO-IB003578.
XX
PR 21-NOV-2003; 2003FR-00013687.
XX
PA (INSP) INST PASTEUR.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (UYLY-) UNIV LYON 1 BERNARD CLAUDE.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;
PI Rusniok C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
PI Jarraud S;
XX
DR WPI; 2005-388305/40.
XX
PT New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX
PS Example 9; SEQ ID NO 7067; 660pp; English.
XX
CC The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain. (I), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (II). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present squence represents a L. pneumophila DNA.
XX
SQ Sequence 243335 BP; 73649 A; 42176 C; 51283 G; 76227 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	831	Length:	243335
Score:	94.50	Matches:	39
Percent Similarity:	43.1%	Conservative:	27
Best Local Similarity:	25.5%	Mismatches:	46
Query Match:	7.9%	Indels:	41
DB:	14	Gaps:	7

US-10-757-745-2_COPY_121_347 (1-227) x AEB42735 (1-243335)

```
Qy         20 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal---------------- 33
              :::|||      :::||||||   :::||||||||||||||||
Db     147024 ATCGCTGGATTAAATCCGGACTTTGTTTTTTTACAAGAAGTGCAAGGTAAACATAGAAAA 147083

Qy         33 ---------------------------------------------------------- 33

Db     147084 CGCGAAAAAAGAATTGTCAGTTGGCCTGATTCCCCTCAATGTGAGTACATTGCTGAGAAT 147143

Qy         34 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 53
              |||  |||:::|||  |||  |||          |||          :::|||
Db     147144 ATATGGCCACATTACGTTTATGCAAAG-----------AATGCAGTATATCAATCTGGC 147191

Qy         54 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 73
              |||         :::      ||||||:::         |||:::         :::|||
Db     147192 CAC---------CACGGAAATGCTATTTTA---------AGCAAATACCCGTTCGAGAGT 147233

Qy         74 GlnGluIleIleProPheProSer---ThrLysMetMetArgAsnLeuLeuCysValHis 92
                 |||   |||         :::      :::      |||   :::|||
Db     147234 TTTGAAAGTATTAATCTTTCAAACATGAACAGAGCATCAAGAGGTATACTTCATACCCAA 147293

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              :::   :::      |||   :::   |||:::      ||||||          :::
Db     147294 TTAAAATTGGATAATAATATTATTCATCTTTTGTGTGTTCATTTAGGCCTATTCAAG--- 147350

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu---Ala 131
                 |||||||||      |||   |||   :::::::::::::::::::
Db     147351 ------GCAGAGCGTACTGAGCAATGTAAGGCATTGATACAGCGTATTAAAGATGTTGTG 147404

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              |||::::::      :::|||   |||||||||   |||
Db     147405 CCTAAAAATGAGCCTTTGATAATGGCAGGAGATTTTAAT 147443
```

RESULT 66
AEB35721
ID AEB35721 standard; DNA; 295644 BP.
XX
AC AEB35721;
XX
DT 08-SEP-2005 (first entry)
XX
DE L. pneumophila DNA SEQ ID NO 53.
XX
KW detection; infection; Antibacterial; Vaccine; ds; gene.
XX
OS Legionella pneumophila.
XX
PN WO2005049642-A2.
XX
PD 02-JUN-2005.
XX
PF 23-SEP-2004; 2004WO-IB003578.
XX
PR 21-NOV-2003; 2003FR-00013687.
XX
PA (INSP) INST PASTEUR.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (UYLY-) UNIV LYON 1 BERNARD CLAUDE.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;

```
PI   Rusniok C,  Bouchier C,  Zidane N,  Magnier A,  Kunst F,  Vandenesch F;
PI   Jarraud S;
XX
DR   WPI; 2005-388305/40.
XX
PT   New genome of Legionella pneumophila Paris strain and derived
PT   polypeptides, useful for detection or identification of the strain and
PT   for treatment and prevention of infections.
XX
PS   Claim 1; SEQ ID NO 53; 660pp; English.
XX
CC   The invention relates to an isolated or purified nucleotide sequences (I)
CC   from Legionella pneumophila Paris strain. (I), and their related
CC   sequences or fragments, are useful as primers and probes for detection
CC   and amplification, including differentiation between the Paris and
CC   Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC   (hybrid) polypeptides (II). (II) are also useful for preparation of
CC   specific antibodies (Ab), also used for detection/identification of
CC   Legionella, and some (I), specifically those involved in synthesis of
CC   surface proteins, are targets for identification of inhibitors. (II), or
CC   vectors that contain (I), are useful as vaccines and immunogenic
CC   compositions, for treatment and prevention of infections by L.
CC   pneumophila. The present squence represents a L. pneumophila DNA.
XX
SQ   Sequence 295644 BP; 87609 A; 52716 C; 63040 G; 92276 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:                1.1e+03        Length:          295644
Score:                    94.50          Matches:         39
Percent Similarity:       43.1%          Conservative:    27
Best Local Similarity:    25.5%          Mismatches:      46
Query Match:              7.9%           Indels:          41
DB:                       14             Gaps:            7

US-10-757-745-2_COPY_121_347 (1-227) x AEB35721 (1-295644)

Qy         20 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal------------------ 33
              :::|||      :::||||||   :::||||||||||||||||
Db     112690 ATCGCTGGATTAAATCCGGACTTTGTTTTTTTACAAGAAGTGCAAGGTAAACATAGAAAA 112749

Qy         33 ------------------------------------------------------------ 33

Db     112750 CGCGAAAAAAGAATTGTCAGTTGGCCTGATTCCCCTCAATGTGAGTACATTGCTGAGAAT 112809

Qy         34 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 53
              |||   |||:::|||   |||   |||             |||            :::|||
Db     112810 ATATGGCCACATTACGTTTATGCAAAG------------AATGCAGTATATCAATCTGGC 112857

Qy         54 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 73
              |||         :::      ||||||:::         |||:::         :::|||
Db     112858 CAC---------CACGGAAATGCTATTTTA---------AGCAAATACCCGTTCGAGAGT 112899

Qy         74 GlnGluIleIleProPheProSer---ThrLysMetMetArgAsnLeuLeuCysValHis 92
                |||   |||         :::     :::      |||   :::|||
Db     112900 TTTGAAAGTATTAATCTTTCAAACATGAACAGAGCATCAAGAGGTATACTTCATACCCAA 112959

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              :::   :::      |||   :::   |||:::      ||||||         :::
Db     112960 TTAAAATTGGATAATAATATTATTCATCTTTTGTGTGTTCATTTAGGCCTATTTAAG--- 113016

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu---Ala 131
                 |||||||||      |||   |||   :::::::::::::::::::::
Db     113017 ------GCAGAGCGTACTGAGCAATGTAAGGCATTGATACAACGTATTAAAGATGTTGTG 113070

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              |||::::::      :::|||   |||||||||   |||
Db     113071 CCTAAAAATGAGCCTATGATAATGGCAGGAGATTTTAAT 113109

RESULT 67
AEB39173
ID   AEB39173 standard; DNA; 298667 BP.
XX
AC   AEB39173;
XX
DT   08-SEP-2005  (first entry)
XX
DE   L. pneumophila DNA SEQ ID NO 3505.
XX
KW   detection; infection; Antibacterial; Vaccine; ds; gene.
XX
OS   Legionella pneumophila.
XX
PN   WO2005049642-A2.
XX
PD   02-JUN-2005.
XX
PF   23-SEP-2004; 2004WO-IB003578.
XX
PR   21-NOV-2003; 2003FR-00013687.
XX
PA   (INSP ) INST PASTEUR.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA   (UYLY-) UNIV LYON 1 BERNARD CLAUDE.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Buchrieser C,  Tichit M,  Etienne J,  Ma L,  Cazalet C,  Glaser P;
PI   Rusniok C,  Bouchier C,  Zidane N,  Magnier A,  Kunst F,  Vandenesch F;
PI   Jarraud S;
XX
DR   WPI; 2005-388305/40.
XX
PT   New genome of Legionella pneumophila Paris strain and derived
PT   polypeptides, useful for detection or identification of the strain and
PT   for treatment and prevention of infections.
XX
PS   Example 9; SEQ ID NO 3505; 660pp; English.
XX
CC   The invention relates to an isolated or purified nucleotide sequences (I)
CC   from Legionella pneumophila Paris strain. (I), and their related
CC   sequences or fragments, are useful as primers and probes for detection
CC   and amplification, including differentiation between the Paris and
CC   Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC   (hybrid) polypeptides (II). (II) are also useful for preparation of
CC   specific antibodies (Ab), also used for detection/identification of
CC   Legionella, and some (I), specifically those involved in synthesis of
CC   surface proteins, are targets for identification of inhibitors. (II), or
CC   vectors that contain (I), are useful as vaccines and immunogenic
CC   compositions, for treatment and prevention of infections by L.
CC   pneumophila. The present squence represents a L. pneumophila DNA.
XX
SQ   Sequence 298667 BP; 89051 A; 52428 C; 62548 G; 94640 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.12e+03       Length:          298667
Score:                    94.50          Matches:         39
Percent Similarity:       43.1%          Conservative:    27
Best Local Similarity:    25.5%          Mismatches:      46
Query Match:              7.9%           Indels:          41
DB:                       14             Gaps:            7

US-10-757-745-2_COPY_121_347 (1-227) x AEB39173 (1-298667)

Qy         20 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal------------------ 33
              :::|||      :::||||||   :::||||||||||||||||
Db     117371 ATCGCTGGATTAAATCCGGACTTTGTTTTTTTACAAGAAGTGCAAGGTAAACATAGAAAA 117430

Qy         33 ------------------------------------------------------------ 33

Db     117431 CGCGAAAAAAGAATTGTCAGTTGGCCTGATTCCCCTCAATGTGAGTACATTGCTGAGAAT 117490

Qy         34 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 53
              |||   |||:::|||   |||   |||             |||            :::|||
Db     117491 ATATGGCCACATTACGTTTATGCAAAG------------AATGCAGTATATCAATCTGGC 117538

Qy         54 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 73
              |||         :::      ||||||:::         |||:::         :::|||
```

```
Db     117539 CAC---------CACGGAAATGCTATTTTA---------AGCAAATACCCGTTCGAGAGT 117580

Qy         74 GlnGluIleIleProPheProSer---ThrLysMetMetArgAsnLeuLeuCysValHis 92

Db     117581 TTTGAAAGTATTAATCTTTCAAACATGAACAGAGCATCAAGAGGTATACTTCATACCCAA 117640

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112

Db     117641 TTAAAATTGGATAATAATATTATTCATCTTTTGTGTGTTCATTTAGGCCTATTCAAG--- 117697

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu---Ala 131

Db     117698 ------GCAGAGCGTACTGAGCAATGTAAGGCATTGATACAGCGTATTAAAGATGTTGTG 117751

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144

Db     117752 CCTAAAAATGAGCCTTTGATAATGGCAGGAGATTTTAAT 117790
```

RESULT 68
ABL65231
ID ABL65231 standard; DNA; 3285 BP.
XX
AC ABL65231;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3568.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 3568; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 3285 BP; 997 A; 574 C; 690 G; 1024 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.91	Length:	3285
Score:	94.00	Matches:	63
Percent Similarity:	37.7%	Conservative:	37
Best Local Similarity:	23.8%	Mismatches:	96
Query Match:	7.9%	Indels:	70
DB:	6	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x ABL65231 (1-3285)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32

Db        703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy         33 ValIleProProTyr----------------------------------------TyrSer 39

Db        763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59

Db        823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75

Db        868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy         76 ----------------------------------------------IleIleProPhePr 80

Db        928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987
```

```
Qy          80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
               |:::|||      :::      ||||||         ::::::|||:::   :::   ::
Db         988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy         100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
               :||||||:::      ||||||:::||||||                |||         ||
Db        1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy         120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
               |       :::|||   :::|||   :::|||                ::::::   ||
Db        1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy         140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
               ||||                ||||||:::||||||      |||              ||
Db        1155 TGGG----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy         160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
               |   ::::::   :::   |||           |||   :::|||      |||
Db        1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy         178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
                :::|||         :::
Db        1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy         187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||         :::::||||||   :::           |||
Db        1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy         204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                  |||   ::::::   |||   |||:::                |||
Db        1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy         221 sLeuAspCysGly 225
               |||   ||||||
Db        1431 TCTAATTTGTGGT 1443
```

```
RESULT 69
ABN97250
ID   ABN97250 standard; DNA; 3285 BP.
XX
AC   ABN97250;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #3748 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030589.
XX
PR   02-OCT-2000; 2000US-0237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer, hepatocellular
PT   carcinoma or metastatic liver tumor in a patient, involves detecting the
PT   level of expression of two or more genes in a liver tissue sample.
XX
PS   Claim 1; SEQ ID NO 3748; 298pp; English.
XX
```

```
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3285 BP; 997 A; 574 C; 690 G; 1024 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:              1.91          Length:         3285
Score:                  94.00         Matches:        63
Percent Similarity:     37.7%         Conservative:   37
Best Local Similarity:  23.8%         Mismatches:     96
Query Match:            7.9%          Indels:         70
DB:                     6             Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x ABN97250 (1-3285)

Qy          13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
               |||:::|||      |||   :::|||                :::||||||
Db         703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy          33 ValIleProProTyr---------------------------------------TyrSer 39
               :::      |||                                          |||
Db         763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy          40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
               |||:::      |||               :::||||||      ||||||        ||
Db         823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy          59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
               ||||      :::  :::|||||||||          :::|||   |||
Db         868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy          76 ------------------------------------------------IleIleProPhePr 80
                                                               |||:::      ||
Db         928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy          80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
               |:::|||      :::      ||||||         ::::::|||:::   :::   ::
Db         988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy         100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
               :||||||:::      ||||||:::||||||                |||         ||
Db        1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy         120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
               |       :::|||   :::|||   :::|||                ::::::   ||
Db        1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy         140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
               ||||                ||||||:::||||||      |||              ||
Db        1155 TGGG----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy         160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
               |   ::::::   :::   |||           |||   :::|||      |||
Db        1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy         178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
                :::|||         :::
Db        1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy         187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
```

```
Db       1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221

Db       1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy        221 sLeuAspCysGly 225

Db       1431 TCTAATTTGTGGT 1443
```

RESULT 70
ADE85050
ID ADE85050 standard; DNA; 3285 BP.
XX
AC ADE85050;
XX
DT 29-JAN-2004 (first entry)
XX
DE Farnesyl transferase inhibitor modulated leukemia associated gene #269.
XX
KW ss; cytostatic; farnesyl transferase inhibitor; gene expression;
KW quinolinone; leukemia; cancer.
XX
OS Homo sapiens.
XX
PN WO2003038129-A2.
XX
PD 08-MAY-2003.
XX
PF 30-OCT-2002; 2002WO-US034784.
XX
PR 30-OCT-2001; 2001US-0338997P.
PR 30-OCT-2001; 2001US-0340081P.
PR 30-OCT-2001; 2001US-0340938P.
PR 30-OCT-2001; 2001US-0341012P.
XX
PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI Raponi M;
XX
DR WPI; 2003-513497/48.
XX
PT Determining whether a patient will respond to treatment with a farnesyl
PT transferase inhibitor, by analyzing the expression of gene that is
PT differentially modulated in the presence of the inhibitor.
XX
PS Disclosure; SEQ ID NO 269; 346pp; English.
XX
CC The invention relates to a method of determining whether a patient will
CC respond to treatment with a farnesyl transferase inhibitor (FTI), by
CC analyzing the expression of gene that is differentially modulated in the
CC presence of an FTI. The method is useful for determining whether a
CC patient will respond to treatment with a FTI such as (B)-6-[amino(4-
CC chlorophenyl)(1-methyl-1H-imidazol-5-yl)methyl]-4-(3-chlorophenyl)-1-
CC methyl-2-(1H)quinolinone, monitoring the therapy of a patient, treating a
CC patient with leukemia with FTI if the analysis indicates that the patient
CC will respond. This sequence corresponds to a gene whose expression may be
CC modulated in the presence of FTI.
XX
SQ Sequence 3285 BP; 997 A; 574 C; 690 G; 1024 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.91	Length:	3285
Score:	94.00	Matches:	63
Percent Similarity:	37.7%	Conservative:	37
Best Local Similarity:	23.8%	Mismatches:	96
Query Match:	7.9%	Indels:	70
DB:	10	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x ADE85050 (1-3285)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32

Db        703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy         33 ValIleProProTyr---------------------------------------TyrSer 39

Db        763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59

Db        823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75

Db        868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy         76 ---------------------------------------------IleIleProPhePr 80

Db        928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100

Db        988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120

Db       1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140

Db       1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160

Db       1155 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178

Db       1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186

Db       1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204

Db       1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221

Db       1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy        221 sLeuAspCysGly 225

Db       1431 TCTAATTTGTGGT 1443
```

RESULT 71
ABL61829
ID ABL61829 standard; DNA; 3397 BP.
XX
AC ABL61829;
XX
DT 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:166.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.

XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 166; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 3397 BP; 1052 A; 581 C; 700 G; 1064 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2	Length:	3397
Score:	94.00	Matches:	63
Percent Similarity:	37.7%	Conservative:	37
Best Local Similarity:	23.8%	Mismatches:	96
Query Match:	7.9%	Indels:	70
DB:	6	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x ABL61829 (1-3397)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||     |||  :::|||                 :::|||||||
Db        679 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 738

Qy         33 ValIleProProTyr----------------------------------------TyrSer 39
              :::      |||                                           |||
Db        739 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 798

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::      |||                :::||||||      ||||||       ||
Db        799 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 843

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
              ||||      :::  :::||||||||||          :::|||   |||
Db        844 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 903

Qy         76 ------------------------------------------------IleIleProPhePr 80
                                                              |||:::      ||
Db        904 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 963

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::      ||||||         :::::::|||:::   :::   ::
Db        964 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1017

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::       ||||||:::||||||           |||      ||
Db       1018 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1076

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |         :::|||   :::|||   :::|||                 ::::::   ||
Db       1077 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1130

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||                 ||||||:::||||||     |||          ||
Db       1131 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1172

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |  ::::::   :::   |||          |||   :::|||      |||
Db       1173 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1232

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
               :::|||          :::
Db       1233 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1292

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||         ::::::||||||   :::          |||
Db       1293 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1346

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                 |||   ::::::   |||   |||:::                |||
Db       1347 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1406

Qy        221 sLeuAspCysGly 225
               |||   ||||||
Db       1407 TCTAATTTGTGGT 1419
```

```
RESULT 72
ABN95906
ID   ABN95906 standard; DNA; 3397 BP.
XX
AC   ABN95906;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #2404 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US030589.
XX
PR   02-OCT-2000; 2000US-0237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer, hepatocellular
PT   carcinoma or metastatic liver tumor in a patient, involves detecting the
PT   level of expression of two or more genes in a liver tissue sample.
XX
PS   Claim 1; SEQ ID NO 2404; 298pp; English.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3397 BP; 1052 A; 581 C; 700 G; 1064 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2           Length:         3397
Score:                  94.00       Matches:        63
Percent Similarity:     37.7%       Conservative:   37
Best Local Similarity:  23.8%       Mismatches:     96
Query Match:            7.9%        Indels:         70
DB:                     6           Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x ABN95906 (1-3397)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||      |||   :::|||                  :::||||||
Db        679 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 738

Qy         33 ValIleProProTyr-------------------------------------TyrSer 39
              :::      |||                                        |||
Db        739 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 798

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::      |||               :::||||||      ||||||       ||
Db        799 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 843

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
              ||||      :::   :::|||||||||         :::|||   |||
Db        844 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 903

Qy         76 ----------------------------------------------IleIleProPhePr 80
                                                            |||:::      ||
Db        904 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 963

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::      ||||||         ::::::|||:::   :::   ::
Db        964 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1017

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::||||||                |||        ||
Db       1018 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1076

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |         :::|||   :::|||   :::|||                 ::::::   ||
Db       1077 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1130

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||                ||||||:::||||||     |||          ||
Db       1131 TGGG----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1172

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |   ::::::   :::   |||             |||   :::|||      |||
Db       1173 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1232

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
              :::|||          :::
Db       1233 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1292

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||         ::::::||||||   :::           |||
Db       1293 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1346

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                    |||   ::::::   |||   |||:::               |||
Db       1347 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1406

Qy        221 sLeuAspCysGly 225
              |||   ||||||
Db       1407 TCTAATTTGTGGT 1419

RESULT 73
ADE85000
ID   ADE85000 standard; DNA; 3397 BP.
XX
AC   ADE85000;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Farnesyl transferase inhibitor modulated leukemia associated gene #219.
XX
KW   ss; cytostatic; farnesyl transferase inhibitor; gene expression;
KW   quinolinone; leukemia; cancer.
XX
OS   Homo sapiens.
XX
PN   WO2003038129-A2.
XX
PD   08-MAY-2003.
XX
PF   30-OCT-2002; 2002WO-US034784.
XX
PR   30-OCT-2001; 2001US-0338997P.
PR   30-OCT-2001; 2001US-0340081P.
PR   30-OCT-2001; 2001US-0340938P.
PR   30-OCT-2001; 2001US-0341012P.
```

```
XX
PA   (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI   Raponi M;
XX
DR   WPI; 2003-513497/48.
XX
PT   Determining whether a patient will respond to treatment with a farnesyl
PT   transferase inhibitor, by analyzing the expression of gene that is
PT   differentially modulated in the presence of the inhibitor.
XX
PS   Disclosure; SEQ ID NO 219; 346pp; English.
XX
CC   The invention relates to a method of determining whether a patient will
CC   respond to treatment with a farnesyl transferase inhibitor (FTI), by
CC   analyzing the expression of gene that is differentially modulated in the
CC   presence of an FTI. The method is useful for determining whether a
CC   patient will respond to treatment with a FTI such as (B)-6-[amino(4-
CC   chlorophenyl)(1-methyl-1H-imidazol-5-yl)methyl]-4-(3-chlorophenyl)-1-
CC   methyl-2-(1H)quinolinone, monitoring the therapy of a patient, treating a
CC   patient with leukemia with FTI if the analysis indicates that the patient
CC   will respond. This sequence corresponds to a gene whose expression may be
CC   modulated in the presence of FTI.
XX
SQ   Sequence 3397 BP; 1052 A; 581 C; 700 G; 1064 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2                 Length:          3397
Score:                    94.00             Matches:         63
Percent Similarity:       37.7%             Conservative:    37
Best Local Similarity:    23.8%             Mismatches:      96
Query Match:              7.9%              Indels:          70
DB:                       10                Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x ADE85000 (1-3397)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||     |||   :::|||           :::||||||
Db        679 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 738

Qy         33 ValIleProProTyr---------------------------------------TyrSer 39
              :::     |||                                           |||
Db        739 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 798

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::     |||             :::||||||     ||||||      ||
Db        799 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 843

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
              ||||     :::   :::||||||||||         :::|||   |||
Db        844 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 903

Qy         76 ------------------------------------------IleIleProPhePr 80
                                                        |||:::      ||
Db        904 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 963

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||     :::   ||||||         ::::::|||:::  :::   ::
Db        964 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1017

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :|||||:::      ||||||:::||||||          |||             ||
Db       1018 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1076

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |         :::|||   :::|||   :::|||                :::::   ||
Db       1077 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1130

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||           ||||||:::||||||      |||             ||
Db       1131 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1172

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |   ::::::   :::   |||             |||   :::|||      |||
Db       1173 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1232

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
              :::|||            :::
Db       1233 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1292

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||           :::::::||||||   :::          |||
Db       1293 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1346

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                  |||   ::::::   |||   |||:::              |||
Db       1347 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1406

Qy        221 sLeuAspCysGly 225
                |||  |||||
Db       1407 TCTAATTTGTGGT 1419

RESULT 74
ADR24971
ID   ADR24971 standard; DNA; 3423 BP.
XX
AC   ADR24971;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Breast cancer prognosis marker #832.
XX
KW   ds; breast cancer; prognosis; gene expression; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004065545-A2.
XX
PD   05-AUG-2004.
XX
PF   15-JAN-2004; 2004WO-US001100.
XX
PR   15-JAN-2003; 2003US-00342887.
XX
PA   (ROSE-) ROSETTA INPHARMATICS LLC.
PA   (NECA-) NETHERLANDS CANCER INST.
XX
PI   Van't Veer LJ,  He Y;
XX
DR   WPI; 2004-593473/57.
XX
PT   Classifying a breast cancer patient according to prognosis comprises
PT   determining the similarity between the level of expression of each of
PT   five genes in a cell sample taken from patient, to control levels.
XX
PS   Disclosure; SEQ ID NO 832; 226pp; English.
XX
CC   The invention relates to a method of classifying a breast cancer patient
CC   according to prognosis by determining the similarity between the level of
CC   expression of each of five genes for which markers are listed in the
CC   specification, in a cell sample taken from the breast cancer patient, to
CC   control levels of expression for each respective five genes to obtain a
CC   patient similarity value. The methods are useful for classifying a breast
CC   cancer patient according to prognosis. Kits and computer program products
CC   are useful for data analysis using the diagnostic, prognostic and
CC   statistical methods of the invention. This sequence corresponds to a
CC   marker used in the method of the invention.
XX
SQ   Sequence 3423 BP; 1054 A; 591 C; 709 G; 1069 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.02              Length:          3423
Score:                    94.00             Matches:         63
Percent Similarity:       37.7%             Conservative:    37
Best Local Similarity:    23.8%             Mismatches:      96
Query Match:              7.9%              Indels:          70
```

```
DB:                    13            Gaps:          10

US-10-757-745-2_COPY_121_347 (1-227) x ADR24971 (1-3423)

Qy          13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
               |||:::|||      |||   :::|||                :::||||||
Db         703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy          33 ValIleProProTyr---------------------------------------TyrSer 39
               :::      |||                                          |||
Db         763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy          40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
               |||:::      |||                :::||||||      ||||||       ||
Db         823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy          59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
               ||||      :::   :::|||||||||          :::|||   |||
Db         868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy          76 ---------------------------------------------IleIleProPhePr 80
                                                           |||:::      ||
Db         928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy          80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
               |:::|||      :::      ||||||         ::::::|||:::  :::   ::
Db         988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy         100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
               :||||||:::     ||||||:::||||||           |||       ||
Db        1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy         120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
               |         :::|||   :::|||   :::|||              ::::::   ||
Db        1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy         140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
               ||||               ||||||:::||||||      |||          ||
Db        1155 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy         160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
               |   ::::::   :::   |||              |||   :::|||      |||
Db        1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy         178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
               :::|||           :::
Db        1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy         187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||        ::::::||||||  :::              |||
Db        1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy         204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                   |||   ::::::   |||   |||:::               |||
Db        1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy         221 sLeuAspCysGly 225
               |||   ||||||
Db        1431 TCTAATTTGTGGT 1443

RESULT 75
ADY19709
ID   ADY19709 standard; DNA; 3423 BP.
XX
AC   ADY19709;
XX
DT   05-MAY-2005  (first entry)
XX
DE   DNA encoding a PRO polypeptide, SEQ ID NO 5515.
XX
KW   Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW   Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW   Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
KW   Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW   Antiallergic; ds; gene; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2005016962-A2.
XX
PD   24-FEB-2005.
XX
PF   11-AUG-2004; 2004WO-US026249.
XX
PR   11-AUG-2003; 2003US-0493546P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Abbas A,  Clark H,  Ouyang W,  Williams MP,  Wood WI,  Wu TD;
XX
DR   WPI; 2005-182330/19.
XX
PT   New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT   treating an immune related disorder, e.g. systemic lupus erythematosus,
PT   rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
PS   Claim 1; SEQ ID NO 5515; 158pp; English.
XX
CC   The invention relates to an isolated nucleic acid encoding a PRO
CC   polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC   composition, and method are useful for diagnosing and treating an immune
CC   related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC   arthritis. The present sequence represents a PRO polypeptide.
XX
SQ   Sequence 3423 BP; 1054 A; 591 C; 709 G; 1069 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.02          Length:          3423
Score:                    94.00         Matches:         63
Percent Similarity:       37.7%         Conservative:    37
Best Local Similarity:    23.8%         Mismatches:      96
Query Match:              7.9%          Indels:          70
DB:                       14            Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x ADY19709 (1-3423)

Qy          13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
               |||:::|||      |||   :::|||                :::||||||
Db         703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy          33 ValIleProProTyr---------------------------------------TyrSer 39
               :::      |||                                          |||
Db         763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy          40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
               |||:::      |||                :::||||||      ||||||       ||
Db         823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy          59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
               ||||      :::   :::|||||||||          :::|||   |||
Db         868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy          76 ---------------------------------------------IleIleProPhePr 80
                                                           |||:::      ||
Db         928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy          80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
               |:::|||      :::      ||||||         ::::::|||:::  :::   ::
Db         988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy         100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
               :||||||:::     ||||||:::||||||           |||       ||
Db        1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy         120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
               |         :::|||   :::|||   :::|||              ::::::   ||
```

```
Db       1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160

Db       1155 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178

Db       1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186

Db       1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204

Db       1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221

Db       1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy        221 sLeuAspCysGly 225

Db       1431 TCTAATTTGTGGT 1443
```

Search completed: March 10, 2006, 15:33:27
Job time : 1238.39 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:40:54 ; Search time 7484.21 Seconds
(without alignments)
1724.091 Million cell updates/sec

Title: US-10-757-745-2_COPY_121_347
Perfect score: 1192
Sequence: 1 IDGLDLNNLSERARGVCSYL.....HIIPRSLDLLGLEKLDCGRF 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170327_4011/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_5954_@runat_09032006_170327_4011 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1192	100.0	1261	8	AF201687	AF201687 Homo sapi
2	1192	100.0	1898	6	BD157138	BD157138 Primer fo
3	1192	100.0	1898	6	AX878304	AX878304 Sequence
4	1192	100.0	1898	8	AK002168	AK002168 Homo sapi
5	1192	100.0	1920	6	BD205490	BD205490 CD40-Inte
6	1192	100.0	1920	6	AR594294	AR594294 Sequence
7	1192	100.0	1920	6	AX011599	AX011599 Sequence
8	1192	100.0	1921	8	HSA269473	AJ269473 Homo sapi
9	1192	100.0	1936	8	AF223469	AF223469 Homo sapi
10	1192	100.0	1948	6	AR339398	AR339398 Sequence
11	1192	100.0	1952	8	BC017553	BC017553 Homo sapi
12	1192	100.0	2151	8	HSA420495	AJ420495 Homo sapi
13	1192	100.0	2499	6	E23195	E23195 Topoisomera
c 14	1099	92.2	150344	14	AC152027	AC152027 Dasypus n
c 15	1090	91.4	1079	6	BD139846	BD139846 Compounds
c 16	1090	91.4	1079	6	AR202950	AR202950 Sequence
c 17	1090	91.4	1079	6	AR208051	AR208051 Sequence
c 18	1090	91.4	1079	6	AX429924	AX429924 Sequence
19	945	79.3	1312	6	BD205491	BD205491 CD40-Inte
20	945	79.3	1312	6	AR594295	AR594295 Sequence
21	945	79.3	1312	6	AX011601	AX011601 Sequence
22	945	79.3	1943	9	MMU251328	AJ251328 Mus muscu
23	883	74.1	1236	5	CR761829	CR761829 Xenopus t
24	742.5	62.3	858	6	CQ726600	CQ726600 Sequence
25	710	59.6	1451	5	BC083404	BC083404 Danio rer
26	707	59.3	1455	5	BC097117	BC097117 Danio rer
27	590	49.5	400	6	CQ663156	CQ663156 Sequence
28	568	47.7	752	6	BD150065	BD150065 Primer fo
29	568	47.7	752	6	AX870003	AX870003 Sequence
30	462	38.8	256608	9	AL589699	AL589699 Mouse DNA
31	431	36.2	20938	8	AY613922	AY613922 Homo sapi
32	431	36.2	102200	8	HS30M3	AL031775 Human DNA
33	431	36.2	167847	8	CR925830	CR925830 Human DNA
34	431	36.2	176819	8	CR942205	CR942205 Human DNA
c 35	380	31.9	245394	14	AC125756	AC125756 Rattus no
36	333	27.9	201	10	BV203236	BV203236 sqnm21148
37	298	25.0	176	6	AX351374	AX351374 Sequence
c 38	298	25.0	179	6	AX339803	AX339803 Sequence
39	232.5	19.5	198752	5	BX957279	BX957279 Zebrafish
40	231.5	19.4	200076	5	BX511258	BX511258 Zebrafish
41	214	18.0	483	6	BD076937	BD076937 5' EST of
c 42	206	17.3	110000	14	AP006486_2	Continuation (3 of
43	201	16.9	1326	15	BT000599	BT000599 Arabidops
44	201	16.9	1557	15	AY085643	AY085643 Arabidops
45	201	16.9	1582	15	AY128286	AY128286 Arabidops
46	182.5	15.3	444	6	CQ069444	CQ069444 Sequence
47	182.5	15.3	691	6	CQ078617	CQ078617 Sequence
c 48	177.5	14.9	33899	2	CEY63D3A	AL032652 Caenorhab
49	162.5	13.6	2036	2	AK113716	AK113716 Ciona int
c 50	153.5	12.9	110000	14	CR954207_1	Continuation (2 of
51	153	12.8	91566	15	F25C20	AC007296 Arabidops
c 52	137.5	11.5	238494	14	AC095717	AC095717 Rattus no
53	137	11.5	110000	15	AP008218_128	Continuation (129
c 54	137	11.5	156724	15	CNS08CB2	AL844879 Oryza sat
c 55	135	11.3	111678	15	AC149575	AC149575 Populus t
c 56	134	11.2	268	6	AR240556	AR240556 Sequence
c 57	130	10.9	103129	15	AP004494	AP004494 Lotus cor
58	126	10.6	86100	15	AC151708	AC151708 Medicago
c 59	126	10.6	133028	15	AC134967	AC134967 Medicago
60	120	10.1	110000	1	CR626927_15	Continuation (16 o
61	120	10.1	110000	1	AP006841_15	Continuation (16 o
62	119	10.0	110000	15	AP008216_130	Continuation (131
63	119	10.0	169270	15	AC083943	AC083943 Genomic s
64	119	10.0	324367	15	AE017091	AE017091 Oryza sat
65	116.5	9.8	304517	1	AE016789	AE016789 Pseudomon
c 66	116	9.7	182686	14	AC145862	AC145862 Pan trogl
c 67	116	9.7	187205	14	AC148939	AC148939 Pan trogl
68	116	9.7	237211	8	AC150749	AC150749 Pan trogl
c 69	116	9.7	347071	1	BX640415	BX640415 Bordetell
70	116	9.7	348706	1	BX640445	BX640445 Bordetell
c 71	114.5	9.6	411	10	HS133N16S	AL110544 H.sapiens
c 72	113	9.5	110000	1	AE016853_39	Continuation (40 o
c 73	113	9.5	225886	14	CR855361	CR855361 Danio rer
74	111.5	9.4	110000	15	CR382132_15	Continuation (16 o
75	111.5	9.4	110000	15	CR382132_16	Continuation (17 o

ALIGNMENTS

```
RESULT 1
AF201687
LOCUS       AF201687                1261 bp    mRNA    linear   PRI 16-MAY-2003
DEFINITION  Homo sapiens ETS1-associated protein 2 (EAP2) mRNA, complete cds.
ACCESSION   AF201687
VERSION     AF201687.1  GI:11493668
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1261)
  AUTHORS   Pei,H., Yordy,J.S., Leng,Q., Zhao,Q., Watson,D.K. and Li,R.
  TITLE     EAPII interacts with ETS1 and modulates its transcriptional
            function
  JOURNAL   Oncogene 22 (18), 2699-2709 (2003)
   PUBMED   12743594
REFERENCE   2  (bases 1 to 1261)
  AUTHORS   Li,R., Pei,H. and Papas,T.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1999) Center for Molecular and Structural
            Biology, Department of Medicine, and Hollings Cancer Center,
            Medical University of South Carolina, 86 Jonathan Lucas St.,
            Charleston, SC 29425, USA
FEATURES             Location/Qualifiers
     source          1. .1261
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22.1-p22.3"
                     /cell_line="Jurkat"
     gene            1. .1261
                     /gene="EAP2"
     CDS             26. .1114
                     /gene="EAP2"
                     /codon_start=1
                     /product="ETS1-associated protein 2"
                     /protein_id="AAG35600.1"
                     /db_xref="GI:11493669"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:              1.69e-114       Length:         1261
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AF201687 (1-1261)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066

RESULT 2
BD157138
LOCUS       BD157138                1898 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD157138
VERSION     BD157138.1  GI:27862896
KEYWORDS    JP 2002191363-A/11981.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1898)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 11981 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/11981
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                        Location/Qualifiers
            FT   CDS             (1). .(1086).
FEATURES             Location/Qualifiers
     source          1. .1898
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                2.71e-114      Length:          1898
Score:                    1192.00        Matches:         227
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BD157138 (1-1898)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTT 1041

RESULT 3
AX878304
LOCUS       AX878304               1898 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 13209 from Patent EP1074617.
ACCESSION   AX878304
VERSION     AX878304.1  GI:40033040
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 13209 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .1089
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE90095.1"
                     /db_xref="GI:40033041"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                2.71e-114      Length:          1898
Score:                    1192.00        Matches:         227
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x AX878304 (1-1898)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900
```

```
Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTT 1041
```

RESULT 4
AK002168

LOCUS	AK002168 1898 bp mRNA linear PRI 30-JAN-2004
DEFINITION	Homo sapiens cDNA FLJ11306 fis, clone PLACE1010031.
ACCESSION	AK002168
VERSION	AK002168.1 GI:7023882
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
TITLE	Complete sequencing and characterization of 21,243 full-length human cDNAs
JOURNAL	Nat. Genet. 36 (1), 40-45 (2004)
PUBMED	14702039
REFERENCE	2
AUTHORS	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 1898)
AUTHORS	Isogai,T. and Otsuki,T.
TITLE	Direct Submission
JOURNAL	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES Location/Qualifiers

source	1. .1898
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="PLACE1010031"
	/tissue_type="placenta"
	/clone_lib="PLACE1"
	/note="cloning vector: pME18SFL3"
CDS	1. .1089
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="BAA92119.1"
	/db_xref="GI:7023883"
	/translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL DLLGLEKLDCGRFPSDHWGLLCNLDIIL"

ORIGIN

Alignment Scores:

Pred. No.:	2.71e-114	Length:	1898
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x AK002168 (1-1898)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840
```

```
Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTT 1041

RESULT 5
BD205490
LOCUS       BD205490                1920 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  CD40-Interacting and TRAF-interacting protein.
ACCESSION   BD205490
VERSION     BD205490.1  GI:33015260
KEYWORDS    JP 2002512796-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1920)
  AUTHORS   Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting protein
  JOURNAL   Patent: JP 2002512796-A 1 08-MAY-2002;
            VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002512796-A/1
            PD   08-MAY-2002
            PF   28-APR-1999 JP 2000546003
            PR   29-APR-1998 EP   98201392.2
            PI   STEFAN MARIA CHRISTIAAN PYPE,
            PI   JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI   DANNY
            FRANCOIS EVELINE HUYLEBROECK
            PC   C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
              A61P37/02,
            PC   A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
            PC   G01N33/566,
            PC   C12N15/00,A61K37/02
            CC   CD40-Interacting and TRAF-interacting protein. FH   Key
                 Location/Qualifiers
            FT   CDS             (20). .(1108).
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.75e-114      Length:         1920
Score:                  1192.00        Matches:        227
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BD205490 (1-1920)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTT 1060

RESULT 6
AR594294
LOCUS       AR594294                1920 bp    DNA     linear   PAT 15-DEC-2004
DEFINITION  Sequence 1 from patent US 6812203.
ACCESSION   AR594294
VERSION     AR594294.1  GI:56643900
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1920)
  AUTHORS   Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting proteins
  JOURNAL   Patent: US 6812203-A 1 02-NOV-2004;
            Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
            Zwijnaarde;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              2.75e-114      Length:         1920
Score:                  1192.00        Matches:        227
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AR594294 (1-1920)
```

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTT 1060
```

```
RESULT 7
AX011599
LOCUS       AX011599               1920 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 1 from Patent WO9955859.
ACCESSION   AX011599
VERSION     AX011599.1  GI:9998123
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 1 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             20. .1108
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07555.1"
                     /db_xref="GI:9998124"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:              2.75e-114       Length:         1920
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AX011599 (1-1920)
```


```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTT 1060
```


RESULT 8

```
HSA269473
LOCUS       HSA269473               1921 bp    mRNA    linear   PRI 27-MAY-2000
DEFINITION  Homo sapiens mRNA for TRAF and TNF receptor associated protein
            (ttrap gene).
ACCESSION   AJ269473
VERSION     AJ269473.1  GI:8247253
KEYWORDS    TRAF and TNF receptor associated protein; ttrap gene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van
            Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,
            Huylebroeck,D. and Remacle,J.E.
  TITLE     TTRAP, a novel protein that associates with CD40, tumor necrosis
            factor (TNF) receptor-75 and TNF receptor-associated factors
            (TRAFs), and that inhibits nuclear factor-kappa B activation
  JOURNAL   J. Biol. Chem. 275 (24), 18586-18593 (2000)
   PUBMED   10764746
REFERENCE   2  (bases 1 to 1921)
  AUTHORS   Pype,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1999) Pype S., VIB07, Dept. Cell Growth,
            Differentiation and Development, Flanders Interuniversity Institute
            for Biotechnology, CELGEN, K.U.Leuven, Gasthuisberg Campus,
            Herestraat 49, B-3000 Leuven, BELGIUM
COMMENT     Related sequence: AL031775.
FEATURES             Location/Qualifiers
     source          1. .1921
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .1921
                     /gene="ttrap"
     5'UTR           <1. .21
                     /gene="ttrap"
     CDS             22. .1110
                     /gene="ttrap"
                     /function="putative role in signal transduction"
                     /codon_start=1
                     /product="TRAF and TNF receptor associated protein"
                     /protein_id="CAB92966.1"
                     /db_xref="GI:8247254"
                     /db_xref="GOA:O95551"
                     /db_xref="UniProt/TrEMBL:O95551"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
     3'UTR           1111. .1921
                     /gene="ttrap"
     polyA_signal    1900. .1905
                     /gene="ttrap"
ORIGIN

Alignment Scores:
Pred. No.:              2.75e-114       Length:         1921
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x HSA269473 (1-1921)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 441

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 501

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 561

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        562 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 621

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        622 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 681

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        682 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 741

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        742 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 801

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        802 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 861

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        862 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 921

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        922 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 981

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        982 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1041

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1042 AAACTGGACTGTGGTAGATTT 1062

RESULT 9
AF223469
LOCUS       AF223469                1936 bp    mRNA    linear   PRI 17-APR-2000
DEFINITION  Homo sapiens AD022 protein (AD022) mRNA, complete cds.
ACCESSION   AF223469
VERSION     AF223469.1  GI:7578788
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1936)
  AUTHORS   Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
  TITLE     A novel gene expressed in human adrenal gland
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1936)
  AUTHORS   Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) Chinese National Human Genome Center at
            Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
            Shanghai 201203, China
FEATURES             Location/Qualifiers
     source          1. .1936
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="adrenal gland"
```

```
     gene            1. .1936
                     /gene="AD022"
     CDS             17. .1105
                     /gene="AD022"
                     /codon_start=1
                     /product="AD022 protein"
                     /protein_id="AAF64144.1"
                     /db_xref="GI:7578789"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVRFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYPTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:              2.78e-114       Length:         1936
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AF223469 (1-1936)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 436

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 496

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 556

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 616

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 676

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 736

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 796

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 856

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 916

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 976

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1036

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
```

```
Db       1037 AAACTGGACTGTGGTAGATTT 1057

RESULT 10
AR339398
LOCUS       AR339398                1948 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 889 from patent US 6569662.
ACCESSION   AR339398
VERSION     AR339398.1  GI:33726255
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Tang,Y.T., Zhou,P. and Drmanac,R.T.
  TITLE     Nucleic acids and polypeptides
  JOURNAL   Patent: US 6569662-A 889 27-MAY-2003;
            Hyseq, Inc.; Sunnyvale, CA
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              2.8e-114        Length:         1948
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AR339398 (1-1948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998
```

```
Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079
```

RESULT 11
BC017553

```
LOCUS       BC017553                1952 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein, mRNA (cDNA
            clone MGC:9099 IMAGE:3920790), complete cds.
ACCESSION   BC017553
VERSION     BC017553.2  GI:34782842
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1952)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1952)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Sep 16, 2003 this sequence version replaced gi:17028464.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 15 Row: m Column: 16

            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 23510347.
FEATURES            Location/Qualifiers
     source         1. .1952
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="MGC:9099 IMAGE:3920790"
                    /tissue_type="Skin, melanotic melanoma."
                    /clone_lib="NIH_MGC_72"
                    /lab_host="DH10B"
                    /note="Vector: pCMV-SPORT6"
     gene           1. .1952
                    /gene="TTRAP"
                    /note="synonyms: AD022, EAP2, MGC9099, dJ30M3.3"
                    /db_xref="GeneID:51567"
                    /db_xref="MIM:605764"
     CDS            38. .1126
                    /gene="TTRAP"
                    /codon_start=1
                    /product="TRAF and TNF receptor-associated protein"
                    /protein_id="AAH17553.1"
                    /db_xref="GI:17028465"
                    /db_xref="GeneID:51567"
                    /db_xref="MIM:605764"
                    /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                    CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                    SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                    LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                    ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                    PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                    DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                2.8e-114        Length:          1952
Score:                    1192.00         Matches:         227
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BC017553 (1-1952)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 457

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 517

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 577

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 637

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 697

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 757

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 817

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 877

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 937

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 997

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1057

Qy        221 LysLeuAspCysGlyArgPhe 227
                 |||||||||||||||||||
Db       1058 AAACTGGACTGTGGTAGATTT 1078

RESULT 12
HSA420495
LOCUS       HSA420495               2151 bp    mRNA    linear   PRI 23-NOV-2001
DEFINITION  Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1761756.
ACCESSION   AJ420495
VERSION     AJ420495.1  GI:17066359
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Auffray,C., Ansorge,W., Ballabio,A., Estivill,X., Gibson,K.,
            Lehrach,H., Poustka,A. and Lundeberg,J.
  TITLE     The European IMAGE consortium for integrated Molecular analysis of
            human gene transcripts
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2151)
  AUTHORS   Persson,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2001) Persson A., Center for Molecular
            Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagsvagen
            30B, 106 91 Stockholm, SWEDEN
COMMENT     This clone is available royalty-free through IMAGE Consortium
            Distributors. IMPORTANT: This sequence represents the full insert
            of this IMAGE cDNA clone. No attempt has been made to verify
            whether this corresponds to the full-length of the original mRNA
            from which it was derived.
FEATURES             Location/Qualifiers
     source          1. .2151
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="EUROIMAGE 1761756"
                     /clone_lib="NCI_CGAP_Kid3"
     polyA_signal    2105. .2110
ORIGIN

Alignment Scores:
Pred. No.:              3.14e-114      Length:         2151
Score:                  1192.00        Matches:        227
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     8              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x HSA420495 (1-2151)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 646

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 706

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 766

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTCCA 826

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 886

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 946

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1006

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1066

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1126

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1186

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1246

Qy        221 LysLeuAspCysGlyArgPhe 227
                 |||||||||||||||||||
Db       1247 AAACTGGACTGTGGTAGATTT 1267

RESULT 13
E23195
LOCUS       E23195                  2499 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Topoisomerase binding protein.
ACCESSION   E23195
VERSION     E23195.1  GI:13024277
KEYWORDS    JP 1999075856-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2499)
  AUTHORS   Takashi,T. and Kazuhiko,Y.
  TITLE     Topoisomerase binding protein
  JOURNAL   Patent: JP 1999075856-A 1 23-MAR-1999;
            TAKASHI TSURUO,CHUGAI PHARMACEUT CO LTD
COMMENT     OS   Unidentified
            PN   JP 1999075856-A/1
            PD   23-MAR-1999
            PF   17-SEP-1997 JP 1997251544
            PR
            PI   TAKASHI TSURUO,KAZUHIKO YAMANE
            PC   C12N15/09,C07K14/47,C12N9/90,C12N15/00
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .2499
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2499
```

```
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                3.74e-114      Length:          2499
Score:                    1192.00        Matches:         227
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x E23195 (1-2499)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 901

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 961

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1021

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1081

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1141

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1201

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 1261

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 1321

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 1381

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 1441

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1501

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||||
Db       1502 AAACTGGACTGTGGTAGATTT 1522

RESULT 14
AC152027/c
LOCUS       AC152027               150344 bp    DNA     linear   HTG 11-NOV-2004
DEFINITION  Dasypus novemcinctus clone VMRC5-12D21, WORKING DRAFT SEQUENCE, 5
            ordered pieces.
ACCESSION   AC152027
VERSION     AC152027.2  GI:55700150
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Dasypus novemcinctus (nine-banded armadillo)
  ORGANISM  Dasypus novemcinctus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Xenarthra; Dasypodidae; Dasypus.
REFERENCE   1  (bases 1 to 150344)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
            Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
            Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Laric,P.,
            Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L.,
            Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A.,
            Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,
            Shah,K., Sison,C., Stantripop,S., Stephen,E., Thomas,J.W.,
            Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
            Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625
            Fishers Lane, Rockville, MD 20852, USA
COMMENT     On Nov 11, 2004 this sequence version replaced gi:54312159.
            ---------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            ---------- Project Information
                 Center project name: ifi
                 Center clone name: 012D21

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            ---------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 149480 bases at least Q40
                 Consensus quality: 149754 bases at least Q30
                 Consensus quality: 149909 bases at least Q20
                 Insert size: 155000; agarose-fp
                 Insert size: 149944; sum-of-contigs
                 Quality coverage: 9.52x in Q20 bases; agarose-fp
                 Quality coverage: 9.84x in Q20 bases; sum-of-contigs
            ----------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    36273: contig of 36273 bp in length
            *    36274    36373: gap of unknown length
            *    36374    55049: contig of 18676 bp in length
            *    55050    55149: gap of unknown length
```

```
            *    55150   101939: contig of 46790 bp in length
            *   101940   102039: gap of unknown length
            *   102040   114806: contig of 12767 bp in length
            *   114807   114906: gap of unknown length
            *   114907   150344: contig of 35438 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150344
                     /organism="Dasypus novemcinctus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9361"
                     /clone="VMRC5-12D21"
                     /clone_lib="VMRC5"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .36273
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    1. .23436
                     /note="clone overlaps with GenBank Accession Number
                     AC152029 clone VMRC5-168B11 (center project name ifj)"
     gap             36274. .36373
                     /estimated_length=unknown
     misc_feature    36374. .55049
                     /note="assembly_fragment"
     gap             55050. .55149
                     /estimated_length=unknown
     misc_feature    55150. .101939
                     /note="assembly_fragment"
     gap             101940. .102039
                     /estimated_length=unknown
     misc_feature    102040. .114806
                     /note="assembly_fragment"
     gap             114807. .114906
                     /estimated_length=unknown
     misc_feature    114907. .150344
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:              2.23e-102       Length:         150344
Score:                  1099.00         Matches:        208
Percent Similarity:     96.0%           Conservative:   10
Best Local Similarity:  91.6%           Mismatches:     9
Query Match:            92.2%           Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AC152027 (1-150344)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db      53237 ATTGATGGATTGGATCTAAACAACCTGCAAGAGAGGGCTCGAGGGGTGTGTTCTTACTTA 53178

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      53177 ACTTTGTACAGTCCAGATGTGATATTTCTACAGGAAGTTATTCCTCCGTACTACAGTTAC 53118

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::|||||||||||||||||||||   |||||||||||||||||||
Db      53117 CTAAAGAAGAGAGCAAGTAATTACGAGATTATTACAGGTCGTGAAGAAGGATATTTCACA 53058

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      53057 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 52998

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db      52997 AGTACCAAAATGATGAGAAACCTATTGTGTGTGCATGTGAGTGTGTCAGGAAATGAACTT 52938

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||   ||||||:::||||||
Db      52937 TACCTTATGACTTCCCATTTAGAGAGCACCAGAGGACATGCTAAGGAACGAATCAATCAG 52878

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||:::|||||||||||||||||||||||||||||:::||||||||||||||||||
Db      52877 TTAAAAGTGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGACAGCTACGGTTATATTTGCA 52818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||:::|||||||||||||||||||||||||||||||||||
Db      52817 GGAGATACGAATTTAAGGGATCAAGAAGTTACCAGATGTGGTGGTTTGCCCAACAACATT 52758

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              :::||||||||||||||||||||||||||:::|||||||||:::|||||||||||||||
Db      52757 TTGGATGTCTGGGAGTTTTTGGGCAAGCCTGAGCATTGCCAGTTTACATGGGATACGCAA 52698

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||   ||||||||||||   |||   ||||||||||||||||   |||||||||
Db      52697 ATGAACTTTAACCTTGGAATACCTGCTACTTGTAAGCTTCGTTTTGATTGAATCTTTTTC 52638

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db      52637 AGAGCAGCAGCAGAAGAGGGCCATATCATTCCCCGAAGTATGGACCTTCTTGGGTTGGAA 52578

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||
Db      52577 AAACTGGACTGTGGCAGATTT 52557

RESULT 15
BD139846/c
LOCUS       BD139846               1079 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Compounds for immunotherapy and diagnosis of breast cancer and
            methods for their use.
ACCESSION   BD139846
VERSION     BD139846.1  GI:23234791
KEYWORDS    JP 2002507387-A/63.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for immunotherapy and diagnosis of breast cancer and
            methods for their use
  JOURNAL   Patent: JP 2002507387-A 63 12-MAR-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002507387-A/63
            PD   12-MAR-2002
            PF   22-DEC-1998 JP 2000526543
            PR   24-DEC-1997 US    08/998253,24-DEC-1997 US    08/998255 PR
            17-JUL-1998 US    09/118627,17-JUL-1998 US    09/118554 PI   STEVEN
            G REED,JIANGCHUN XU
            PC   C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,
            PC   C07K14/47,
            PC
            C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/ PC
             53,
            PC   G01N33/577,C12N15/00,A61K37/02,C12N5/00
            CC   Compounds for immunotherapy and diagnosis
            of breast cancer and
            CC          methods for
            CC   their use
            FH   Key             Location/Qualifiers
            FT   source          1. .1079
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              6.13e-104       Length:         1079
```

```
Score:                 1090.00       Matches:       211
Percent Similarity:    99.5%         Conservative:  0
Best Local Similarity: 99.5%         Mismatches:    1
Query Match:           91.4%         Indels:        1
DB:                    6             Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x BD139846 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 16
AR202950/c
LOCUS       AR202950               1079 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 63 from patent US 6365348.
ACCESSION   AR202950
VERSION     AR202950.1  GI:21499210
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for diagnosis of Breast cancer and methods for their use
  JOURNAL   Patent: US 6365348-A 63 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
```

```
Pred. No.:             6.13e-104     Length:        1079
Score:                 1090.00       Matches:       211
Percent Similarity:    99.5%         Conservative:  0
Best Local Similarity: 99.5%         Mismatches:    1
Query Match:           91.4%         Indels:        1
DB:                    6             Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x AR202950 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 17
AR208051/c
LOCUS       AR208051               1079 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 63 from patent US 6379951.
ACCESSION   AR208051
VERSION     AR208051.1  GI:21507966
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for immunotherapy of breast cancer and methods for their
            use
  JOURNAL   Patent: US 6379951-A 63 30-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              6.13e-104       Length:         1079
Score:                  1090.00         Matches:        211
Percent Similarity:     99.5%           Conservative:   0
Best Local Similarity:  99.5%           Mismatches:     1
Query Match:            91.4%           Indels:         1
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AR208051 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 18
AX429924/c
LOCUS       AX429924               1079 bp    DNA     linear   PAT 21-JUN-2002
DEFINITION  Sequence 63 from Patent WO0198339.
ACCESSION   AX429924
VERSION     AX429924.1  GI:21541088
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.
  TITLE     Compositions and methods for the therapy and diagnosis of breast
            cancer
  JOURNAL   Patent: WO 0198339-A 63 27-DEC-2001;
            CORIXA CORPORATION (US)

FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              6.13e-104       Length:         1079
Score:                  1090.00         Matches:        211
Percent Similarity:     99.5%           Conservative:   0
Best Local Similarity:  99.5%           Mismatches:     1
Query Match:            91.4%           Indels:         1
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AX429924 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 19
BD205491
LOCUS       BD205491               1312 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  CD40-Interacting and TRAF-interacting protein.
ACCESSION   BD205491
VERSION     BD205491.1  GI:33015261
KEYWORDS    JP 2002512796-A/2.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
```

```
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting protein
  JOURNAL   Patent: JP 2002512796-A 2 08-MAY-2002;
            VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT     OS   Mus musculus (mouse)
            PN   JP 2002512796-A/2
            PD   08-MAY-2002
            PF   28-APR-1999 JP 2000546003
            PR   29-APR-1998 EP   98201392.2
            PI   STEFAN MARIA CHRISTIAAN PYPE,
            PI   JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI   DANNY
            FRANCOIS EVELINE HUYLEBROECK
            PC   C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
              A61P37/02,
            PC   A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
            PC   G01N33/566,
            PC   C12N15/00,A61K37/02
            CC   CD40-Interacting and TRAF-interacting protein. FH   Key
                 Location/Qualifiers
            FT   CDS             (122). .(1234).
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
ORIGIN

Alignment Scores:
Pred. No.:                1.03e-88        Length:          1312
Score:                    945.00          Matches:         179
Percent Similarity:       89.0%           Conservative:    23
Best Local Similarity:    78.9%           Mismatches:      23
Query Match:              79.3%           Indels:          2
DB:                       6               Gaps:            1

US-10-757-745-2_COPY_121_347 (1-227) x BD205491 (1-1312)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||     ||||||   ||||||||||||||||||||||||   |||
Db        512 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 571

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||:::|||||||||||||||||||||||||||||   :::|||
Db        572 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTAC 631

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||:::::::::|||   ||||||||||||:::|||||||||||||||||||
Db        632 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 691

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||:::|||||||||   |||||||||   ||||||||||||||||||||||||||||
Db        692 GCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 751

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::|||||||||||||||||||||||||||||:::|||::::::   |||||||||
Db        752 AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 811

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||   |||:::|||||||||:::   |||
Db        812 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 871

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||   ||||||   ||||||||||||||||||:::|||   |||||||||||||||
Db        872 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 931

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||:::||||||   :::|||||||||||||||:::|||:::
Db        932 GGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTT 991

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        992 TTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAA 1051

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||:::||||||   |||   ||||||   |||   ||||||||||||||||||||||
Db       1052 GCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTC 1111

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||   ||||||||||||||:::||||||::||||||||||:::|||||||||
Db       1112 AGA------GCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 1165

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||
Db       1166 AAACTGGACTGTGGTAGATTT 1186

RESULT 20
AR594295
LOCUS       AR594295               1312 bp    DNA     linear   PAT 15-DEC-2004
DEFINITION  Sequence 3 from patent US 6812203.
ACCESSION   AR594295
VERSION     AR594295.1  GI:56643901
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting proteins
  JOURNAL   Patent: US 6812203-A 3 02-NOV-2004;
            Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
            Zwijnaarde;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.03e-88        Length:          1312
Score:                    945.00          Matches:         179
Percent Similarity:       89.0%           Conservative:    23
Best Local Similarity:    78.9%           Mismatches:      23
Query Match:              79.3%           Indels:          2
DB:                       6               Gaps:            1

US-10-757-745-2_COPY_121_347 (1-227) x AR594295 (1-1312)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||     ||||||   ||||||||||||||||||||||||   |||
Db        512 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 571

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||:::|||||||||||||||||||||||||||||   :::|||
Db        572 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTAC 631

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||:::::::::|||   ||||||||||||:::|||||||||||||||||||
Db        632 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 691

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||:::|||||||||   |||||||||   ||||||||||||||||||||||||||||
Db        692 GCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 751

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::|||||||||||||||||||||||||||||:::|||::::::   |||||||||
Db        752 AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 811

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||   |||:::|||||||||:::   |||
Db        812 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 871

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||   ||||||   ||||||||||||||||||:::|||   |||||||||||||||
Db        872 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 931
```

```
Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||:::||||||   :::|||||||||||||||:::|||:::
Db        932 GGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTT 991

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        992 TTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAA 1051

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||:::|||||  |||  ||||||  |||  ||||||||||||||||||||
Db       1052 GCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTC 1111

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||      ||||||||||||||:::|||||||::|||||||||||:::|||||||||
Db       1112 AGA------GCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 1165

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||
Db       1166 AAACTGGACTGTGGTAGATTT 1186

RESULT 21
AX011601
LOCUS       AX011601                1312 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 3 from Patent WO9955859.
ACCESSION   AX011601
VERSION     AX011601.1  GI:9998125
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 3 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
     CDS             122. .1234
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07556.1"
                     /db_xref="GI:9998126"
                     /translation="MASGSSSDAAEPAGPAGRAASAPEAAQAEEDRVKRRRLQCLGFA
                     LVGGCDPTMVPSVLRENDWQTQKALSAYFELPENDQGWPRQPPTSFKSEAYVDLTNED
                     ANDTTILEASPSGTPLEDSSTISFITWNIDGLDGCNLPERARGVCSCLALYSPDVVFL
                     QEVIPPYCAYLKKRAASYTIITGNEEGYFTAILLKKGRVKFKSQEIIPFPNTKMMRNL
                     LCVNVSLGGNEFCLMTSHLESTREHSAERIRQLKTVLGKMQEAPDSTTVIFAGDTNLR
                     DQEVIKCGGLPDNVFDAWEFLGKPKHCQYTWDTKANNNLRIPAAYKHRFDRIFFRAEE
                     GHLIPQSLDLVGLEKLDCGRFPSDHWGLLCTLNVVL"
ORIGIN

Alignment Scores:
Pred. No.:              1.03e-88        Length:         1312
Score:                  945.00          Matches:        179
Percent Similarity:     89.0%           Conservative:   23
Best Local Similarity:  78.9%           Mismatches:     23
Query Match:            79.3%           Indels:         2
DB:                     6               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x AX011601 (1-1312)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||      ||||||   ||||||||||||||||||||||||   |||
Db        512 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 571

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||:::|||||||||||||||||||||||||||  :::|||
Db        572 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTAC 631

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::::::|||  |||||||||||:::|||||||||||||||||||
Db        632 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 691

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||:::||||||||  ||||||||  ||||||||||||||||||||||||||||
Db        692 GCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 751

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::||||||||||||||||||||||||||||||:::|||:::::   |||||||||
Db        752 AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 811

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||  |||:::|||||||||:::  |||
Db        812 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 871

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||  ||||||  ||||||||||||||||||:::|||  |||||||||||||||
Db        872 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 931

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||:::|||||||  :::|||||||||||||||:::|||:::
Db        932 GGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTT 991

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        992 TTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAA 1051

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||:::|||||  |||  ||||||  |||  ||||||||||||||||||||
Db       1052 GCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTC 1111

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||      ||||||||||||||||:::|||||:::|||||||||||||:::|||||||||
Db       1112 AGA------GCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 1165

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||
Db       1166 AAACTGGACTGTGGTAGATTT 1186

RESULT 22
MMU251328
LOCUS       MMU251328               1943 bp    mRNA    linear   ROD 27-MAY-2000
DEFINITION  Mus musculus mRNA for putative TRAF and TNF receptor associated
            protein (TTRAP gene).
ACCESSION   AJ251328
VERSION     AJ251328.1  GI:8247282
KEYWORDS    ttrap gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van
            Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,
            Huylebroeck,D. and Remacle,J.E.
  TITLE     TTRAP, a novel protein that associates with CD40, tumor necrosis
            factor (TNF) receptor-75 and TNF receptor-associated factors
            (TRAFs), and that inhibits nuclear factor-kappa B activation
  JOURNAL   J. Biol. Chem. 275 (24), 18586-18593 (2000)
   PUBMED   10764746
REFERENCE   2  (bases 1 to 1943)
  AUTHORS   Pype,S.M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-NOV-1999) Pype S.M.C., VIB07, Department Cell Growth,
            Differentiation and Developme, Flanders Interuniversity Institute
            for Biotechnology, K.U.Leuven, Campus Gasthuisberg, Herestraat 49,
            3000 Leuven, BELGIUM
```

```
COMMENT     Related sequence: AJ269473.
FEATURES             Location/Qualifiers
     source          1. .1943
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /note="IMAGE cDNA EST clone 1262914"
     gene            1. .1943
                     /gene="TTRAP"
     CDS             48. .1160
                     /gene="TTRAP"
                     /codon_start=1
                     /product="putative TRAF and TNF receptor associated
                     protein"
                     /protein_id="CAB92971.1"
                     /db_xref="GI:8247283"
                     /db_xref="GOA:Q9JJX7"
                     /db_xref="UniProt/TrEMBL:Q9JJX7"
                     /translation="MASGSSSDAAEPAGPAGRAASAPEAAQAEEDRVKRRRLQCLGFA
                     LVGGCDPTMVPSVLRENDWQTQKALSAYFELPENDQGWPRQPPTSFKSEAYVDLTNED
                     ANDTTILEASPSGTPLEDSSTISFITWNIDGLDGCNLPERARGVCSCLALYSPDVVFL
                     QEVIPPYCAYLKKRAASYTIITGNEEGYFTAILLKKGRVKFKSQEIIPFPNTKMMRNL
                     LCVNVSLGGNEFCLMTSHLESTREHSAERIRQLKTVLGKMQEAPDSTTVIFAGDTNLR
                     DQEVIKCGGLPDNVFDAWEFLGKPKHCQYTWDTKANNNLRIPAAYKHRFDRIFFRAEE
                     GHLIPQSLDLVGLEKLDCGRFPSDHWGLLCTLNVVL"
ORIGIN

Alignment Scores:
Pred. No.:              1.63e-88        Length:         1943
Score:                  945.00          Matches:        179
Percent Similarity:     89.0%           Conservative:   23
Best Local Similarity:  78.9%           Mismatches:     23
Query Match:            79.3%           Indels:         2
DB:                     9               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x MMU251328 (1-1943)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||      ||||||   |||||||||||||||||||||||||   |||
Db        438 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 497

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||:::|||||||||||||||||||||||||||   :::|||
Db        498 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTAC 557

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||:::::::::|||   ||||||||||:::|||||||||||||||||||
Db        558 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 617

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||:::|||||||||   |||||||||   ||||||||||||||||||||||||||
Db        618 GCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 677

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::||||||||||||||||||||||||||||:::|||::::::   |||||||||
Db        678 AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 737

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||   |||:::||||||||:::   |||
Db        738 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 797

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||   ||||||   ||||||||||||||||||:::|||   ||||||||||||||||
Db        798 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 857

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||:::|||||||   :::|||||||||||||||:::|||:::
Db        858 GGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTT 917

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
                 |||   ||||||||||||||||||||||||||||||||||||||||||||||:::
Db        918 TTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAA 977

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||:::||||||   |||   ||||||   |||   ||||||||||||||||||||||
Db        978 GCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTC 1037

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||      ||||||||||||||:::||||||:::||||||||||||:::||||||||||
Db       1038 AGA------GCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 1091

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1092 AAACTGGACTGTGGTAGATTT 1112

RESULT 23
CR761829
LOCUS       CR761829               1236 bp    mRNA    linear   VRT 15-SEP-2004
DEFINITION  Xenopus tropicalis finished cDNA, clone TGas102g02.
ACCESSION   CR761829
VERSION     CR761829.1  GI:51967738
KEYWORDS    .
SOURCE      Xenopus tropicalis (Silurana tropicalis)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1236)
  AUTHORS   Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
            Francis,M.D., Garrett,N., Gilchrist,M.J., Grafham,D.V.,
            McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
            Voigt,J. and Zorn,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
COMMENT     Sanger Xenopus tropicalis EST/cDNA project.
            This sequence is from a Xenopus Gene Collection (XGC) library, from
            a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
            extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
            ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'
            end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1-blue.
FEATURES             Location/Qualifiers
     source          1. .1236
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TGas102g02"
                     /clone_lib="XGC-gastrula"
                     /dev_stage="gastrula (stage 10.5-13 mixed)"
ORIGIN

Alignment Scores:
Pred. No.:              2.82e-82        Length:         1236
Score:                  883.00          Matches:        162
Percent Similarity:     85.5%           Conservative:   32
Best Local Similarity:  71.4%           Mismatches:     33
Query Match:            74.1%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CR761829 (1-1236)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||   :::|||::::::|||||||||||||||||||||||||||
Db        442 ATAGATGGCCTTGATGAATCAAATGTCGCAGAAAGGGCTCGTGGTGTGTGTTCCTATTTG 501

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||:::|||||||||||||||||||||||||||||   |||
Db        502 GCCCTGTACAGTCCAGATGTAGTATTCTTACAGGAGGTTATTCCACCATACTATGAGTAC 561

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||:::   :::|||   ||||||||||:::|||:::   |||||||||
Db        562 CTGAAGAAGAGAGCTGTCAGCTACACAATTATTACAGGTAATGAGGATGAATACTTCACT 621

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
```

```
Db        622 GCTATGATGCTGAAAAAGAGCAGGGTAAAGTTAATAAGCCAAGAAATTGTACCATACCCA 681

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100

Db        682 TCCACGGTGATGATGAGAAATTTGCTTGTAGCAAATGTGAATATCTCTGGCAACAGTATA 741

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120

Db        742 TGCTTAATGACTTCTCATCTAGAGAGCACAAAAGATCATTCAAAGGAGCGATTAAAGCAG 801

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140

Db        802 CTGGACATTGTGCTAAAAAAGATGATGGATGCTCCCCCTTTAGCAACTGTAATATTTGGA 861

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db        862 GGGGATACAAACCTGCGAGATCAGGAGGTTGCTAAAATTGGAGGCATGCCAAATAATATA 921

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180

Db        922 TTGGATGTTTGGGACTTCTTGGGTAAGCCTGAGCACTGTCGCTATACCTGGGACACAAAA 981

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200

Db        982 GTGAACAAGAACCTGCGTGCTCCCTACATTTGCAGGTTGCGCTTTGACCGCATATTTTTT 1041

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220

Db       1042 AGAGCTTCGCAGGAGGGAAGTCAAGTGATCCCCCAGTCTCTTTATCTTGCTGGCACTGAA 1101

Qy        221 LysLeuAspCysGlyArgPhe 227

Db       1102 AAGCTGGACTGTGGCAGATTT 1122
```

```
RESULT 24
CQ726600
LOCUS       CQ726600                858 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 12534 from Patent WO02068579.
ACCESSION   CQ726600
VERSION     CQ726600.1  GI:42290170
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 12534 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .858
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              8.34e-68        Length:         858
Score:                  742.50          Matches:        150
Percent Similarity:     66.1%           Conservative:   0
Best Local Similarity:  66.1%           Mismatches:     0
Query Match:            62.3%           Indels:         77
DB:                     6               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x CQ726600 (1-858)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
```

```
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGG---------- 405

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40

Db        405 ---------------------------------------------------------- 405

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60

Db        405 ---------------------------------------------------------- 405

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80

Db        405 ---------------------------------------------------------- 405

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100

Db        406 ----------------------------------GTGAATGTGTCAGGAAATGAGCTT 429

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120

Db        430 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 489

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140

Db        490 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 549

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db        550 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 609

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180

Db        610 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 669

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200

Db        670 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 729

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220

Db        730 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 789

Qy        221 LysLeuAspCysGlyArgPhe 227

Db        790 AAACTGGACTGTGGTAGATTT 810
```

```
RESULT 25
BC083404
LOCUS       BC083404                1451 bp    mRNA    linear   VRT 27-MAY-2005
DEFINITION  Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7238334), partial
            cds.
ACCESSION   BC083404
VERSION     BC083404.1  GI:53734140
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 1451)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
```

```
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1451)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: John Postlethwait, University of Oregon
            cDNA Library Preparation: Dr. Sumio Sugano
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 192 Row: c Column: 18.
FEATURES             Location/Qualifiers
     source          1. .1451
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:7238334"
                     /tissue_type="Ovary"
                     /clone_lib="NIH_ZGC_5"
                     /lab_host="DH10B"
                     /note="Vector: pME18S-FL3"
     gene            <1. .1451
                     /gene="TTRAP"
                     /db_xref="GeneID:553516"
     CDS             <1. .1141
                     /gene="TTRAP"
                     /codon_start=2
                     /product="TTRAP protein"
                     /protein_id="AAH83404.1"
                     /db_xref="GI:53734141"
                     /db_xref="GeneID:553516"
                     /translation="FLSGNQTDQQMSALEESRTRLCDQFAFVSGSDSAVAQCYLAENE
                     WDMERALNSFFEAHMDSVFDEEAEKTEVTGNKRKDDTAEASGTKKKLKTDNADCIDLT
                     AEEPTCSITVNSKENQAENGTAKSEVEDSKLSIISWNVDGLDTLNLADRARGLCSYLA
                     LYTPDVVFLQELIPAYVQYLKKRAVSYLFFEGSDDGYFTGIMLRKSRVKFLESEIICF
                     PTTQMMRNLLIAQVTFSGQKLYLMTSHLESCKNQSQERTKQLRVVLQKIKEAPEDAIV
                     IFAGDTNLRDAEVANVGGLPAGVCDVWEQLGKQEHCRYTWDTKANSNKTVPYVSRCRF
                     DRIFLRSAKTGPPVTPDHMALIGMEKLDCGRYTSDHWGIYCTFNT"
ORIGIN

Alignment Scores:
Pred. No.:                3.78e-64        Length:          1451
Score:                    710.00          Matches:         134
Percent Similarity:       74.0%           Conservative:    34
Best Local Similarity:    59.0%           Mismatches:      59
Query Match:              59.6%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BC083404 (1-1451)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              :::|||||||||||||       ||||||::::::||||||||||||:::||||||||||||

Db        419 GTGGACGGTTTGGACACATTAAATCTTGCAGATCGTGCCAGGGGCTTGTGCTCATATCTA 478

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||:::||||||||:::||||||||||||:::||||||   |||       |||

Db        479 GCTCTATACACACCAGATGTGGTTTTTCTTCAAGAGCTCATTCCAGCTTATGTTCAGTAT 538

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||||||||||:::   :::|||           |||   ::::::||||||||||||

Db        539 CTTAAGAAACGTGCCGTCAGCTACTTGTTTTTTGAAGGAAGCGATGATGGATACTTTACC 598

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||||||||:::|||||||||||||||           |||||||||   ||||||

Db        599 GGGATAATGTTGAGGAAATCAAGAGTCAAATTTTTGGAAAGTGAGATCATCTGCTTTCCA 658

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::|||:::||||||||||||||||||           |||       ||||||   :::|||

Db        659 ACAACACAAATGATGAGGAATCTGCTCATTGCTCAGGTGACTTTCTCAGGTCAGAAGCTA 718

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||   :::       :::   ||||||      |||

Db        719 TACTTGATGACATCCCATTTAGAAAGCTGTAAAAACCAATCCCAGGAGAGAACGAAACAA 778

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||::::::||||||:::|||::::::||||||||||||   |||   ||||||||||||

Db        779 TTGCGGGTTGTGCTTCAAAAGATTAAGGAAGCGCCAGAAGATGCCATTGTCATTTTTGCA 838

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||   ||||||      ||||||||||||      :::

Db        839 GGTGACACAAATCTCAGAGACGCAGAGGTGGCCAATGTTGGAGGTTTGCCTGCAGGTGTG 898

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||   |||||||||   :::||||||:::||||||||||||||||:::

Db        899 TGTGATGTGTGGGAACAGCTGGGCAAACAAGAACACTGCAGGTACACCTGGGACACCAAA 958

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||||||||      :::           :::   ||||||||||||||||||

Db        959 GCTAACAGCAATAAGACTGTGCCTTACGTCAGCAGATGTCGCTTCGACCGCATATTCCTG 1018

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||:::|||           :::   |||      :::   |||:::|||:::|||

Db       1019 AGGTCTGCTAAAACCGGTCCGCCGGTCACTCCTGACCACATGGCCTTGATTGGAATGGAG 1078

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||:::
Db       1079 AAACTGGACTGCGGTCGTTAT 1099

RESULT 26
BC097117
LOCUS       BC097117                1455 bp    mRNA    linear   VRT 23-JUN-2005
DEFINITION  Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7431816), partial
            cds.
ACCESSION   BC097117
VERSION     BC097117.1  GI:66910444
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 1455)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
```

```
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1455)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Will Talbot
            cDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 226 Row: g Column: 18
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 53734140.
FEATURES             Location/Qualifiers
     source          1. .1455
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:7431816"
                     /tissue_type="Embryo, whole, 2-8 hr, mixed wild type, pool
                     of 2500 embryos"
                     /clone_lib="NIH_ZGC_15"
                     /lab_host="DH10B"
                     /note="Vector: pME18S-FL3"
     gene            <1. .1455
                     /gene="TTRAP"
                     /db_xref="GeneID:553516"
     CDS             <1. .438
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TTRAP protein"
                     /protein_id="AAH97117.1"
                     /db_xref="GI:66910445"
                     /db_xref="GeneID:553516"
                     /translation="VFSGNQTDQQMSALGESRTRLCDQFAFVSGSDSAVAQCYLAENE
                     WDMERALNSFFEAHMDSVFDEEAEKTEVTGNKRKDDTAEASGTKKKLKTDNADCIDLT
                     AEEPTCSITVNSKKIRLKMAQPNLKWRTASSPSSAGMWTVWTH"
ORIGIN

Alignment Scores:
Pred. No.:              7.79e-64        Length:         1455
Score:                  707.00          Matches:        134
Percent Similarity:     73.6%           Conservative:   33
Best Local Similarity:  59.0%           Mismatches:     60
Query Match:            59.3%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BC097117 (1-1455)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              :::|||||||||||||       ||||||:::::::|||||||||||||:::|||||||||||||
Db        417 GTGGACGGTTTGGACACATTAAATCTTGCAGATCGTGCCAGGGGCTTGTGCTCATATCTA 476

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              |||||||||:::|||||||||:::|||||||||||||:::||||||   |||      |||
Db        477 GCTCTATACACACCAGATGTGGTTTTTCTTCAAGAGCTCATTCCAGCTTATGTTCAGTAT 536

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||||||||||||:::   :::|||             |||   ::::::|||||||||||||
Db        537 CTTAAGAAACGTGCCGTCAGCTACTTGTTTTTTGAAGGAAGCGATGATGGATACTTTACC 596

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||:::|||||||||||||||||           |||||||||   ||||||
Db        597 GGGATAATGTTAAGGAAATCAAGAGTAAAATTTTTGGAAAGCGAGATCATCTGTTTTCCA 656

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::|||:::|||||||||||||||||||||      |||      ||||||   :::|||
Db        657 ACAACACAAATGATGAGGAATCTGCTCATTGCTCAGGTGACTTTCTCAGGTCAGAAGCTA 716

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||   :::      :::   ||||||       |||
Db        717 TACTTGATGACATCCCATTTAGAAAGCTGTAAAAACCAATCCCAGGAGAGAACGAAACAA 776

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||::::::||||||:::|||::::::|||||||||||   |||   |||||||||||
Db        777 TTGCGGGTTGTGCTTCAAAAGATTAAGGAAGCGCCAGAAGATGCCATTGTCATTTTTGCA 836

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||   ||||||      ||||||||||||      :::
Db        837 GGCGACACAAATCTCAGAGACGCAGAGGTGGCCAATGTTGGAGGTTTGCCTGCAGGTGTG 896

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||||||||||   |||||||||   :::||||||:::||||||||||||||||:::
Db        897 TGTGATGTGTGGGAACAGCTGGGCAAACAAGAACACTGCAGGTACACCTGGGACACCAAA 956

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||||||||      :::           :::   ||||||||||||||||||
Db        957 GCTAACAGCAATAAGACTGTGCCTTACGTCAGCAGATGTCGCTTCGACCGCATATTCCTG 1016

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||:::|||               :::   |||          |||:::|||:::||||
Db       1017 AGGTCTGCTAAAACCGCTCCGCCGGTCACTCCTGACCACAAGGCCTTGATTGGAATGGAG 1076

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||:::
Db       1077 AAACTGGACTGCGGTCGTTAT 1097

RESULT 27
CQ663156
LOCUS       CQ663156                 400 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 8082 from Patent WO02070737.
ACCESSION   CQ663156
VERSION     CQ663156.1  GI:42140949
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 8082 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.77e-52        Length:         400
Score:                  590.00          Matches:        111
Percent Similarity:     100.0%          Conservative:   1
```

```
Best Local Similarity:  99.1%          Mismatches:     0
Query Match:            49.5%          Indels:         0
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CQ663156 (1-400)

Qy        103 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 122
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GTGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAA 124

Qy        123 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGAT 184

Qy        143 ThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGAT 244

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 GTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 304

Qy        183 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAla 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCA 364

Qy        203 AlaAlaGluGluGlyHisIleIleProArgSerLeu 214
              |||||||||||||||||||||||||||||||||||
Db        365 GCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 400

RESULT 28
BD150065
LOCUS       BD150065                752 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD150065
VERSION     BD150065.1  GI:27855823
KEYWORDS    JP 2002191363-A/4908.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 4908 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/4908
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                    Location/Qualifiers
            FT   source          1. .752
            FT                   /organism='Homo sapiens (human)'.
FEATURES              Location/Qualifiers
     source           1. .752
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.14e-49       Length:         752
Score:                  568.00         Matches:        119
Percent Similarity:     90.9%          Conservative:   1
Best Local Similarity:  90.2%          Mismatches:     9
Query Match:            47.7%          Indels:         3
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BD150065 (1-752)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArg-AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 660

Qy        100 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 120
              |||||||   ||||||   |||    :::|||   ||||||      ||||||||||||||
Db        661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy        120 lnLeuLysMetValLeu-LysLysMetGln 129
              ||   ||||||   |||   |||||||||
Db        721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750

RESULT 29
AX870003
LOCUS       AX870003                752 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 4908 from Patent EP1074617.
ACCESSION   AX870003
VERSION     AX870003.1  GI:40024866
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 4908 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES              Location/Qualifiers
     source           1. .752
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.14e-49       Length:         752
Score:                  568.00         Matches:        119
Percent Similarity:     90.9%          Conservative:   1
Best Local Similarity:  90.2%          Mismatches:     9
Query Match:            47.7%          Indels:         3
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AX870003 (1-752)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr  40

Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr  60

Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro  80

Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArg-AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100

Db        601 AGTACCAAAATGATGAGAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCT 660

Qy        100 uCysLeuMetThrSerHisLeu-GluSerThrArgGlyHisAlaAlaGluArgMetAsnG 120

Db        661 TTGCCTTAAGACATCCCAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATC 720

Qy        120 lnLeuLysMetValLeu-LysLysMetGln 129

Db        721 AGTTTAAAATGGGTTTAAANGAAAATGCAA 750
```

```
RESULT 30
AL589699
LOCUS       AL589699               256608 bp    DNA     linear   ROD 17-DEC-2004
DEFINITION  Mouse DNA sequence from clone RP23-92G13 on chromosome 13 Contains
            the 3' end of the gene for a novel protein similar to KIAA0386,
            five novel genes, the Gmnn gene for geminin, gene PNAS-27, the
            Ttrap gene for Traf and Tnf receptor associated protein, the gene
            for a novel protein similar to KIAA0319, the gene for the ortholog
            of human and rat aldehyde dehydrogenase 5A1 (succinate-semialdehyde
            dehydrogenase) (ALDH5A1) (SSADH), the 5' end of the Gpld1 gene for
            glycosylphosphatidylinositol specific phospholipase D1 and four CpG
            islands, complete sequence.
ACCESSION   AL589699
VERSION     AL589699.4  GI:13810107
KEYWORDS    HTG; Aldh5a1; CpG island; geminin; Gmnn; Gpld1; KIAA0319; KIAA0386;
            phospholipase; SSADH; Ttrap.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 256608)
  AUTHORS   Pelan,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 26, 2001 this sequence version replaced gi:13751568.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep ----------
            Genome Center
            Center: UK Medical Research Council
            Center code: UK-MRC
            Web site: http://mrcseq.har.mrc.ac.uk
            Contact: mouseq@har.mrc.ac.uk
            ----------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence.  Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
            RP23-92G13 is from the RPCI-23 Mouse BAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES                Location/Qualifiers
     source             1. .256608
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10090"
                        /chromosome="13"
                        /clone="RP23-92G13"
                        /clone_lib="RPCI-23"
     gene               join(AL513014.9:11808. .79376,4770. .28457)
                        /locus_tag="RP23-461O24.1-003"
     mRNA               join(AL513014.9:11808. .12037,AL513014.9:43061. .43187,
                        AL513014.9:44856. .45011,AL513014.9:46133. .46211,
                        AL513014.9:47748. .47771,AL513014.9:48876. .48929,
                        AL513014.9:50825. .50974,AL513014.9:53729. .53792,
                        AL513014.9:62567. .62610,AL513014.9:66867. .66992,
                        AL513014.9:67218. .67366,AL513014.9:68674. .68803,
                        AL513014.9:74163. .74831,AL513014.9:79192. .79376,
                        4770. .4932,8263. .8398,9931. .10092,12188. .12346,
                        16272. .16474,18350. .18433,19587. .19677,26226. .28457)
                        /locus_tag="RP23-461O24.1-003"
                        /note="match: cDNAs: AK007145.1 AB002384.1
                        match: ESTs: AW822704.1 BB674120.1 AA041998.1 BM144846.1
                        Z42557.1 R20596.1"
     CDS                join(AL513014.9:11989. .12037,AL513014.9:43061. .43187,
                        AL513014.9:44856. .45011,AL513014.9:46133. .46211,
                        AL513014.9:47748. .47771,AL513014.9:48876. .48929,
                        AL513014.9:50825. .50974,AL513014.9:53729. .53792,
                        AL513014.9:62567. .62610,AL513014.9:66867. .66992,
                        AL513014.9:67218. .67366,AL513014.9:68674. .68803,
                        AL513014.9:74163. .74831,AL513014.9:79192. .79376,
                        4770. .4932,8263. .8398,9931. .10092,12188. .12346,
                        16272. .16474,18350. .18433,19587. .19677,26226. .26326)
                        /locus_tag="RP23-461O24.1-003"
                        /standard_name="OTTMUSP00000000566"
                        /note="match: proteins: Q9D9C2 Q9Y4F9"
                        /codon_start=1
                        /protein_id="CAI26080.1"
                        /db_xref="GI:56237839"
                        /translation="MPPGTKRLRALGAFSAGLPTRLPEIMLVGSQSFSPGGPNGIIRS
                        QSFAGFSGLQERRSRCNSFIENASALKKPQAKLKKMHNLGHKNNNTPKEPQPKRVEEV
                        YRALKNGLDEYLEFHQTELDKLTAQLKDMKRNSRLGVLYDLDKQIKTIERYMRRLEFH
                        ISKVDELYEAYCIQRRLQDGASKMKQAFATSPASKAARESLSEINRSYKEYTENMCAI
                        EAELESLLGEFSIKMKGLAGFARLCPGDQYEIFMKYGRQRWKLKGKIEVNGKQSWDGA
                        ETVFLPLIVGFISIKVTELKGLATHILVGSVTCETKELFAARPQVVAVDINDLGTIKL
                        NLEITWYPFDVEDTTPSSGPGNKTAALQRRMSMYSQGTPETPTFKDQSFFSNLPDDIF
                        ESGKAAEEKRPLSLSFSDLQDGDCVFTSSSATSPSSSHSAHPEITITPAELTHSSLSS
                        QNEGTEDSSSASSRNSLGEDHEPKSHPKSDTVEPGKPGVATRSGTESLFLESSVAEAL
                        LQESDEASELKPVELDTFEGNITKQLVKRLTSAEGPVTTDKLFFEGSVGSESEAGRSF
                        LDGSLEDAFNGLFLALDPHKKQYKEFQDLNQEVTHLDDVLKCKPAGSRSRSSSLSLTV
                        ESALESFDFLNTSDFDEEEEDGDDVCHVGGGADSVFSDTETEKSGSVHPEARGHLSEA
                        LTEDTGVGTSVAGSPLPLTTGNESLDITIVKHLQYCTQLIQQIVFSSKTPFVARSLLE
                        KLSRQVLVMQKLAAVSDENLGNITSVVEAIPEFHKKLSLLAFWTKCCSPSGVYHSSAA
                        RLIKQLEASFARSINKDYPGLAEPVFRTLVSQILDRAEPLLSSSLSSEVITVFQYYSF
                        FTSHGVSDLETYLGQLTRQVAMVQTLQSLRDEKLLQTMSDLAPSNLPAQQEVLRTLAL
                        LLTRDDNEVSEAVTLYLAAASKNEHFREKALLYYCEALTKANLQLQKAACLALKSLEA
                        TESIKMLVTLCQSDTEEIRTVASETLLSLGEDGRLAYEQLDKFPRDCVKVGGRHGTEV
                        ATAF"
     polyA_signal       26424. .26429
                        /locus_tag="RP23-461O24.1-003"
     polyA_site         26446
                        /locus_tag="RP23-461O24.1-003"
     polyA_site         26447
                        /locus_tag="RP23-461O24.1-003"
     polyA_site         26457
                        /locus_tag="RP23-461O24.1-003"
     gene               join(35873. .36201,40700. .41233)
```

```
                /locus_tag="RP23-92G13.3-001"
mRNA            join(35873. .36201,40700. .41233)
                /locus_tag="RP23-92G13.3-001"
                /note="match: cDNAs: AK006029.1
                match: ESTs: AV045211 BF018417 BB012811 BB014319 BB020000
                AV263251 BF319523 AV206925 AV041162 AV259155 AV041846"
polyA_signal    39211. .39216
                /locus_tag="RP23-92G13.3-001"
polyA_site      39233
                /locus_tag="RP23-92G13.3-001"
polyA_site      complement(44486)
polyA_site      complement(44487)
polyA_site      complement(44488)
polyA_signal    complement(44511. .44516)
polyA_signal    complement(44515. .44520)
gene            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56454. .56564))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-001"
mRNA            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56454. .56564))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-001"
                /note="match: ESTs: BI789536.1 BQ178448.1 BF722692.1
                AA510015.1 BG797776.1 BB868504.1 AA014584.1 AA032471.1
                BI454155.1 BE448331.1 W13399.1 BF556131.1 BG085437.1
                AI528457.1 AA277821.1 AA681520.1 W99871.1 AA004051.1"
gene            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56228. .56502))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-002"
mRNA            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56228. .56502))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-002"
                /note="match: cDNAs: AF068780.1 BC020061.1
                match: ESTs: BF160648.1 AA437903.1"
CDS             complement(join(46710. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54851))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-001"
                /standard_name="OTTMUSP00000000569"
                /note="match: proteins: O88513 O75496"
                /codon_start=1
                /protein_id="CAI26081.1"
                /db_xref="GI:56237840"
                /translation="MNLSMKQKQEGAQENVKNSPVPRRTLKMIQPSADGSLVGRENEL
                PKGLFKRKLWDDQLASQTSSCGPEANENKDVGDLTQEAFDLISKENPSSQYWKEVAEQ
                RRKALYEALKENEKLHKEIEQKDSEIARLRKENKDLAEVAEHVQYMAEVIERLSNEPL
                DNFESPDSQEFDSEEEAVEYSELEDSGAGTCAEETVSSSTDARPCT"
gene            96664. .108654
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
mRNA            join(96664. .98150,99423. .99599,100503. .100625,
                104527. .104661,107278. .108654)
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
                /note="match: cDNAs: BC002217.1 BC005537.1 AL136632.1
                AK022681.1 AJ420475.1
                match: ESTs: BI078093.1 BG295445.1 BI696541.1 BG717389.1
                BI108344.1 BG975946.1 BI452750.1 BG168325.1 BG295139.1
                BG914583.1 BI753721.1"
CDS             join(98022. .98150,99423. .99599,100503. .100625,
                104527. .104661,107278. .107403)
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
                /standard_name="OTTMUSP00000000563"
                /note="match: proteins: Q9H1Z2 Q99LU8 Q9GZU0 O95550"
                /codon_start=1
                /protein_id="CAI26082.1"
                /db_xref="GI:56237841"
                /translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
                KKSALFEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
                DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEFHHGDYEKQFLHV
                LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
                IEDHLRPYMPE"
polyA_signal    106607. .106612
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
polyA_site      106654
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
```

Alignment Scores:

Pred. No.:	1.17e-35	Length:	256608
Score:	462.00	Matches:	104
Percent Similarity:	38.3%	Conservative:	15
Best Local Similarity:	33.4%	Mismatches:	14
Query Match:	38.8%	Indels:	178
DB:	9	Gaps:	2

US-10-757-745-2_COPY_121_347 (1-227) x AL589699 (1-256608)

```
Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              |||::::::   |||||||||  ||||||||||||||||||||||||||||||||||
Db     135029 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 135088

Qy         113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              |||:::|||||||||:::   |||||||||   ||||||   ||||||||||||||||||
Db     135089 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 135148

Qy         133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr--- 151
              :::|||   |||||||||||||||||||||||||||||||||:::||||||:::
Db     135149 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTAAGTAAC 135208

Qy         151 ------------------------------------------------------------ 151

Db     135209 AGAATTGATTCAATGTTTGTGGGGAGAGTATTGCTATGGAAATTAATTTTGATTCATTTT 135268

Qy         151 ------------------------------------------------------------ 151

Db     135269 ATCCAAAGTCTTATGGACTTGAAATACAAGTTAGAAAACACTGTTGTAGCTTTTGTTTAA 135328

Qy         151 ------------------------------------------------------------ 151

Db     135329 AATATGCACAAAAGCTATTAGGAGAAATCCTCTTATATGTTAGATAAACAAGTATTAAAA 135388

Qy         151 ------------------------------------------------------------ 151

Db     135389 TAAGATAATAATTATAAGAATTCTAATTTTAAAAGTTGAGGGTATTTGTAGCATCCCTGT 135448

Qy         151 ------------------------------------------------------------ 151

Db     135449 GTATAAGAAGATCTAATTCTCTTCTGAGTTGAAATCAACTTTTTAAAATACTTTCTTGCT 135508

Qy         151 ------------------------------------------------------------ 151

Db     135509 GGGTGTGGTGGCCCATGCCTTAATCTCAGCACTTGGGAGGCGGAGACAGGCAGATCTCTG 135568

Qy         151 ------------------------------------------------------------ 151

Db     135569 AGTTACAGAGTGAATTCCAGCACAGCTCAGATACACAGAAACCCTGTCTTGAAAAACAAC 135628

Qy         151 ------------------------------------------------------------ 151

Db     135629 AACAAAAGCTGTACTTACAAATGATAATATCAGACACCTTTTAAATTGGGGGCAGGGAGT 135688

Qy         152 ----------------------------------------------ArgCysGlyGlyLeu 156
                                                            :::|||||||||||
Db     135689 AAGTATATAAACCTTTCATAATATTCTTTCTCTTGCCAGGTTATCAAATGTGGTGGTTTA 135748

Qy         157 ProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 176
              |||:::|||:::   |||  ||||||||||||||||||||||||||||||||||||||
Db     135749 CCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACA 135808

Qy         177 TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAsp 196
```

```
              ||||||||||:::   |||:::||||||   |||   ||||||   |||   |||||||||
Db     135809 TGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGAT 135868

Qy        197 ArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeu 216
              ||||||||||||||     ||||||||||||||:::||||||:::|||||||||||||
Db     135869 CGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTT 135922

Qy        217 LeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              :::|||||||||||||||||||||||||||||
Db     135923 GTTGGGTTGGAAAAACTGGACTGTGGTAGATTT 135955

RESULT 31
AY613922
LOCUS       AY613922               20938 bp    DNA     linear   PRI 11-MAY-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein (TTRAP) gene,
            complete cds.
ACCESSION   AY613922
VERSION     AY613922.1  GI:47060296
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 20938)
  AUTHORS   Rieder,M.J., Daniels,R.L., da Ponte,S.H., Hastings,N.C.,
            Ahearn,M.O., Rajkumar,N., Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2004) Genome Sciences, University of Washington,
            1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  SeattleSNPs. NHLBI HL66682 Program
            for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.gs.washington.edu).
FEATURES             Location/Qualifiers
     source          1. .20938
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     repeat_region   1. .32
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     variation       81
                     /frequency="0.67"
                     /replace=""
     repeat_region   87. .368
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     repeat_region   369. .673
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       449
                     /frequency="0.96"
                     /replace=""
     variation       540
                     /frequency="0.01"
                     /replace="t"
     variation       586
                     /frequency="0.01"
                     /replace="a"
     variation       650
                     /frequency="0.03"
                     /replace="g"
     variation       658. .660
                     /frequency="0.01"
                     /replace=""
     repeat_region   683. .989
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       695
                     /frequency="0.18"
                     /replace="t"
     variation       922
                     /frequency="0.02"
                     /replace="t"
     variation       1319
                     /frequency="0.15"
                     /replace="t"
     variation       1456
                     /frequency="0.19"
                     /replace="a"
     variation       1532
                     /frequency="0.01"
                     /replace="c"
     variation       1711. .1713
                     /frequency="0.04"
                     /replace=""
     variation       1795
                     /frequency="0.10"
                     /replace="a"
     variation       1821
                     /frequency="0.02"
                     /replace="t"
     variation       1830. .1862
                     /frequency="0.04"
                     /replace=""
     variation       1913. .1914
                     /frequency="0.13"
                     /replace=""
     gene            <2005. .18938
                     /gene="TTRAP"
     mRNA            join(<2005. .2194,2281. .2366,10172. .10345,11003. .11094,
                     14376. .14494,15753. .15923,17846. .18938)
                     /gene="TTRAP"
                     /product="TRAF and TNF receptor associated protein"
     CDS             join(2030. .2194,2281. .2366,10172. .10345,11003. .11094,
                     14376. .14494,15753. .15923,17846. .18127)
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TRAF and TNF receptor associated protein"
                     /protein_id="AAT09764.1"
                     /db_xref="GI:47060297"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
     variation       2381
                     /gene="TTRAP"
                     /frequency="0.08"
                     /replace="a"
     variation       2767
                     /gene="TTRAP"
                     /frequency="0.08"
                     /replace="g"
     variation       2800
                     /gene="TTRAP"
                     /frequency="0.79"
                     /replace=""
     repeat_region   2919. .3323
                     /rpt_family="L2"
                     /rpt_type=dispersed
     variation       2927
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="g"
     variation       3004
                     /gene="TTRAP"
                     /frequency="0.04"
                     /replace="t"
     variation       3061
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="t"
     variation       3062
```

```
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="t"
     variation       3114
                     /gene="TTRAP"
                     /frequency="0.21"
                     /replace="c"
     variation       3418
                     /gene="TTRAP"
                     /frequency="0.03"
                     /replace="c"
     variation       3781
                     /gene="TTRAP"
                     /frequency="0.15"
                     /replace="g"
     variation       3988
                     /gene="TTRAP"
                     /frequency="0.01"
                     /replace="a"
     variation       4191
                     /gene="TTRAP"
                     /frequency="0.01"
                     /replace="c"
     variation       4393
                     /gene="TTRAP"
                     /frequency="0.97"
                     /replace=""
     variation       4414
                     /gene="TTRAP"
                     /frequency="0.03"
                     /replace="c"
     variation       4437
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="g"
     repeat_region   4446. .4530
                     /rpt_family="MER1_type"
                     /rpt_type=dispersed
     repeat_region   4559. .4609
                     /rpt_family="MER1_type"
                     /rpt_type=dispersed
     variation       4565
                     /gene="TTRAP"
                     /frequency="0.52"
                     /replace=""
     repeat_region   4625. .4760
                     /rpt_family="MER1_type"
                     /rpt_type=dispersed
     variation       4658
                     /gene="TTRAP"
                     /frequency="0.01"
                     /replace="a"
     variation       4861
                     /gene="TTRAP"
                     /frequency="0.02"
                     /replace="a"
     variation       4941
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="g"
     variation       4983
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="t"
     variation       5024
                     /gene="TTRAP"
                     /frequency="0.07"

Alignment Scores:
Pred. No.:                1.08e-33      Length:          20938
Score:                    431.00        Matches:         78
Percent Similarity:       100.0%        Conservative:    1
Best Local Similarity:    98.7%         Mismatches:      0

Query Match:              36.2%         Indels:          0
DB:                       8             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x AY613922 (1-20938)

Qy          149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
                ::::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        17843 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 17902

Qy          169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 188
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        17903 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 17962

Qy          189 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 208
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        17963 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACAC 18022

Qy          209 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        18023 ATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 18079

RESULT 32
HS30M3
LOCUS       HS30M3                 102200 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP1-30M3 on chromosome 6p22.1-22.3
            Contains the 5' end of gene KIAA0319, the gene for TRAF and TNF
            receptor-associated protein (TTRAP)(EAP2, AD022, MGC9099), gene
            FLJ20501, the 3' end of gene FLJ12619 (DKFZP564G182) and two CpG
            islands, complete sequence.
ACCESSION   AL031775
VERSION     AL031775.1  GI:4071041
KEYWORDS    HTG; AD022; CpG island; DKFZP564G182; FLJ12619; FLJ20501; KIAA0319;
            TNF; TRAF; TTRAP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 102200)
  AUTHORS   Phillips,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 29, 1998 this sequence version replaced gi:4007165.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP1-30M3 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES                Location/Qualifiers
     source             1. .102200
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704M0330"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="p22.1-22.3"
                     /clone="RP1-30M3"
                     /clone_lib="RPCI-1"
     gene            join(complement(AL133264.10:4611. .6907),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-001"
     mRNA            join(complement(AL133264.10:4611. .6907),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-001"
                     /product="chromosome 6 open reading frame 62"
                     /note="match: ESTs: AA084220 AA093151 AA168907 AA197971
                     AA200765 AA237701 AA292313 AA295166 AA402388 AA478637
                     AA591228 AA612170 AA638148 AA647423 AA647435 AA647742
                     AA672357 AA798123 AA798467 AA815944 AA833812 AA863774
                     AI053561 AI054330 AI122327 AI156738 AI191279 AI194277
                     AI284633 AI316639 AI390324 AI458673 AI688647 AI694146
                     AI694199 AI788240 AI802319 AI807413 AI831412 AI932363
                     AI935446 AI982078 AJ396992 AJ398211 AL529211 AL563388.1
                     AL601897.1 AU051650 AU125496 AU135776 AU137791 AU138464
                     AV656408 AV686324 AV698092 AV703346 AV708342 AV708873
                     AV709145 AV709526 AV739602 AV742411 AV743612 AV744587
                     AV745018 AW050065 AW063931 AW270998 AW300107 AW300121
                     AW301984 AW302318 AW302718 AW302841 AW302861 AW303029
                     AW355281 AW822881 AW859606 AW959243 BB067326 BB123906
                     BB609972 BB624305 BB628344 BB654156 BB823708 BE007675
                     BE049029 BE049182 BE120619 BE178692 BE178693 BE244058
                     BE246922 BE287246 BE289710 BE311050 BE369398 BE377181
                     BE386264 BE465032 BE484573 BE554360 BE571717 BE625094
                     BE766876 BE766886 BE770856 BE785982.1 BE813002 BE817780
                     BE829341 BE829344 BE829355 BE829359 BE893564 BE895461
                     BE896141 BE908143 BE913030 BE945891 BE991271 BE992722
                     BF015474 BF021751 BF022025 BF029248 BF037603 BF060909
                     BF064161 BF109576 BF111215 BF122334 BF137583 BF139381
                     BF140373 BF145038 BF148549 BF150174 BF165619 BF207643
                     BF326037 BF326041 BF326042 BF326045 BF334390 BF341149
                     BF376614 BF378829 BF380539 BF380550 BF465841 BF535383
                     BF578793 BF582307 BF748822 BF748938 BF748963 BF765346
                     BF766403 BF784199 BF789335 BF789696 BF885392 BF896088
                     BF940327 BF984391 BG030354 BG035327 BG088552 BG116557
                     BG148023 BG167018 BG168325 BG247693 BG252842 BG295272
                     BG295445 BG342500 BG610013 BG655426 BG667600 BG677384.1
                     BG678705 BG692520 BG717389 BG801183 BG802707 BG803992
                     BG863302 BG914583 BG917993 BG940215 BG943615 BG975946
                     BI080618 BI092275 BI106992 BI107793 BI108344 BI222784
                     BI258149 BI326830 BI651746 BI659175 BI666111 BI688353
                     BI689042 BI694119 BI696541 BI696575 BI696638 BI733610
                     BI753721 BI754669 BI851772 BI856689.1 BI872006 BI872428
                     BI903238 BI903250 BI960823 BI986171 BM016286 BM069622
                     BM069876 BM147734 BM237580 BM238719 BM238848 BM423047
                     BM423107 F11344 F20959 H17162 H98057 N56057 R25494 T11219
                     T51727 W47233
                     match: cDNAs: AB046018.1 AJ420475.1 AK022681.1 AL136632.1
                     BC002217.1 BC005537.1"
     gene            join(complement(AL133264.10:101. .211),5119. .5253,
                     7768. .8194)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-002"
     mRNA            join(complement(AL133264.10:101. .211),5119. .5253,
                     7768. .8194)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-002"
                     /product="chromosome 6 open reading frame 62"
                     /note="match: ESTs: BG772628 BG772628.1"
     gene            join(complement(AL133264.10:5931. .6069),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-003"
     mRNA            join(complement(AL133264.10:5931. .6069),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-003"
                     /product="chromosome 6 open reading frame 62"
                     /note="match: ESTs: BG342500.1 BG692520.1 BG803992.1"
     CDS             join(complement(AL133264.10:4611. .4739),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-003"
                     /standard_name="OTTHUMP00000016093"
                     /note="match: proteins: O95550 Q961D0 Q99K03 Q99LU8 Q9GZU0
                     Q9H1Z2 Q9N0B2 Q9QEJ3 Q9W384 Q9XWI7"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 62"
                     /protein_id="CAI20135.1"
                     /db_xref="GI:56202469"
                     /translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
                     KKSALFEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
                     DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEFHHGDYEKQFLHV
                     LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
                     IEDHLRPYMPE"
     CDS             join(complement(AL133264.10:5931. .5972),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-003"
                     /standard_name="OTTHUMP00000016094"
                     /codon_start=1
                     /product="chromosome 6 open reading frame 62"
                     /protein_id="CAI20136.1"
                     /db_xref="GI:56202470"
                     /db_xref="UniProt/TrEMBL:Q5JVZ2"
                     /translation="MSENLSDPVSPVVRKKKSALFEVSEVIPVMTNNYEENILKGVRD
                     SSYSLESSLELLQKDVVQLHAPRYQSMRRDVIGCTQEMDFILWPRNDIEKIVCLLFSR
                     WKESDEPFRPVQAKFEFHHGDYEKQFLHVLSRKDKTGIVVNNPNQSVFLFIDRQHLQT
                     PKNKATIFKLCSICLYLPQEQLTHWAVGTIEDHLRPYMPE"
     polyA_signal    9124. .9129
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-001"
     gene            complement(join(12316. .12571,15963. .16147,46686. .46959))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-001"
     mRNA            complement(join(12316. .12571,15963. .16147,46686. .46959))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-001"
                     /product="thioesterase superfamily member 2"
                     /note="match: ESTs: AA614605 AA928745 AI017897 AI052396
                     BF407527 BF475709 BF594161
                     match: cDNAs: AF155649.1 AF220186.1 AF274952.1 AK000508.1
                     AK002261.1 AK008624.1 BC000894.1"
     gene            complement(join(12316. .12571,15963. .16147,26337. .26628))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-002"
     mRNA            complement(join(12316. .12571,15963. .16147,26337. .26628))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-002"
                     /product="thioesterase superfamily member 2"
                     /note="match: ESTs: BG219890 BG256069"
     polyA_site      12318
     polyA_signal    12336. .12341
     CDS             complement(join(12415. .12571,15963. .16147,46686. .46766))
                     /gene="THEM2"
                     /locus_tag="RP1-30M3.2-001"
                     /standard_name="OTTHUMP00000016090"
                     /note="match: proteins: AAA28024 AAH18165 AAL48890 O95549
                     P34419 P87304 P93828 P95914 Q18187 Q9ASQ1 Q9NPJ3 Q9VZZ5
```

```
                     Q9VZZ6"
                     /codon_start=1
                     /product="thioesterase superfamily member 2"
                     /protein_id="CAD92509.1"

Alignment Scores:
Pred. No.:                6.84e-33        Length:         102200
Score:                    431.00          Matches:        78
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    98.7%           Mismatches:     0
Query Match:              36.2%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x HS30M3 (1-102200)

Qy         149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
               :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62958 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 63017

Qy         169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 188
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63018 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 63077

Qy         189 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 208
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63078 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACAC 63137

Qy         209 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63138 ATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 63194

RESULT 33
CR925830
LOCUS       CR925830              167847 bp    DNA     linear   PRI 08-JAN-2005
DEFINITION  Human DNA sequence from clone RP11-195J19 on chromosome 6, complete
            sequence.
ACCESSION   CR925830
VERSION     CR925830.3  GI:56798018
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 167847)
  AUTHORS   Palmer,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 24, 2004 this sequence version replaced gi:56542150.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-195J19 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES              Location/Qualifiers
     source           1. .167847
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="6"
                      /clone="RP11-195J19"
                      /clone_lib="RPCI-11.1"
ORIGIN

Alignment Scores:
Pred. No.:                1.22e-32        Length:         167847
Score:                    431.00          Matches:        78
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    98.7%           Mismatches:     0
Query Match:              36.2%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CR925830 (1-167847)

Qy         149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
               :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       43201 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 43260

Qy         169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 188
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       43261 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 43320

Qy         189 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 208
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       43321 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACAC 43380

Qy         209 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       43381 ATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 43437

RESULT 34
CR942205
LOCUS       CR942205              176819 bp    DNA     linear   PRI 29-APR-2005
DEFINITION  Human DNA sequence from clone RP11-948M1 on chromosome 6, complete
            sequence.
ACCESSION   CR942205
VERSION     CR942205.4  GI:62989869
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 176819)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:62551221.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6
RP11-948M1 is from the library RPCI-11.4 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES Location/Qualifiers
source 1. .176819
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-948M1"
/clone_lib="RPCI-11.4"

ORIGIN

Alignment Scores:

Pred. No.:	1.3e-32	Length:	176819
Score:	431.00	Matches:	78
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	98.7%	Mismatches:	0
Query Match:	36.2%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x CR942205 (1-176819)

```
Qy        149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37184 CAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 37243

Qy        169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37244 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACT 37303

Qy        189 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37304 GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACAC 37363

Qy        209 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37364 ATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 37420
```

RESULT 35
AC125756/c
LOCUS AC125756 245394 bp DNA linear HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-42H3, WORKING DRAFT SEQUENCE, 2 unordered pieces.
ACCESSION AC125756
VERSION AC125756.5 GI:30578987
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 245394)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 245394)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 245394)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On May 13, 2003 this sequence version replaced gi:23265926.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold,

```
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GJVU
                Center clone name: CH230-42H3
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 238380 bases at least Q40
                Consensus quality: 239660 bases at least Q30
                Consensus quality: 240760 bases at least Q20
                Estimated insert size: 255581; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
            * NOTE: This sequence may represent more than one clone.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   242009: contig of 242009 bp in length
            *   242010   242109: gap of unknown length
            *   242110   245394: contig of 3285 bp in length.
FEATURES             Location/Qualifiers
     source          1. .245394
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-42H3"
     misc_feature    1. .1025
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    1076. .2080
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    3723. .4536
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH348771"
     misc_feature    128656. .128715
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH260088"
     misc_feature    236171. .238412
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    239605. .242009
                     /note="wgs_end_extension
                     clone_end:Sp6"
     gap             242010. .242109
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:              3.97e-27        Length:         245394
Score:                  380.00          Matches:        102
Percent Similarity:     24.9%           Conservative:   17

Best Local Similarity:  21.4%           Mismatches:     14
Query Match:            31.9%           Indels:         344
DB:                     14              Gaps:           2

US-10-757-745-2_COPY_121_347 (1-227) x AC125756 (1-245394)

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              |||::::::   |||||||||   |||||||||||||||||||||||||||||||||||
Db     141402 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACAAGAAAA 141343

Qy         113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              |||:::|||||||||:::|||||||||||||   |||   :::|||||||||||||||||
Db     141342 CATTCTGCTGAACGAATAAATCAATTAAAAACTGTTTTTCAAAAAATGCAAGAGGCTACA 141283

Qy         133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr--- 151
              :::|||   |||||||||||||||||||||||||||||||:::|||||||:::
Db     141282 GATTCAACTACTGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTAAGTAAC 141223

Qy         151 ------------------------------------------------------------ 151

Db     141222 AGAATTGATTCAGTGTTTGTGGGGAGAGTATTGCTATTTAATTTGGATTCATTTTATTCA 141163

Qy         151 ------------------------------------------------------------ 151

Db     141162 AGGTATTATGGACTTGAAATACAAGGTAGGAAACACTGTTGTATCTTTTGTATAAAATGT 141103

Qy         151 ------------------------------------------------------------ 151

Db     141102 GCACAAAAGAGGCCTTTACACTATTAGGGGAAACCCTGCTATATTATATTAGTTAGATAA 141043

Qy         151 ------------------------------------------------------------ 151

Db     141042 ATGAGTATTAGAATAAGATAAGAATTATAAGAATTTTAATTAAAGTTAAAATAGGGTATT 140983

Qy         151 ------------------------------------------------------------ 151

Db     140982 TGTAGCATTCCTACGTATAAGAAGACCTAATCCTCTTATGACTTGAAATCAACTTTTTAA 140923

Qy         151 ------------------------------------------------------------ 151

Db     140922 AATACTTTATTGCTGGGTGCTGTGGCCCGTGCCTTTAATCATAGCACTTGGGTGTCAGAG 140863

Qy         151 ------------------------------------------------------------ 151

Db     140862 ACAGGCAGATCTCTGAATTCAAGGCCAGCCTGATATGCAGAGTGAATTGTAGTACAGCCA 140803

Qy         151 ------------------------------------------------------------ 151

Db     140802 GAATTACACAGAAACCCTGTCTTGAAAAACAATAGTCCTGAGAGACACATCCAGAGCATG 140743

Qy         151 ------------------------------------------------------------ 151

Db     140742 TCCAATACAGAGGTCAGTGCTAGCAGCAAACCACCGAACTGAGAACAGGACCCCCTTGGG 140683

Qy         151 ------------------------------------------------------------ 151

Db     140682 GGGAATTAGAGAAAGGACTGAAAGAGCTAAAGGGGCTTGCAACCCCATAAAAACAACAAT 140623

Qy         151 ------------------------------------------------------------ 151

Db     140622 GCCAACCAACCAGAGCTTCCAGGGACTAAACCACTACGGAAAGACTATATGTGGACTGAC 140563

Qy         151 ------------------------------------------------------------ 151

Db     140562 CCAGGGCTCCAACTGCATATGTAGCAGAGAATAGCCTTGTTGGGGCACCAGTGGAAGGGG 140503

Qy         151 ------------------------------------------------------------ 151

Db     140502 AAGCCCTTGGTCCTGCCAAGGTTAGACCCCCAGTGCAGGGGATTTGGGGGGGGGGCAATA 140443

Qy         151 ------------------------------------------------------------ 151

Db     140442 AGGGAAATGTATAGGGGAATACCAGTATGGGAGCAGGAGGAGGGAATGGGGACTTATGGA 140383
```

```
Qy        151 ------------------------------------------------------------ 151

Db     140382 CAGGAAACTGGGAAGGGGAATGACATTTGAAATGTAAGTAAAGAAATATATCTAATAAAC 140323

Qy        151 ------------------------------------------------------------ 151

Db     140322 AATTAAAAAAAAGAAAAACAACAAAAGTTGCACTTACAAGTGATAATATCAAACACCTTT 140263

Qy        151 ------------------------------------------------------------ 151

Db     140262 TAAGTTGAGGTGGGGAGTAAATATATAAACCTTACATAATATTTTTTCTCTTGCCAGGTT 140203

Qy        152 ---ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 170
                  :::|||||||||||||||||:::|||:::   |||   ||||||||||||||||||||
Db     140202 ATCAAATGTGGTGGTTTACCTGACAATGTTTTTGATGCCTGGGAATTTTTGGGCAAACCT 140143

Qy        171 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 190
              |||||||||:::|||||||||||||||:::   |||   ||||||   |||   ||||||
Db     140142 AAGCATTGCCGGTATACATGGGATACGAAAGCAAACGATAATCTCAGGATTCCTGCCGCT 140083

Qy        191 CysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210
              ||||||   |||||||||||||||||||||||      |||||||||||||||:::|||
Db     140082 TGTAAGCATCGTTTTGATCGAATATTCTTCAGA------GCAGAAGAGGGACACCTTATT 140029

Qy        211 ProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||:::|||||||||||||:::||||||||:::|||||||||||||||||
Db     140028 CCCCAAAGTTTAGACCTAATTGGGTTGGAAAGACTGGACTGTGGTAGATTT 139978
```

RESULT 36
BV203236

LOCUS	BV203236 201 bp DNA linear STS 10-JUN-2004
DEFINITION	sqnm211480 Human DNA (Sequenom) Homo sapiens STS genomic, sequence tagged site.
ACCESSION	BV203236
VERSION	BV203236.1 GI:48172559
KEYWORDS	STS.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 201)
AUTHORS	Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A.
TITLE	Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions
JOURNAL	Genome Res. (2004) In press
COMMENT	Contact: Andreas Braun Pharmaceuticals division Sequenom, Inc. 3595 John Hopkins Court, San Diego, CA 92121, USA Tel: 18582029018 Fax: 18582029020 Email: abraun@sequenom.com Primer A: No primer sequence submitted Primer B: No primer sequence submitted STS size: 201.

FEATURES	Location/Qualifiers
source	1. .201 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone_lib="Human DNA (Sequenom)"
STS	<1. .>201

ORIGIN

Alignment Scores:

Pred. No.:	8e-26	Length:	201
Score:	333.00	Matches:	66
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	27.9%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x BV203236 (1-201)

```
Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 62

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCys 101
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTTTGC 122

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTA 182

Qy        122 LysMetValLeuLysLys 127
              ||||||||||||||||||
Db        183 AAAATGGTTTTAAAGAAA 200
```

RESULT 37
AX351374

LOCUS	AX351374 176 bp DNA linear PAT 06-FEB-2002
DEFINITION	Sequence 121 from Patent WO0196390.
ACCESSION	AX351374
VERSION	AX351374.1 GI:18616721
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
TITLE	Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL	Patent: WO 0196390-A 121 20-DEC-2001; CORIXA CORPORATION (US)

FEATURES	Location/Qualifiers
source	1. .176 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	3.07e-22	Length:	176
Score:	298.00	Matches:	58
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	25.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x AX351374 (1-176)

```
Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 62

Qy         44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 122

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176
```

RESULT 38
AX339803/c

LOCUS	AX339803 179 bp DNA linear PAT 10-JAN-2002
DEFINITION	Sequence 50 from Patent WO0196388.
ACCESSION	AX339803

```
VERSION     AX339803.1  GI:18135784
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Harlocker,S.L. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196388-A 50 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .179
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                3.13e-22       Length:          179
Score:                    298.00         Matches:         58
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              25.0%          Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x AX339803 (1-179)

Qy          24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy          44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy          64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 39
BX957279
LOCUS       BX957279               198752 bp    DNA     linear   VRT 29-APR-2005
DEFINITION  Zebrafish DNA sequence from clone DKEY-218N20 in linkage group 19,
            complete sequence.
ACCESSION   BX957279
VERSION     BX957279.13  GI:62989971
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 198752)
  AUTHORS   Clark,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:61673710.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-218N20
            is from a Zebrafish BAC library
            VECTOR: pIndigoBAC-5.
FEATURES             Location/Qualifiers
     source          1. .198752
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-218N20"
                     /clone_lib="DanioKey"
ORIGIN

Alignment Scores:
Pred. No.:                7.55e-12       Length:          198752
Score:                    232.50         Matches:         67
Percent Similarity:       42.2%          Conservative:    38
Best Local Similarity:    26.9%          Mismatches:      77
Query Match:              19.5%          Indels:          67
DB:                       5              Gaps:            5

US-10-757-745-2_COPY_121_347 (1-227) x BX957279 (1-198752)

Qy          39 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 58
               |||:::|||   |||:::     :::     :::       :::   :::   |||
Db      147680 AGTTTTTTAGATAAACAAGAATGTTCACTTGCCATGTTTCTTAAATATCTGCAAACATAT 147739

Qy          59 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 78
                   |||||||||   |||||||||                              :::
Db      147740 TAATGAATTATTATGCTTTAAAAGAGTCGAAAA--------------------CTTACA 147778

Qy          79 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal---------------- 93
               :::   :::   |||:::::   :::||||||   :::::::::
Db      147779 TACAGCACCTTTAAGGTCATTATTAGCCTCCTTAAGCTATATATTTTTTCGATAATCCAC 147838

Qy          94 ------AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 111
                     |||                    ||||||:::|||              |||
Db      147839 AGAACAAACCACTGTTATACAATTACTTGCCTAATTACCCCAACTTGCCTAGTTACCCTA 147898

Qy         112 GlyHisAlaAlaGluArgMetAsn---------------------------------- 119
               :::       :::   :::|||
Db      147899 ATAAACATAGTTAAGCCTTTAAATGTAACTTTAAACTGAATACTAGTATCTTGAAAAATA 147958

Qy         119 ------------------------------------------------------------ 119

Db      147959 TCTAGTCAAATATTATTTATTGTCATCATGGTAAAATAAAAGATAAAATAAATCAGTTAT 148018

Qy         120 ----------------------GlnLeuLysMetValLeuLysLysMetGlnGluAla 131
                                    |||:::   |||||||||   :::
Db      148019 TAGAAATGAGTTATTATCACTATTATGTTAAGAAATGTGTTGAAAAATCTCTCCATTAAA 148078

Qy         132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsn---------LeuArgAspArg 148
```

```
Db      148079 CAGAAATTGGGGAATAATATATACAGGGGGGCTAATAATTGACTTCAACTGTATAATAAA 148138

Qy         149 GluValThrArgCys----------------------------GlyGlyLeuProAsn 158

Db      148139 GGAATGTTCGTGTGTGTATGTGTGTGTGTGTAGGTGGCCAATGTTGGAGGTTTGCCTGCA 148198

Qy         159 AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAsp 178

Db      148199 GGTGTGTGTGATGTGTGGGAACAGCTGGGCAAACAAGAACACTGCAGGTACACCTGGGAC 148258

Qy         179 ThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 198

Db      148259 ACCAAAGCTAACAGCAATAAGACTGTGCCTTACGTCAGCAGATGTCGCTTCGACCGCATA 148318

Qy         199 PhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGly 218

Db      148319 TTCCTGAGGTCTGCTAAAACCGCTCCGCCGGTCACTCCTGACCACATGGCCTTGATTGGA 148378

Qy         219 LeuGluLysLeuAspCysGlyArgPhe 227

Db      148379 ATGGAGAAACTGGACTGCGGTCGTTAT 148405
```

RESULT 40
BX511258
LOCUS BX511258 200076 bp DNA linear VRT 02-APR-2004
DEFINITION Zebrafish DNA sequence from clone CH211-81E5 in linkage group 19, complete sequence.
ACCESSION BX511258
VERSION BX511258.6 GI:42820869
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 200076)
AUTHORS Heath,P.
TITLE Direct Submission
JOURNAL Submitted (01-APR-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Feb 25, 2004 this sequence version replaced gi:42538803.
---------- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-81E5 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.
FEATURES Location/Qualifiers
source 1. .200076
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-81E5"
/clone_lib="CHORI-211"
ORIGIN

Alignment Scores:

Pred. No.:	9.67e-12	Length:	200076
Score:	231.50	Matches:	48
Percent Similarity:	55.8%	Conservative:	19
Best Local Similarity:	40.0%	Mismatches:	40
Query Match:	19.4%	Indels:	13
DB:	5	Gaps:	2

US-10-757-745-2_COPY_121_347 (1-227) x BX511258 (1-200076)

```
Qy         121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140

Db       90172 TTAAGAAATGTGTTGAAAAATCTCTCCATTAAACAGAAATTGGGGAATAATATATACAGG 90231

Qy         141 GlyAspThrAsn---------LeuArgAspArgGluValThrArgCys------------ 153

Db       90232 GGGGCTAATAATTGACTTCAACTGTATAATAAAGGAATGTTCGTGTGTGTATGTGTGTGT 90291

Qy         154 ------------------GlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167

Db       90292 GTGTAGGTGGCCAATGTTGGAGGTTTGCCTGCAGGTGTGTGTGATGTGTGGGAACAGCTG 90351

Qy         168 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 187

Db       90352 GGCAAACAAGAACACTGCAGGTACACCTGGGACACCAAAGCTAACAGCAATAAGACTGTG 90411

Qy         188 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 207

Db       90412 CCTTACGTCAGCAGATGTCGCTTCGACCGCATATTCCTGAGGTCTGCTAAAACCCGTCCA 90471

Qy         208 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227

Db       90472 GCGGTCACTCCTGACCACATGGCCTTGATTGGAATGGAGAAACTGGACTGCGGTCGTTAT 90531
```

RESULT 41
BD076937
LOCUS BD076937 483 bp DNA linear PAT 27-AUG-2002
DEFINITION 5' EST of secretory protein expressed in prostate.
ACCESSION BD076937
VERSION BD076937.1 GI:22622540
KEYWORDS JP 2001512013-A/184.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 483)
AUTHORS Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
TITLE 5' EST of secretory protein expressed in prostate
JOURNAL Patent: JP 2001512013-A 184 21-AUG-2001;
GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001512013-A/184
PD 21-AUG-2001
PF 31-JUL-1998 JP 2000505291
PR 01-AUG-1997 US 08/905144
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO PI LACROIX

```
          PC   C12N15/09,C07K14/47,C12P21/02,C12Q1/02,C12Q1/68,C12N15/00 CC
          blastn
          CC   identity 99
          CC   region 1. .310
          CC   id T26956
          CC   est
          CC   blastn
          CC   identity 98
          CC   region 1. .315
          CC   id T31666
          CC   est
          CC   blastn
          CC   identity 100
          CC   region 137. .267
          CC   id R14990
          CC   est
          CC   blastn
          CC   identity 100
          CC   region 63. .137
          CC   id R14990
          CC   est
          CC   blastn
          CC   identity 100
          CC   region 1. .50
          CC   id R14990
          CC   est
          CC   Von Heijne matrix
          CC   score 5.2
          CC   seq LCVEFASVASCDA/AV
          FH   Key             Location/Qualifiers
          FT   misc_feature    105. .414
          FT   misc_feature    45. .359
          FT   misc_feature    202. .332
          FT   misc_feature    127. .201
          FT   misc_feature    65. .114
          FT   sig_peptide     1. .120.
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              5.76e-13        Length:         483
Score:                  214.00          Matches:        41
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            18.0%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BD076937 (1-483)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 Leu 41
              |||
Db        481 CTA 483

RESULT 42
AP006486_2/c
WPCOMMENT
  Sequence split into 6 fragments  LOCUS AP006486 Accession AP006486
     Fragment Name      Begin     End
     AP006486_0             1    110000
     AP006486_1        100001    210000
     AP006486_2        200001    310000
     AP006486_3        300001    410000
     AP006486_4        400001    510000
     AP006486_5        500001    513790
Continuation (3 of 6) of AP006486 from base 200001 (AP006486 Cyanidioschyzon merolae st

Alignment Scores:
Pred. No.:              2.2e-09         Length:         110000
Score:                  206.00          Matches:        64
Percent Similarity:     42.8%           Conservative:   46
Best Local Similarity:  24.9%           Mismatches:     93
Query Match:            17.3%           Indels:         54
DB:                     14              Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x AP006486_2 (1-110000)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              ||||||          :::         :::             |||   |||:::
Db      35476 GAGCGGTTTACAGCTATTGCCGAAGAGGTCCTGGGTGCAAAACCCGGCGTGGTGCTTTTT 35417

Qy         31 GlnGluValIlePro---ProTyrTyrSerTyr-------------------------- 40
              |||||||||         |||:::   |||:::
Db      35416 CAAGAAGTCTCTGATTTACCGTTCTCCTCTTGGCAGGGCGAGAAAGAGCAGTCGACGATT 35357

Qy         41 ------------------------------------LeuLysLysArgSerSerAsn 47
                                                   |||::::::     :::
Db      35356 ACGTCCGCTCGTCAGCATGATTCTTGTTACAAGGTCCTCTTGCGTCGTCTTGAGGCTCTA 35297

Qy         48 TyrGluIleIleThrGly-----------HisGluGluGlyTyrPheThrAlaIleMet 63
              ||||||   |||             |||:::   ||||||
Db      35296 TATGAGCCAATTGTTGATCTCGGCTTACCCTCAGAGCAGGCATACTTCGTGATGGCGTTC 35237

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 83
              :::   :::|||:::   ::::::          |||          |||   ::::::
Db      35236 GTTGACCGAAGCAAGTATCGAGTTCACGGATACGAATACACTGAATTTGAAAACTCTTCA 35177

Qy         84 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCys------ 101
              |||   ||||||          :::   :::          :::::::|||
Db      35176 CAGGGAAGAGGTCTGCTCGCA------ATACAGCTCGAGCCCAGAGATATGTGCGGAGAT 35123

Qy        102 -----------LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                        ::::::|||||||||||||||||||          :::       |||
Db      35122 TTCAAGTTTCGAGTGCTTACATCACACCTCGAATCCGGAACTGCCTGTTCGGAAGTCCGA 35063

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
              :::   |||::::::   :::       :::|||   |||              |||
Db      35062 GTTCAGCAGATGAGGCAAATCGCGTATGAGATGCTCGAGGGTGATACCCCGGCA------ 35009

Qy        138 IlePheAlaGlyAspThrAsnLeuArgAspArgGlu----------------------- 149
              ||||||||||||||||||   |||::::::|||
Db      35008 ATATTTGCAGGAGATACCAATTTTCGAGAGAAAGAAAAAGCAGGATTTGAGAAGGAAATC 34949

Qy        150 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLys 169
                       |||                 :::   |||::::::      :::|||:::
Db      34948 GCTAAAGCGATCGGCTCAGAAGAGATGCGACTTTTTGACCTATTTGTTGAGGTAGGGGAG 34889

Qy        170 ProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn-----------Leu 185
              |||   |||   ::::::|||||||||||||:::   |||:::|||           |||
Db      34888 CCTTCACATTTGCGTTTCACCTGGGACACACGCACAAACACGAACAAAGTCAAGGAGTTG 34829

Qy        186 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 205
                          :::   ||||||||:::::   |||          :::
Db      34828 CCACCTCGGTTCAGGCCGCGGGCTCGTTTCGACCAAGTGCTTTTCACTAGGCCTTTTCAA 34769

Qy        206 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 222
                             ::::::   |||||||||   |||:::|||
Db      34768 --------------GCGAACGCCATGCGGCTTTTGGGTCGAGAGCGACTC 34733

RESULT 43
BT000599
LOCUS       BT000599               1326 bp    mRNA    linear   PLN 25-SEP-2002
DEFINITION  Arabidopsis thaliana unknown protein  (At1g11800/F25C20_3) mRNA,
            complete cds.
ACCESSION   BT000599
```

```
VERSION     BT000599.1  GI:23308396
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1326)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hayashizaki,Y.,
            Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
            Nguyen,M., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
            Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1326)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hayashizaki,Y.,
            Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
            Nguyen,M., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
            Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-SEP-2002) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
            Chen,H., Kim,C.J., Shinn,P., Banh,J. Bowser,L., Chan,M.M.,
            Chang,E., Dale,J.M., Deng, J.M., Goldsmith,A.D., Jones,T.,
            Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamada,K., Yamamura, Y., Yu,G., Davis,R.W.,
            Theologis,A., and Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1326
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="U22176"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51"
     CDS             1. .1326
                     /note="unknown protein"
                     /codon_start=1
                     /product="At1g11800/F25C20_3"
                     /protein_id="AAN18168.1"
                     /db_xref="GI:23308397"
                     /translation="MSTTIRFLRIVTSRAMSSSWSCNKCTFLNSASQKLNCMICLAPV
                     SLPSLSPPPPSLSISANDEAKWACKACTPLNTYKNSICDVCGTRSPTSSLLGFQDLTD
                     SGLESNDADSSVGSVFFPLRRCIKRKAMDDDVVEVDGASVVCSESQGVMKKNKEIETK
                     GVASDSGTPLTCLKILSYNVWFREDLELNLRMRAIGHLIQLHSPHLICFQEVTPEIYD
                     IFRKSNWWKAYSCSVSVDVAVSRGYYCMLLSKLGVKSFSSKSFGNSIMGRELSIAEVE
                     VPGRKPLVFATSHLESPCPGPPKWDQMFSRERVEQAKEAIEILRPNANVIFGGDMNWC
                     DKLDGKFPLPDKWVDVWEVLKPGDLGFTYDTKANPMLSGNRALQKRLDRILCRLDDYK
                     LGGIEMVGKEAIPGLSYVKEKKVRGDIKKLELPVLPSDHFGLLVTLSPK"

ORIGIN

Alignment Scores:
Pred. No.:                4.24e-11        Length:         1326
Score:                    201.00          Matches:        69
Percent Similarity:       46.3%           Conservative:   32
Best Local Similarity:    31.7%           Mismatches:     85
Query Match:              16.9%           Indels:         32
DB:                       15              Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x BT000599 (1-1326)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
              :::   |||:::|||||||           |||   |||   :::       :::
Db        550 GAAGACCTTGAGTTGAAT---------CTAAGGATGAGAGCGATTGGTCATCTTATTCAG 600

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
              |||:::||||||   :::|||       |||||||||   |||       |||
Db        601 CTTCATTCTCCTCATCTCATATGTTTTCAGGAAGTCACTCCTGAGATATATGATATTTTC 660

Qy         42 LysLysArg-----------------SerSerAsnTyrGluIleIleThrGlyHisGlu 55
              :::|||                   |||   :::       :::
Db        661 CGCAAGTCGAATTGGTGGAAAGCGTATTCTTGTTCTGTCTCAGTTGATGTGGCAGTTTCG 720

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||:::       ::::::         |||::::::   :::||||||
Db        721 AGAGGCTACTACTGCATGCTGTTG---------AGCAAGTTGGGAGTGAAATCTTTCAGC 771

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
                    |||   ::::::   |||   |||   |||           |||   |||
Db        772 AGCAAAAGCTTTGGGAACTCGATAATGGGAAGAGAACTATCTATCGCAGAAGTTGAAGTT 831

Qy         96 SerGlyAsnGlu---LeuCysLeuMetThrSerHisLeuGluSer---------------- 109
              |||   :::   |||               |||||||||||||||||
Db        832 CCAGGGAGAAAGCCACTTGTTTTTGCAACTAGCCATCTTGAGAGTCCCTGTCCAGGTCCT 891

Qy        110 ---------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
                      :::       :::   ||||||:::   |||   |||       ::::::
Db        892 CCAAAATGGGACCAAATGTTCAGTCGAGAACGCGTTGAACAAGCCAAGGAGGCAATTGAG 951

Qy        127 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 146
              ::::::           :::|||   |||||||||   ||||||   |||
Db        952 ATTCTCAGA---------CCTAATGCCAATGTGATATTCGGGGGAGACATGAACTGGTGC 1002

Qy        147 AspArgGluValThrArgCysGlyGly---------LeuProAsnAsnIleValAspVal 163
              |||             :::       |||         ||||||:::       ||||||||||
Db       1003 GAC------------AAGCTAGATGGGAAGTTTCCTTTGCCAGACAAATGGGTGGATGTT 1050

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
              ||||||   |||   ||||||           :::|||:::|||||||:::   |||
Db       1051 TGGGAGGTTTTG---AAACCAGGTGATTTGGGTTTACATATGACACAAAAGCAAACCCA 1107

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
              |||             |||   :::   |||   |||||||||       |||
Db       1108 ATGTTGTCCGGTAACCGGGCATTGCAGAAACGGTTAGACCGGATTTTGTGTCGT 1161

RESULT 44
AY085643
LOCUS       AY085643               1557 bp    mRNA    linear   PLN 14-APR-2003
DEFINITION  Arabidopsis thaliana clone 16491 mRNA, complete sequence.
ACCESSION   AY085643
VERSION     AY085643.1  GI:21404353
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1557)
  AUTHORS   Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
            Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
  TITLE     Full-length messenger RNA sequences greatly improve genome
            annotation
```

```
  JOURNAL   Genome Biol. 3 (6), RESEARCH0029 (2002)
   PUBMED   12093376
REFERENCE   2  (bases 1 to 1557)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Full-Length cDNA from Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1557)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
            Malibu, CA 90265, USA
COMMENT     This clone sequence is one of 5,000 Ceres full-length cDNAs made
            available to TIGR and Genbank. The following quality assessment of
            this set was done by comparison with known proteins: two percent of
            the clones are estimated to be 5'-truncated; less than one percent
            are 3'-truncated; approximately two percent represent alternative
            splice variants, including unspliced introns and spliced exons; one
            percent may contain premature stop codons; five percent may have
            frame shifts in a coding region. A sequence is considered to be
            5'-truncated if it lacks the translation initiation start (ATG). A
            sequence is considered to be 3'-truncated if it lacks the
            C-terminal end of the encoded protein. Please note that these cDNA
            sequences are derived from the Ws or LAer ecotypes and therefore
            may contain polymorphisms when compared to sequences from Col-0.
            Genset carried out the library production and sequencing of the
            full-length clones. Ceres, Inc. carried out the clustering of the
            5' sequences, selection of clones, and sequence assembly.
FEATURES             Location/Qualifiers
     source          1. .1557
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="16491"
     CDS             35. .1315
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAM62864.1"
                     /db_xref="GI:21553771"
                     /translation="MSSSWSCNKCTFLNSASQKLNCMICLAPVSLPSLSPPPPSLSIS
                     ANDEAKWACKACTFLNTYKNSICDVCGTRSPTSSLLGFQDLTDSGLESNDADSSVGSV
                     FFPLRRCIKRKAMDDDVVEVDGASVVCSASQGVMKKNKEIETKGVASDSGTPLTCLKI
                     LSYNVWFREDLELNLRMRAIGHLIQLHSPHLICFQEVTPEIYDIFRKSNWWKAYSCSV
                     SVDVAVSRGYYCMLLSKLGVKSFSSKSFGNSIMGRELSIAEVEVPGRKPLVFATSHLE
                     SPCPGPPKWDQMFSRERVEQAKEAIEILRPNANVIFGGDMNWCDKLDGKFPLPDKWVD
                     VWEVLKPGDLGFTYDTKANPMLSGNRALQKRLDRILCRLDDYKLGGIEMVGKEAIPGL
                     SYVKEKKVRGDIKKLELPVLPSDHFGLLVTLSPK"
ORIGIN

Alignment Scores:
Pred. No.:              5.12e-11        Length:         1557
Score:                  201.00          Matches:        69
Percent Similarity:     46.3%           Conservative:   32
Best Local Similarity:  31.7%           Mismatches:     85
Query Match:            16.9%           Indels:         32
DB:                     15              Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x AY085643 (1-1557)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
              ::   |||:::||||||          |||   |||   :::         :::
Db        539 GAAGACCTTGAGTTGAAT---------CTAAGGATGAGAGCGATTGGTCATCTTATTCAG 589

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
              |||:::|||||||   :::|||       |||||||||   |||       |||
Db        590 CTTCATTCTCCTCATCTCATATGTTTTCAGGAAGTCACTCCTGAGATATATGATATTTTC 649

Qy         42 LysLysArg-----------------SerSerAsnTyrGluIleIleThrGlyHisGlu 55
              :::|||                  |||   :::       :::
Db        650 CGCAAGTCGAATTGGTGGAAAGCGTATTCTTGTTCTGTCTCAGTTGATGTGGCAGTTTCG 709

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
                ||||||:::       ::::::          |||:::::::   :::||||||
Db        710 AGAGGCTACTACTGCATGCTGTTG---------AGCAAGTTGGGAGTGAAATCTTTCAGC 760

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
                  |||   ::::::   |||   |||   |||                 |||   |||
Db        761 AGCAAAAGCTTTGGGAACTCGATAATGGGAAGAGAACTATCTATCGCAGAAGTTGAAGTT 820

Qy         96 SerGlyAsnGlu---LeuCysLeuMetThrSerHisLeuGluSer--------------- 109
              |||   :::   |||          ||||||||||||||||||
Db        821 CCAGGGAGAAAGCCACTTGTTTTTGCAACTAGCCATCTTGAGAGTCCCTGTCCAGGTCCT 880

Qy        110 ---------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
                      :::     :::   ||||||:::   |||   |||     ::::::
Db        881 CCAAAATGGGACCAAATGTTCAGTCGAGAACGCGTTGAACAAGCCAAGGAGGCAATTGAG 940

Qy        127 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 146
              ::::::          :::|||   |||||||||   ||||||   |||
Db        941 ATTCTCAGA---------CCTAATGCCAATGTGATATTCGGGGGAGACATGAACTGGTGC 991

Qy        147 AspArgGluValThrArgCysGlyGly---------LeuProAsnAsnIleValAspVal 163
              |||             :::     |||         ||||||:::      |||||||||
Db        992 GAC------------AAGCTAGATGGGAAGTTTCCTTTGCCAGACAAATGGGTGGATGTT 1039

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
              ||||||   |||   ||||||           :::|||:::|||||||:::   |||
Db       1040 TGGGAGGTTTTG---AAACCAGGTGATTTGGGTTTTACATATGACACAAAAGCAAACCCA 1096

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                 |||          |||   :::   |||   |||||||||       |||
Db       1097 ATGTTGTCCGGTAACCGGGCATTGCAGAAACGGTTAGACCGGATTTTGTGTCGT 1150

RESULT 45
AY128286
LOCUS       AY128286               1582 bp    mRNA    linear   PLN 07-AUG-2002
DEFINITION  Arabidopsis thaliana At1g11800/F25C20_3 mRNA, complete cds.
ACCESSION   AY128286
VERSION     AY128286.1  GI:22135815
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1582)
  AUTHORS   Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Banh,J., Bowser,L.,
            Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
            Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
            and Ecker,J.R.
  TITLE     Arabidopsis cDNA clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1582)
  AUTHORS   Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Banh,J., Bowser,L.,
            Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
            Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
            and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUN-2002) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.
```

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Banh,J. Bowser,L., Chang,E., Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura, Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

```
FEATURES             Location/Qualifiers
     source          1. .1582
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="RAFL09-10-B22(R22176)"
                     /ecotype="Columbia"
     CDS             28. .1353
                     /note="unknown protein"
                     /codon_start=1
                     /product="At1g11800/F25C20_3"
                     /protein_id="AAM91094.1"
                     /db_xref="GI:22135816"
                     /translation="MSTTIRFLRIVTSRAMSSSWSCNKCTFLNSASQKLNCMICLAPV
                     SLPSLSPPPPSLSISANDEAKWACKACTFLNTYKNSICDVCGTRSPTSSLLGFQDLTD
                     SGLESNDADSSVGSVFFPLRRCIKRKAMDDDVVEVDGASVVCSESQGVMKKNKEIETK
                     GVASDSGTPLTCLKILSYNVWFREDLELNLRMRAIGHLIQLHSPHLICFQEVTPEIYD
                     IFRKSNWWKAYSCSVSVDVAVSRGYYCMLLSKLGVKSFSSKSFGNSIMGRELSIAEVE
                     VPGRKPLVFATSHLESPCPGPPKWDQMFSRERVEQAKEAIEILRPNANVIFGGDMNWC
                     DKLDGKFPLPDKWVDVWEVLKPGDLGFTYDTKANPMLSGNRALQKRLDRILCRLDDYK
                     LGGIEMVGKEAIPGLSYVKEKKVRGDIKKLELPVLPSDHFGLLVTLSPK"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              5.21e-11        Length:         1582
Score:                  201.00          Matches:        69
Percent Similarity:     46.3%           Conservative:   32
Best Local Similarity:  31.7%           Mismatches:     85
Query Match:            16.9%           Indels:         32
DB:                     15              Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x AY128286 (1-1582)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
              :::   |||:::||||||           |||   |||   :::         :::
Db        577 GAAGACCTTGAGTTGAAT---------CTAAGGATGAGAGCGATTGGTCATCTTATTCAG 627

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
              |||:::||||||   :::|||      |||||||||   |||       |||
Db        628 CTTCATTCTCCTCATCTCATATGTTTTCAGGAAGTCACTCCTGAGATATATGATATTTTC 687

Qy         42 LysLysArg-----------------SerSerAsnTyrGluIleIleThrGlyHisGlu 55
              :::|||                    |||   :::         :::
Db        688 CGCAAGTCGAATTGGTGGAAAGCGTATTCTTGTTCTGTCTCAGTTGATGTGGCAGTTTCG 747

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
                ||||||:::      ::::::         |||::::::   :::||||||
Db        748 AGAGGCTACTACTGCATGCTGTTG---------AGCAAGTTGGGAGTGAAATCTTTCAGC 798

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
                   |||  ::::::  |||  |||   |||             |||   |||
Db        799 AGCAAAAGCTTTGGGAACTCGATAATGGGAAGAGAACTATCTATCGCAGAAGTTGAAGTT 858

Qy         96 SerGlyAsnGlu---LeuCysLeuMetThrSerHisLeuGluSer--------------- 109
              |||   :::   |||         ||||||||||||||||||
Db        859 CCAGGGAGAAAGCCACTTGTTTTTGCAACTAGCCATCTTGAGAGTCCCTGTCCAGGTCCT 918

Qy        110 ---------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
                      :::   :::   ||||||:::   |||   |||      ::::::
Db        919 CCAAAATGGGACCAAATGTTCAGTCGAGAACGCGTTGAACAAGCCAAGGAGGCAATTGAG 978
```

```
Qy        127 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 146
                  ::::::          :::|||   |||||||||   ||||||    |||
Db        979 ATTCTCAGA---------CCTAATGCCAATGTGATATTCGGGGGAGACATGAACTGGTGC 1029

Qy        147 AspArgGluValThrArgCysGlyGly---------LeuProAsnAsnIleValAspVal 163
              |||             :::      |||         ||||||:::      |||||||||
Db       1030 GAC------------AAGCTAGATGGGAAGTTTCCTTTGCCAGACAAATGGGTGGATGTT 1077

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
              ||||||   |||   ||||||           :::|||:::|||||:::   |||
Db       1078 TGGGAGGTTTTG---AAACCAGGTGATTTGGGTTTTACATATGACACAAAAGCAAACCCA 1134

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                 |||             |||  :::   |||   |||||||||      |||
Db       1135 ATGTTGTCCGGTAACCGGGCATTGCAGAAACGGTTAGACCGGATTTTGTGTCGT 1188

RESULT 46
CQ069444
LOCUS       CQ069444                 444 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 5244 from Patent WO0157278.
ACCESSION   CQ069444
VERSION     CQ069444.1  GI:41039313
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 5244 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL031775.1
                     EXPRESSED IN HELA, SIGNAL = 1.9"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              1.01e-09        Length:         444
Score:                  182.50          Matches:        37
Percent Similarity:     74.5%           Conservative:   4
Best Local Similarity:  67.3%           Mismatches:     11
Query Match:            15.3%           Indels:         3
DB:                     6               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x CQ069444 (1-444)

Qy         15 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 31
              |||:::|||       |||               ||||||||||||||||||||||||||
Db        221 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 280

Qy         32 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 340

Qy         52 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 66
              ||||||:::      |||   |||   :::       :::|||
Db        341 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 385

RESULT 47
CQ078617
LOCUS       CQ078617                 691 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 14417 from Patent WO0157278.
ACCESSION   CQ078617
VERSION     CQ078617.1  GI:41048486
```

```
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 14417 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL031775.1
                     EXPRESSED IN HELA, SIGNAL = 1.9
                     SWISSPROT HIT: P18004, EVALUE 7.00e-01
                     NT HIT: AF201687.1, EVALUE 3.00e-45
                     EST_HUMAN HIT: AV717253.1, EVALUE 5.00e-45"
ORIGIN

Alignment Scores:
Pred. No.:                1.69e-09        Length:         691
Score:                    182.50          Matches:        37
Percent Similarity:       74.5%           Conservative:   4
Best Local Similarity:    67.3%           Mismatches:     11
Query Match:              15.3%           Indels:         3
DB:                       6               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x CQ078617 (1-691)

Qy          15 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 31
               |||:::|||       |||                |||||||||||||||||||||||||
Db          83 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 142

Qy          32 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 51
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         143 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 202

Qy          52 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 66
               ||||||:::          |||   |||   :::      :::|||
Db         203 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 247

RESULT 48
CEY63D3A/c
LOCUS       CEY63D3A                33899 bp    DNA     linear   INV 09-AUG-2005
DEFINITION  Caenorhabditis elegans YAC Y63D3A, complete sequence.
ACCESSION   AL032652
VERSION     AL032652.1  GI:3810723
KEYWORDS    HTG.
SOURCE      Caenorhabditis elegans
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 33899)
  AUTHORS   .
  CONSRTM   C. elegans Sequencing Consortium
  TITLE     Genome sequence of the nematode C. elegans: a platform for
            investigating biology
  JOURNAL   Science 282 (5396), 2012-2018 (1998)
   PUBMED   9851916
REFERENCE   2  (bases 1 to 33899)
  AUTHORS   White,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-OCT-1998) Nematode Sequencing Project, Sanger
            Institute, Hinxton, Cambridge CB10 1SA, England and Department of
            Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
            worm@sanger.ac.uk
COMMENT     On May 14, 2001 this sequence version replaced gi:3646980.
            Coding sequences below are predicted from computer analysis, using
            predictions from Genefinder (P. Green, U. Washington), and other
            available information.
            Current sequence finishing criteria for the C. elegans genome
            sequencing consortium are that all bases are either sequenced
            unambiguously on both strands, or on a single strand with both a
            dye primer and dye terminator reaction, from distinct subclones.
            Exceptions are indicated by an explicit note.
            IMPORTANT:  This sequence is NOT necessarily the entire insert of
            the specified clone.  It may be shorter because we only sequence
            overlapping sections once, or longer because we arrange for a small
            overlap between neighbouring submissions.
            For a graphical representation of this sequence and its analysis
            see:- http://www.wormbase.org/perl/ace/elegans/seq/sequence?
            name=Y63D3A;class=Sequence
            IMPORTANT: This sequence is not the entire insert of clone Y63D3A.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we arrange for a small overlap between
            neighbouring submissions.
            The true left end of clone K03D10 is at 33796 in this sequence. The
            true right end of clone F47G4 is at 103 in this sequence. The true
            right end of clone Y71A12 is at 113 in this sequence. The start of
            this sequence (1. .103) overlaps with the end of sequence Z99171.
            The end of this sequence (33796. .33899) overlaps with the start of
            sequence Z81561.
FEATURES             Location/Qualifiers
     source          1. .33899
                     /organism="Caenorhabditis elegans"
                     /mol_type="genomic DNA"
                     /strain="Bristol N2"
                     /db_xref="taxon:6239"
                     /chromosome="I"
                     /clone="Y63D3A"
     gene            complement(join(49. .345,612. .1065,1117. .1199))
                     /locus_tag="Y63D3A.9"
     CDS             complement(join(49. .345,612. .1065,1117. .1199))
                     /locus_tag="Y63D3A.9"
                     /standard_name="Y63D3A.9"
                     /note="contains similarity to Pfam domain PF00646 (F-box
                     domain)"
                     /codon_start=1
                     /product="Hypothetical protein Y63D3A.9"
                     /protein_id="CAA21711.1"
                     /db_xref="GI:3881237"
                     /db_xref="InterPro:IPR001810"
                     /db_xref="UniProt/TrEMBL:Q9XWG0"
                     /translation="MSSPSFPILRLPSKALRTTLQNFDFIDLLSFSFVSPTPTKSLNM
                     KVTDFIIIKNYGIYIVARCCGIALRPITSQIFVDGQPQERQWEIQSITVKKCFENFLA
                     AFNHFEVEYICVRGEDDISYNLEEIKVVKNLEIFSADCGIFGQLYRVANKVTLSQRSL
                     LHEKPRDFFTENMDSLQLEKAVIYRPFTLSDLNLFLQHWRKGSNPRLEEIYIGLKEDN
                     LVDFKKTLLKGIEYWKYGIFRVYNNSPSIEGLEIERTDGIRAIIFVNPMHRSIGMGVL
                     Q"
     gene            join(1524. .1606,1651. .2086,4038. .4352)
                     /locus_tag="Y63D3A.10"
     CDS             join(1524. .1606,1651. .2086,4038. .4352)
                     /locus_tag="Y63D3A.10"
                     /standard_name="Y63D3A.10"
                     /note="contains similarity to Pfam domain PF00646 (F-box
                     domain)"
                     /codon_start=1
                     /product="Hypothetical protein Y63D3A.10"
                     /protein_id="CAA21706.1"
                     /db_xref="GI:3881232"
                     /db_xref="InterPro:IPR001810"
                     /db_xref="UniProt/TrEMBL:Q9XWG4"
                     /translation="MAHVAFNLLHLPAKTLRNVFQLMSHSELFIFSFASICAKRLVKS
                     MNFEMLELDLVYSRPCKLKISFGEHEEFIIEFDEFGTLSKQEPSDAFVFLIAYKLDIR
                     DCINHVLEVYKISKLHTLKVIGAINNPQNLLKIIGKLDLKVLWIFFNFSENSTTISVY
                     REILNHLGARCDVRSLRRGFYDKEPQKSFFLLRLSLSYALSQEPHSRPTATLAARRLK
                     RRQDTQPLDATTSVLMLRSCRTVGRGTSLPTHPPFSFVIAPHYYSCTCSDIQGEPDKL
                     Q"
     gene            complement(join(2689. .3289,3365. .3484,3604. .3686))
                     /locus_tag="Y63D3A.11"
```

```
CDS             complement(join(2689. .3289,3365. .3484,3604. .3686))
                /locus_tag="Y63D3A.11"
                /standard_name="Y63D3A.11"
                /note="contains similarity to Pfam domain PF00646 (F-box
                domain)"
                /codon_start=1
                /product="Hypothetical protein Y63D3A.11"
                /protein_id="CAB76748.1"
                /db_xref="GI:7160769"
                /db_xref="InterPro:IPR001810"
                /db_xref="UniProt/TrEMBL:Q9NA68"
                /translation="MSSPSLPILRLPSKALRTTLQNFEYLDLLSFSLVSPTPVKSLNM
                KVTDFILISNDEIVIFGRATMFFRKSFENFLIAFNHFEVEHVYVNIENDIPYNLEEIK
                VVKDLFVRSADRGIFSQFYRTANKISLSNRSLHNDKSRNFFTENIDSLWLDSSFHLEF
                NDLLAINAKKAVIYRPFTLSDLNLFLQHWRKGSNPRLEEIYIGLEEDNLVDFNETLLK
                GIVHWKSCLSKVYNNSPSIEGLEIERTDGTRAIIFVNPMKRSFEMGVLQ"
gene            join(5095. .5177,5974. .6859)
                /locus_tag="Y63D3A.3"
CDS             join(5095. .5177,5974. .6859)
                /locus_tag="Y63D3A.3"
                /standard_name="Y63D3A.3"
                /note="contains similarity to Pfam domain PF00646 (F-box
                domain)"
                /codon_start=1
                /product="Hypothetical protein Y63D3A.3"
                /protein_id="CAB63397.2"
                /db_xref="GI:7160770"
                /db_xref="InterPro:IPR001810"
                /db_xref="UniProt/TrEMBL:Q9NA67"
                /translation="MPNYPFPILWLPPKPLGHLLRNMDFRNLIIFSLTSKPAKWLTEL
                LNMDAPMISVHVDHQINIIVSSVKSVTYHASNGSLRESRGKIDRQINFPGSLKQFVEH
                LLHIFKEHKIDDLEIMCNSVYQYKPDIQALKGIKMRTLKLRYQESILKMFGTSADRIE
                FSTIKSAELENSKDILIQNFDYLKFDFEGSLTADNLIGINGKVVNIWAADIEMKDVNI
                FLKHWLKGSNPRLEKLILKTYRLDTDGNSVEDSVADLFRDLDCESIAESQFEILRDDE
                SIAKISILEHGYWYNKIHCYFSLNEAMNLYTTYFWDIVHLQEKILH"
gene            complement(join(7747. .8617,13260. .13342))
                /locus_tag="Y63D3A.2"
CDS             complement(join(7747. .8617,13260. .13342))
                /locus_tag="Y63D3A.2"
                /standard_name="Y63D3A.2"
                /note="contains similarity to Pfam domain PF00646 (F-box
                domain)"
                /codon_start=1
                /product="Hypothetical protein Y63D3A.2"
                /protein_id="CAA21709.1"
                /db_xref="GI:3881235"
                /db_xref="InterPro:IPR001810"
                /db_xref="UniProt/TrEMBL:Q9XWG1"
                /translation="MPALPFPILSLPPKSRRIALKCMDTLDLINISFTHPLAKRLAIE
                LSVGIEVGSLRLLVRNGFKVFFSFYAPPHNIIADISYLSADGETFDLIKDEDVLWINS
                KIKLRDLLVDLLDVLHLRKLGEVSVHTSWDLILDELRGIEIHQLIVGESITRQFKLLK
                LQAEELELNVPLTQADVICHNLHRIALKSPGNIQCEDLIGMNMMHIEVFMMFDHSSLT
                HREINLFLLHWVKGSNPRMRYYRSPLQEPLQFEEQFVAQILKGIPHRAITPGELEKLR
                VQFRGKTFQFRGYHIERNEGDAKAMISVDPWSPIVEVFVWF"
gene            complement(join(14704. .14833,15019. .15302,16493. .16724,
                16858. .16952,17625. .17845,17895. .17969,18033. .18084))
                /locus_tag="Y63D3A.4"
CDS             complement(join(14704. .14833,15019. .15302,16493. .16724,
                16858. .16952,17625. .17845,17895. .17969,18033. .18084))
                /locus_tag="Y63D3A.4"
                /standard_name="Y63D3A.4"
                /note="contains similarity to Pfam domain PF03372
                (Endonuclease/Exonuclease/phosphatase family)"
                /codon_start=1
                /product="Hypothetical protein Y63D3A.4"
                /protein_id="CAA21707.1"
                /db_xref="GI:3881233"
                /db_xref="InterPro:IPR005135"
                /db_xref="UniProt/TrEMBL:Q9XWG3"
                /translation="MSNSDDEIQEIEAKRQKMSQEDSEVEIEILDEPEQGKLKNSSMS
                DEQKLHEFAIITATDEAFAQSILQDVDWDLKKALDVFYGSEAFAEARSAAVMGASSSM
                ASSGAAVMTAEDLKGFEVSVMSWNIDGLDGRSLLTRMKAVAHIVKNVNPDILFLQEVV
                DRDLAPIDKLQSLYKIYYSNKGCQYYTAILVSKMFDVEKHDVIHFQNSGMYRTLQILE
                GSIGGLKVFLLNTHLESTREHRPQRCAQFGFCMDKVREIIAQNPGALVFFGGDLNLRD
                EEVSRVPDGVKDAWEAAGSDNKTKFTWDTFKNDNKQGFHGAKMRFDRLYWSGPLDKVK
                FTLEGRQRIRSCLCFPSDHWAINATFFA"
gene            join(19048. .19228,19281. .19426,20077. .21210)
                /gene="tfg-1"
                /locus_tag="Y63D3A.5"
CDS             join(19048. .19228,19281. .19426,20077. .21210)
                /gene="tfg-1"
                /locus_tag="Y63D3A.5"
                /standard_name="Y63D3A.5"
                /note="C. elegans TFG-1 protein; contains similarity to
                Pfam domain PF00564 (PB1 domain)"
                /codon_start=1
                /product="Hypothetical protein Y63D3A.5"
                /protein_id="CAB63398.1"
                /db_xref="GI:6580329"
                /db_xref="InterPro:IPR000270"
                /db_xref="UniProt/TrEMBL:Q9U1W1"
                /translation="MVHSNGAITSTILKARHADVVRKTSLHHANDLTLIDLVLNVQRL
                LALPSDANFVLKYKDEEGDLVTLAEDSDLLLALHTSGATLDVTVVVDSRAREAVHDVQ
                KQVEQIKLDVGKLLGALSALDVAQIAEQSNTSVANLSAPKQSHHDNIVFQKSPEAAPP
                SPVPSEKAELPATIQPSVHEQFNHRPAHVEEEIPLENHYAPPPHQQIPDDLNTSFSSQ
                PPPPIEQFGAIPPPNATIPSFPTSNAASPPVQEFAPPPPQQQQQQFQAPPPPMASHSS
                ISSTPVQQQGFAPPQQFGGPPPSGPPSEYGGYAPPQQQQQQFGAPPPQGAPQQGFGAP
                PQGPPQGGPPQGSFGAPPPQQFHAPSPQSFGGPPPPVSSAPGNFAPPPQSGPPGAFAP
                PPSAFGAPQGPGGPGGYGPPPPGGPGAPGSYGPPQGGPGGFGPPPPGGPGAYGPPPTG
                FPPVGAPPPGAAGAPGGNPFARGPSATGYRQSPYQQ"
gene            complement(join(21795. .21920,22061. .22165,23337. .23672,
                24863. .25354,26233. .26446,27066. .27309,27804. .27912,
                27973. .28209,29577. .29732,29793. .29908,29969. .30092))
```

Alignment Scores:

Pred. No.:	5.25e-07	Length:	33899
Score:	177.50	Matches:	55
Percent Similarity:	41.6%	Conservative:	22
Best Local Similarity:	29.7%	Mismatches:	57
Query Match:	14.9%	Indels:	51
DB:	2	Gaps:	6

US-10-757-745-2_COPY_121_347 (1-227) x CEY63D3A (1-33899)

```
Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 59
              :::||||||||||         |||:::|||                       ||||||
Db      15565 TTTTTGAAGAAAAATGAGGGAAATTTCGAG----------------------TATTTT 15530

Qy         60 ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer--------------Gln 74
                 |||   |||:::|||           |||:::|||
Db      15529 CATCAGATTTCGCTTCGGAAATGTTTGGTGAGGCTAAGCCGAGTTTGTAGTTTGTAACCT 15470

Qy         75 GluIleIlePro-PheProSerThrLysMetMetArgAsnLeuLeuCysValHis----- 92
              |||      ||| ||||||      |||:::         |||:::   :::
Db      15469 GAGCCACAACCACTTTCCCGAAAAAAAACTGAAGAATTTCCTCGTTTTAATTGATTTTTT 15410

Qy         93 ------------------------------ValAsnValSerGlyAsn----------- 98
                                           :::     |||   |||
Db      15409 TGAAAGAAAACTGAACATTTCTATTAAAATCATTCGCTTTTCGGCAAATTTAACAAATTT 15350

Qy         99 ----------------------------------------GluLeuCysLeuMetT 104
                                                      ::::::   |||:::
Db      15349 TTGTAGTTTGTAGCCTGAGCCTCAACCAATTTTCTCGTTTTTTGCAGGTATTCCTACTTA 15290

Qy        104 hrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetV 124
               :::|||||||||||||||||    |||       :::|||       |||
Db      15289 ACACCCATCTCGAATCAACACGTGAGCACCGCCCGCAACGTTGTGCTCAATTCGGCTTCT 15230

Qy        124 alLeuLysLysMetGlnGlu---------AlaProGluSerAlaThrValIlePheAlaG 141
                 :::  |||::::::|||                    |||   |||   |||   |
Db      15229 GCATGGACAAAGTCCGTGAAATCATCGCTCAAAACCCGGGCGCCCTCGTCTTTTTCGGCG 15170

Qy        141 lyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleV 161
              |||||   ||||||||||||    ||||||:::|||                :::|||:::   :::
Db      15169 GCGATTTGAACTTACGAGACGAGGAGGTCAGCCGT---------GTGCCTGACGGTGTTA 15119
```

```
Qy        161 alAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnM 181
Db      15118 AGGATGCCTGGGAGGCTGCCGGCAGTGATAATAAGACAAAATTCACGTGGGACACCTTCA 15059

Qy        181 etAsnSerAsn 184
Db      15058 AAAATGACAAT 15048

RESULT 49
AK113716
LOCUS       AK113716               2036 bp    mRNA    linear   INV 30-NOV-2002
DEFINITION  Ciona intestinalis cDNA, clone:cicl005m10, full insert sequence.
ACCESSION   AK113716
VERSION     AK113716.1  GI:23584533
KEYWORDS    FLI_CDNA.
SOURCE      Ciona intestinalis
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1
  AUTHORS   Satou,Y., Yamada,L., Mochizuki,Y., Takatori,N., Kawashima,T.,
            Sasaki,A., Hamaguchi,M., Awazu,S., Yagi,K., Sasakura,Y.,
            Nakayama,A., Ishikawa,H., Inaba,K. and Satoh,N.
  TITLE     A cDNA resource from the basal chordate Ciona intestinalis
  JOURNAL   Genesis 33 (4), 153-154 (2002)
   PUBMED   12203911
REFERENCE   2  (bases 1 to 2036)
  AUTHORS   Satou,Y. and Satoh,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2002) Nori Satoh, Kyoto University, Department of
            Zoology; Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            (E-mail:satoh@ascidian.zool.kyoto-u.ac.jp, Tel:81-75-753-4095,
            Fax:81-75-705-1113)
COMMENT     Ciona intestinalis cDNA Project (URL:
            http://ghost.zool.kyoto-u.ac.jp/indexr1.html).
FEATURES             Location/Qualifiers
     source          1. .2036
                     /organism="Ciona intestinalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:7719"
                     /clone="cicl005m10"
ORIGIN

Alignment Scores:
Pred. No.:              7.26e-07        Length:         2036
Score:                  162.50          Matches:        54
Percent Similarity:     40.8%           Conservative:   24
Best Local Similarity:  28.3%           Mismatches:     90
Query Match:            13.6%           Indels:         23
DB:                     2               Gaps:           4

US-10-757-745-2_COPY_121_347 (1-227) x AK113716 (1-2036)

Qy          8 AsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspVal 27
Db        193 AACAGAACTGAAAGAATAGAAAATGTTTGTCAAACAATACAAAGGGAGAATCCAACATTG 252

Qy         28 IlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsn 47
Db        253 GTTAGTCTTCAAGAGGTT------GGGGTTTATGAATATAATAAAATTCATAATCATTTC 306

Qy         48 TyrGluIleIleThrGlyHisGlu--------------------------------Glu 56
Db        307 GTAAATATCTATGGTTTCCACTCAGTGAAAAGTTCCATTGGATTCGTTGAAGACTTGGAT 366

Qy         57 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 76
Db        367 TTCTGTATCACAGTCGTAATGTTCAAGAAAGATGTTGTTTTTCATTATTCTGGAAATCAT 426

Qy         77 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 96
Db        427 TATAAATTCCACAGCCTTCGAATGTGTGAATATTTGCATGTGTTTCAAGGAAAAGTTAAA 486

Qy         97 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
Db        487 GGAAATGAATTTGTATTGTTTACATCTCAACTTGAAAAAAGAAAATCAAGTTCTAAAGAA 546

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
Db        547 AGACAAGACCAATTGGGCCAGGTTATTCATTTCCTTAACAAAACTCCAGCCACATCAACA 606

Qy        137 ValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg------CysGly 154
Db        607 GCTATATTTGCTGGTAGTTTATACATTACAGATTTTGAGCTAGAGAAGGAAATGTCGCAA 666

Qy        155 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 174
Db        667 CGAGATGAGAGAAATGTTTGTGAGTTGTGGAAAGAATTAAAACTCAAAAAA--------- 717

Qy        175 TyrThrTrpAspThrGlnMetAsnSerAsnLeu 185
Db        718 ---------------TCTTTCAATGAAAATCTT 735

RESULT 50
CR954207_1/c
WPCOMMENT
  Sequence split into 8 fragments  LOCUS CR954207 Accession CR954207
     Fragment Name      Begin     End
     CR954207_0             1     110000
     CR954207_1        100001     210000
     CR954207_2        200001     310000
     CR954207_3        300001     410000
     CR954207_4        400001     510000
     CR954207_5        500001     610000
     CR954207_6        600001     710000
     CR954207_7        700001     739027
Continuation (2 of 8) of CR954207 from base 100001 (CR954207 Ostreococcus tauri strain C

Alignment Scores:
Pred. No.:              0.00066         Length:         110000
Score:                  153.50          Matches:        61
Percent Similarity:     44.2%           Conservative:   39
Best Local Similarity:  27.0%           Mismatches:     95
Query Match:            12.9%           Indels:         31
DB:                     14              Gaps:           11

US-10-757-745-2_COPY_121_347 (1-227) x CR954207_1 (1-110000)

Qy         12 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 31
Db      16390 AGATTGGAGGCGCTTGGAAAGGTCATCGCGGAGTGCGACCCACACGTGTTGTGTCTGCAG 16331

Qy         32 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 51
Db      16330 GAGGTGACGCCGCAGTCTCTGTTCGCACTTCGCGCGCACGAGTGGTGGAGCGAGTACGTC 16271

Qy         52 ThrGlyHisGlu---------GluGlyTyrPheThrAlaIleMetLeuLysLys------ 66
Db      16270 GTTGGTCCGAAACCGCAATTACAGGAGTACTTCAGCGTCGTGCTGTTCAAGAGGGAGATC 16211

Qy         67 ---SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db      16210 GGGAGCGAGCTAGTGCGGCGAGACCAACAT------CCTTTTGCGAACTCGAAGATGGGG 16157

Qy         86 ArgAsnLeu---------LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu 102
Db      16156 CGGTGCCTGGACATCGTGAGTTACGTACGGTGTCCACTCAGTGGGACGAACTTCGCGGTC 16097

Qy        103 MetThrSerHisLeuGlu-----------SerThrArgGlyHisAlaAlaGluArgMet 118
Db      16096 GGGAGCTCGCATCTCGAGTCCTACATCAGTTCCAATCGTACCAGCGCTCAGGAGCGACGA 16037

Qy        119 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 138
Db      16036 GTGCAGATAGAAGAGTGTTTCTCGCGGCTGAAC---------AAGCACGCGAACGCGATT 15986
```

```
Qy        139 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly------Leu 156
              |||   |||||||||||||      |||::::::            |||         |||
Db      15985 TTCATGGGCGACACGAATTGGGACGATAAGGAT------------GGAGACGTACCTCTG 15938

Qy        157 ProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 176
              |||           |||  |||     |||   :::|||                :::|||
Db      15937 CCGGAGGGATGGCGAGACGCGTGGCTCGAGCTT---CGCCCGGATGAACCTGGTTTCACG 15881

Qy        177 TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAsp 196
              :::|||   :::   |||      |||              :::   |||   |||
Db      15880 TACGACGCGCGCAAAAACGACATGCTA------CGCGGCTACCTACAGAAGCGGCTGGAT 15827

Qy        197 ArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeu 216
              |||   |||   :::                |||         ::::::::::::::
Db      15826 CGAGCGTTCGTTAAGCTCAAG-----------CACTTTGACGTGAAAAACATCGAGATG 15779

Qy        217 LeuGlyLeuGluLysLeu 222
              :::|||   :::   :::
Db      15778 GTGGGCACGCAACCGATT 15761
```

RESULT 51
F25C20

```
LOCUS       F25C20                  91566 bp    DNA     linear   PLN 15-MAY-1999
DEFINITION  Arabidopsis thaliana chromosome 1 BAC F25C20 sequence, complete
            sequence.
ACCESSION   AC007296
VERSION     AC007296.2  GI:4760411
KEYWORDS    HTG.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 91566)
  AUTHORS   Vysotskaia,V.S., Schwartz,J.R., Yu,G., Toriumi,M., Lenz,C., Liu,S.,
            Li,J., Lee,J.M., Kremenetskaia,I., Luros,J., Ngan,I., Liu,A.,
            Gonzalez,A., Altafi,H., Araujo,R., Chao,Q., Conn,L., Conway,A.B.,
            Dunn,P., Hansen,N., Huizar,L., Kim,C., Palm,C., Rowley,D.,
            Shinn,P., Walker,M., Davis,R.W., Ecker,J.R., Federspiel,N.A. and
            Theologis,A.
  TITLE     Arabidopsis thaliana chromosome 1 BAC F25C20 sequence
  JOURNAL   Unpublished (1999)
  REMARK    This sequence is of BAC F25C20 from Arabidopsis thaliana chromosome
            1.  The sequence does not represent the sequence of the entire
            insert of this clone.  It is shorter by 1656 bp because we submit
            only the unique sequence of the clone.  However, in order to
            facilitate the joining of overlapping clones in the future for
            creation of larger contigs, we provide a small overlap (200 bp)
            between overlapping submitted clones.  The 5' end of this sequence
            overlaps by 200 bp the 3' end of the sequence of the BAC F12F1.
REFERENCE   2  (bases 1 to 91566)
  AUTHORS   Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-APR-1999) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
REFERENCE   3  (bases 1 to 91566)
  AUTHORS   Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-MAY-1999) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
REFERENCE   4  (bases 1 to 91566)
  AUTHORS   Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-1999) Plant Gene Expression Center, 800 Buchanan
            St., Albany, CA 94710, USA
COMMENT     On May 8, 1999 this sequence version replaced gi:4582412.
            The sequence of BAC F25C20 from Arabidopsis thaliana chromosome 1.
FEATURES             Location/Qualifiers
     source          1. .91566
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="F25C20"
                     /ecotype="Columbia"
gene                 <4. .294
                     /gene="F25C20.1"
CDS                  join(<4. .50,162. .294)
                     /gene="F25C20.1"
                     /note="EST gb|AA605362 comes from this gene."
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAD30264.1"
                     /db_xref="GI:4835798"
                     /translation="GDAYVSGHGSCIFPGSKKVGNRTQTVVNSTEVAAGEATSRSLSR
                     GFCVTIMILVTFSIL"
gene                 complement(542. .1675)
                     /gene="F25C20.2"
CDS                  complement(542. .1675)
                     /gene="F25C20.2"
                     /note="EST gb|AA605362 comes from this gene."
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAD30240.1"
                     /db_xref="GI:4835774"
                     /translation="MTTTMSTLPVVLVDEILARVPITSLRSLRSTCKKWEASSKTNLV
                     GGKATARKSSHVGFILIGDKICSMKLDLNGGDDFVDTSVNQVSAFDGFAISQLFHCDG
                     LLFCISNNHYSNYTLMVCNMYLGETRLIQNRSLFESYQNFCSYAFGYDSSKNRNHKIL
                     RNNSVSGGYEIYSLKSDSWKDLNVDLEKSIHLWRLGSVSLKGNAYFRVIKVIEEGVWE
                     YNLLCFDFTRESFGKLLSLPFESLGDEGEGIMVISCVRDDHLAVLYQRDTLGIWISTE
                     IEPNKVSWREFLQVDLATLDGFPDVFIAGRFIVDEEKQVVVVFGQETELDLNHGNAFI
                     FGRDGYFTSFTVGDAPPEDFTSYVPSLVSLQIDKTGKRKARDD"
gene                 4175. .5918
                     /gene="F25C20.3"
CDS                  join(4175. .4655,4887. .5035,5129. .5239,5334. .5918)
                     /gene="F25C20.3"
                     /note="EST gb|F14156 comes from this gene."
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAD30241.1"
                     /db_xref="GI:4835775"
                     /translation="MSTTIRFLRIVTSRAMSSSWSCNKCTFLNSASQKLNCMICLAPV
                     SLPSLSPPPPSLSISANDEAKWACKACTFLNTYKNSICDVCGTRSPTSSLLGFQDLTD
                     SGLESNDADSSVGSVFFPLRRCIKRKAMDDDVVEVDGASVVCSESQGVMKKNKEIETK
                     GVASDSGTPLTCLKILSYNVWFREDLELNLRMRAIGHLIQLHSPHLICFQEVTPEIYD
                     IFRKSNWWKAYSCSVSVDVAVSRGYYCMLLSKLGVKSFSSKSFGNSIMGRELSIAEVE
                     VPGRKPLVFATSHLESPCPGPPKWDQMFSRERVEQAKEAIEILRPNANVIFGGDMNWC
                     DKLDGKFPLPDKWVDVWEVLKPGDLGFTYDTKANPMLSGNRALQKRLDRILCRLDDYK
                     LGGIEMVGKEAIPGLSYVKEKKVRGDIKKLELPVLPSDHFGLLVTLSPK"
gene                 complement(6476. .9962)
                     /gene="F25C20.4"
CDS                  complement(join(6476. .6532,6615. .6696,6776. .6843,
                     6930. .7019,7103. .7180,7360. .7455,7533. .7613,7697. .7897,
                     7988. .8080,8167. .8234,9698. .9962))
                     /gene="F25C20.4"
                     /note="Similar to gi|2392772 T32N15.11 putative
                     chloroplast prephenate dehydratase from Arabidopsis
                     thaliana BAC gb|AC002534 and is a member of the PF|00800
                     Prephenate dehydratase family.  ESTs gb|T21562 and
                     gb|T21062 come from this gene."
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAD30242.1"
                     /db_xref="GI:4835776"
                     /translation="MALRCFPIWVCPQTTHHRSPLMGLAEFDADKRRRFCLWECSSSA
                     SQRAVTAIEGEIPFSRELKKSSDELGLTQETQSLSFHRDLSMLPKPLTANSLYSSDGD
                     DSKVRISFQGIPGAYSETAALKAFPNCETVPCEQFEAAFQAVELWLVDKAVLPIENSV
                     GGSIHRNYDLLLRHRLHIVQEVHLPVNHCLLGVPGVKKEDIKCVLSHPQALDQCVNSL
                     NNLGIQRISAKDTATAAQTVSSSGKIDVGAIASVRAANIYGLDILAENIQDDVNNVTR
                     FLILAREPMIPRTDRPYKTSIVFSLEEGPGVLFKALAVFALRSINLSKIESRPQRRRP
                     LRVVDGSNNGSAKYFDYLFYIDFEASMADTRAQHALGHLQEFASFIRILGCYPMDLVR
                     "
gene                 11062. .11290
                     /gene="F25C20.5"
CDS                  join(11062. .11142,11180. .11290)
                     /gene="F25C20.5"
```

```
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F25C20.5"
                     /protein_id="AAD30243.1"
                     /db_xref="GI:4835777"
                     /translation="MRGERDETSNGREEEVIKVVVVCMVCMWSCVNCGRNNGDGGRDD
                     SVMVVSMVVTIAVVMVMMV"
gene                 12261. .13824
                     /gene="F25C20.6"
CDS                  join(12261. .12367,12447. .12595,12676. .12989,13077. .13215,
                     13302. .13368,13453. .13601,13685. .13824)
                     /gene="F25C20.6"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F25C20.6"
                     /protein_id="AAD30244.1"
                     /db_xref="GI:4835778"
                     /translation="MYESANVSDDADRTAFRRAEKKYKLYYEQDSKFSRKKKLPKPID
                     LSELLDFNLISQNFNNDGVLPDGIRVSKVDSSPVFCIDNRPGFYFIPDALSLKEQCKW
                     IKESLTSFPQPPNRTNHNAIYGPIDDLFDSAKENKVLVQDDLTNNKWKFYEEVDIEKA
                     TRSSCKSVSASVLLRKLRWSTLGLQFDWSKRNYDVSLPHNNIPDALCQLAKTHAAIAM
                     PDGEEFRPEGAIVNYFGIGDTLGGHLDDMEADWSKPIVSMSLGCKAIFLLGGKSKDDP
                     PHAMYLRSGDVVLMAGEARECFHGNLLHFQLDGIPRIFTGEENADIGALESELSHESG
                     HFFAEYIKTSRININIRQVF"
gene                 complement(14149. .15759)
                     /gene="F25C20.7"
CDS                  complement(14149. .15759)
                     /gene="F25C20.7"
                     /note="Strong similarity to gb|AF049347 berberine bridge
                     enzyme from Berberis stolonifera."
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAD30246.1"
                     /db_xref="GI:4835780"
                     /translation="MKIFCLILFLISSFISTSLAVEPPPETIYQNFLQCFTNQTKAPP
                     NSLADVVLPKTAAAFTPVLRAYIRNARFNTTATPKPAIVIAARSESHVQAAVICTKSL
                     NIQLKTRSGGHDYEGVSYISHVPFFVLDMSNLRNITVDPATESAWVGAGATLGEVYYR
                     IWEKTKSHGFPAGVCPTVGAGGHISGGGYGNMIRKYGLSVDYVTDAKIVDVNGQVLDR
                     KGMGEDMFWAINGGGGASFGVILAFKIKLVPVPPTVTVFRVEKNLVENATEMVHKWQF
                     VAPKTDPGLFMRLLLQPVTRNKMQTVRASVVALFLGDQNTVMSMLTKEFPELGLKKEN
                     CTEMTWIQSVMWWANNDNATQIKPEILLDRNPDMATFGKRKSDFVEKEITKDGLDFLF
                     KKMIEVGKIGLVFNPYGGIMSTVATTKTPFPHRKKLYKIQHSMNWKDPGTEAETSFLQ
                     KAKSFYSYMAPFVTKNPRHTYINYRDLDIGVNTPGPNSYRVAEVFGRMYFGENFDRLV
                     KVKTAVDPQNFFRDEQSIPTLPGKPARR"
gene                 complement(17116. .17529)
                     /gene="F25C20.8"
CDS                  complement(17116. .17529)
                     /gene="F25C20.8"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F25C20.8"
                     /protein_id="AAD30245.1"
                     /db_xref="GI:4835779"
                     /translation="MKNTSIFLVVFALCMISNAYGHSNPFKGKKTSLLFRNSLSHNKW
                     LKVRCKSGDNDVRERYMKPGEDWGFSFHDDVMGETLFWCTMYKGADYKVQKKFDAYVQ
                     DKRMPHGGSYNYVAQEDGIYHSTLIHRMRKMYDWN"
gene                 18142. .21450
                     /gene="F25C20.9"
CDS                  join(18142. .18195,18614. .19044,19945. .19996,20154. .20424,
                     20494. .20555,20686. .20881,20962. .21058,21147. .21229,
                     21443. .21450)
                     /gene="F25C20.9"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="F25C20.9"
                     /protein_id="AAD30247.1"
                     /db_xref="GI:4835781"
                     /translation="MKKKKKKESDTNRTLRSSPGRDQMDNIVDSLNKAYEKFVLASAGV
                     LESKESAGGQKALLTDTALENFKEKWELFRVACDQAEEFVESVKQRIGSECLVDEATG
                     LTTTAAGGQAPAAVTGAATSLPPISAVRLEQMSRAVRWLVLELQRGSGVAPGSVHSSS
                     TGYMDSNQSMRLLSAWIGQIGDLGLNLLWRFIHIVVSLWYIVSGIFEAIESYAITLGL
                     NKKYGSIDLEKLRCLAVVVDIEAAQDVANVVELLQWLTTIGVKQVGLFDSQGLLKKSK
                     DLILETVPGSMLLEEIEKDVAPDGKRIALEFISSSDNKEAVMKAANILLQRYLKSSHP
```

```
                     EDDKGEDFFTESHLNDALRVVGENVHVPDLLLVYGPIRSHLGFPAWRLRYTEIVHMGT
                     LKYMRYGSLLKAIHKFTGVHQNYGT"
gene                 complement(21902. .23828)

Alignment Scores:
Pred. No.:                0.000601      Length:          91566
Score:                    153.00        Matches:         72
Percent Similarity:       37.4%         Conservative:    32
Best Local Similarity:    25.9%         Mismatches:      85
Query Match:              12.8%         Indels:          89
DB:                       15            Gaps:            11

US-10-757-745-2_COPY_121_347 (1-227) x F25C20 (1-91566)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
              :::   |||:::||||||            |||  |||  :::        :::
Db       4955 GAAGACCTTGAGTTGAAT---------CTAAGGATGAGAGCGATTGGTCATCTTATTCAG 5005

Qy         22 LeuTyrSerProAspValIlePheLeuGln---------------------------- 31
              |||:::||||||   :::|||      |||
Db       5006 CTTCATTCTCCTCATCTCATATGTTTTCAGGTCACTTTTACATTTTGCTATAGAATGATT 5065

Qy         31 ------------------------------------------------------------ 31

Db       5066 TCTATCTGCAACATTTAGAAATAGCTCAGCTTTCTTGACAAGTGGTTTTTGTGTGGTTTG 5125

Qy         32 ---GluValIleProProTyrTyrSerTyrLeuLysLysArg------------------ 44
                 ||||||   |||      |||         :::|||
Db       5126 TAGGAAGTCACTCCTGAGATATATGATATTTTCCGCAAGTCGAATTGGTGGAAAGCGTAT 5185

Qy         45 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe--------------- 59
              |||   :::      :::                 ||||||:::
Db       5186 TCTTGTTCTGTCTCAGTTGATGTGGCAGTTTCGAGAGGCTACTACTGCATGCTGGTAAAG 5245

Qy         59 ------------------------------------------------------------ 59

Db       5246 ATTGTGCTTGCTTGGCTTAGGCTCTTGCCGTCTGTATCTGAATAATTCCCTGTGGTTCTT 5305

Qy         60 ---Thr-AlaIleMet------LeuLysLysSerArgValLysLeuLysSerGlnGluIl 76
                 |||    :::|||      |||:::  |||::::::  :::||||||
Db       5306 GTAACTGACCCTTATGTGTAATTTGCAGTTGAGCAAGTTGGGAGTGAAATCTTTCAGCAG 5365

Qy         76 eIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSe 96
                 |||   ::::::   |||   |||  |||          |||   |||
Db       5366 CAAAAGCTTTGGGAACTCGATAATGGGAAGAGAACTATCTATCGCAGAAGTTGAAGTTCC 5425

Qy         96 rGlyAsnGlu---LeuCysLeuMetThrSerHisLeuGluSer----------------- 109
                 |||   :::   |||          ||||||||||||||||||
Db       5426 AGGGAGAAAGCCACTTGTTTTTGCAACTAGCCATCTTGAGAGTCCCTGTCCAGGTCCTCC 5485

Qy        110 -------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLy 127
                      :::     :::   ||||||:::   |||   |||       ::::::
Db       5486 AAAATGGGACCAAATGTTCAGTCGAGAACGCGTTGAACAAGCCAAGGAGGCAATTGAGAT 5545

Qy        127 sMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAs 147
                :::::           :::|||   |||||||||   ||||||   |||       ||
Db       5546 TCTCAGA---------CCTAATGCCAATGTGATATTCGGGGGAGACATGAACTGGTGCGA 5596

Qy        147 pArgGluValThrArgCysGlyGly---------LeuProAsnAsnIleValAspValTr 164
                |             :::      |||         ||||||:::      ||||||||||
Db       5597 C------------AAGCTAGATGGGAAGTTTCCTTTGCCAGACAAATGGGTGGATGTTTG 5644

Qy        164 pGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAs 184
                ||||   |||   ||||||            :::|||:::|||||:::   |||
Db       5645 GGAGGTTTTG---AAACCAGGTGATTTGGGTTTTACATATGACACAAAAGCAAACCCAAT 5701

Qy        184 nLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                |||            |||   :::   |||   |||||||||      |||
Db       5702 GTTGTCCGGTAACCGGGCATTGCAGAAACGGTTAGACCGGATTTTGTGTCGT 5753

RESULT 52
AC095717/c
LOCUS       AC095717              238494 bp    DNA     linear   HTG 13-NOV-2002
```

```
DEFINITION  Rattus norvegicus clone CH230-9K13, WORKING DRAFT SEQUENCE.
ACCESSION   AC095717
VERSION     AC095717.6  GI:24940991
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 238494)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 238494)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 238494)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 13, 2002 this sequence version replaced gi:22773148.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GDDG
                Center clone name: CH230-9K13
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 221551 bases at least Q40
                Consensus quality: 223436 bases at least Q30
                Consensus quality: 224655 bases at least Q20
                Estimated insert size: 229610; sum-of-contigs estimation
                Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   238494: contig of 238494 bp in length.
FEATURES            Location/Qualifiers
     source         1. .238494
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-9K13"
     misc_feature   8872. .10696
                    /note="wgs_contig"
ORIGIN

Alignment Scores:
Pred. No.:                0.0759          Length:         238494
Score:                    137.50          Matches:        25
Percent Similarity:       72.9%           Conservative:   10
Best Local Similarity:    52.1%           Mismatches:     13
Query Match:              11.5%           Indels:         1
DB:                       14              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x AC095717 (1-238494)

Qy        152 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 171
              :::|||||||||||||||||   |||:::::|||   |||||||||:::||||||||::: 
Db     200723 AAATGTGGTGGTTTACCTGAAAATGTTCTTGATGCCTGGGAATTTGTGGGCAAACCTGAA 200664

Qy        172 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 191
              ||||||||||||||        ::::::|||:::||||||   :::||||||
Db     200663 CATTGCCAGTATACATGATTTATGAAAGTAAATAATAATCTCAG-GTTACTGCCTGGGCC 200605

Qy        192 LysLeuArgPheAspArgIlePhe 199
                    |||       :::|||
Db     200604 TATTTCTGGTTCTGCTCTCTGTTT 200581
```

RESULT 53
AP008218_128
WPCOMMENT
Sequence split into 276 fragments LOCUS AP008218 Accession AP008218

Fragment Name	Begin	End
AP008218_000	1	110000
AP008218_001	100001	210000
AP008218_002	200001	310000
AP008218_003	300001	410000
AP008218_004	400001	510000
AP008218_005	500001	610000
AP008218_006	600001	710000
AP008218_007	700001	810000
AP008218_008	800001	910000
AP008218_009	900001	1010000
AP008218_010	1000001	1110000
AP008218_011	1100001	1210000
AP008218_012	1200001	1310000
AP008218_013	1300001	1410000
AP008218_014	1400001	1510000
AP008218_015	1500001	1610000
AP008218_016	1600001	1710000
AP008218_017	1700001	1810000
AP008218_018	1800001	1910000
AP008218_019	1900001	2010000
AP008218_020	2000001	2110000
AP008218_021	2100001	2210000
AP008218_022	2200001	2310000
AP008218_023	2300001	2410000
AP008218_024	2400001	2510000
AP008218_025	2500001	2610000
AP008218_026	2600001	2710000
AP008218_027	2700001	2810000
AP008218_028	2800001	2910000
AP008218_029	2900001	3010000
AP008218_030	3000001	3110000
AP008218_031	3100001	3210000
AP008218_032	3200001	3310000
AP008218_033	3300001	3410000
AP008218_034	3400001	3510000
AP008218_035	3500001	3610000
AP008218_036	3600001	3710000
AP008218_037	3700001	3810000
AP008218_038	3800001	3910000
AP008218_039	3900001	4010000
AP008218_040	4000001	4110000
AP008218_041	4100001	4210000
AP008218_042	4200001	4310000
AP008218_043	4300001	4410000
AP008218_044	4400001	4510000
AP008218_045	4500001	4610000
AP008218_046	4600001	4710000
AP008218_047	4700001	4810000
AP008218_048	4800001	4910000
AP008218_049	4900001	5010000
AP008218_050	5000001	5110000
AP008218_051	5100001	5210000
AP008218_052	5200001	5310000
AP008218_053	5300001	5410000
AP008218_054	5400001	5510000
AP008218_055	5500001	5610000
AP008218_056	5600001	5710000
AP008218_057	5700001	5810000
AP008218_058	5800001	5910000
AP008218_059	5900001	6010000
AP008218_060	6000001	6110000
AP008218_061	6100001	6210000
AP008218_062	6200001	6310000
AP008218_063	6300001	6410000
AP008218_064	6400001	6510000
AP008218_065	6500001	6610000
AP008218_066	6600001	6710000
AP008218_067	6700001	6810000
AP008218_068	6800001	6910000
AP008218_069	6900001	7010000
AP008218_070	7000001	7110000
AP008218_071	7100001	7210000
AP008218_072	7200001	7310000
AP008218_073	7300001	7410000
AP008218_074	7400001	7510000
AP008218_075	7500001	7610000
AP008218_076	7600001	7710000
AP008218_077	7700001	7810000
AP008218_078	7800001	7910000
AP008218_079	7900001	8010000
AP008218_080	8000001	8110000
AP008218_081	8100001	8210000
AP008218_082	8200001	8310000
AP008218_083	8300001	8410000
AP008218_084	8400001	8510000
AP008218_085	8500001	8610000
AP008218_086	8600001	8710000
AP008218_087	8700001	8810000
AP008218_088	8800001	8910000
AP008218_089	8900001	9010000
AP008218_090	9000001	9110000
AP008218_091	9100001	9210000
AP008218_092	9200001	9310000
AP008218_093	9300001	9410000
AP008218_094	9400001	9510000
AP008218_095	9500001	9610000
AP008218_096	9600001	9710000
AP008218_097	9700001	9810000
AP008218_098	9800001	9910000
AP008218_099	9900001	10010000
AP008218_100	10000001	10110000
AP008218_101	10100001	10210000
AP008218_102	10200001	10310000
AP008218_103	10300001	10410000
AP008218_104	10400001	10510000
AP008218_105	10500001	10610000
AP008218_106	10600001	10710000
AP008218_107	10700001	10810000
AP008218_108	10800001	10910000
AP008218_109	10900001	11010000
AP008218_110	11000001	11110000
AP008218_111	11100001	11210000
AP008218_112	11200001	11310000
AP008218_113	11300001	11410000
AP008218_114	11400001	11510000
AP008218_115	11500001	11610000
AP008218_116	11600001	11710000
AP008218_117	11700001	11810000
AP008218_118	11800001	11910000
AP008218_119	11900001	12010000
AP008218_120	12000001	12110000
AP008218_121	12100001	12210000
AP008218_122	12200001	12310000
AP008218_123	12300001	12410000
AP008218_124	12400001	12510000
AP008218_125	12500001	12610000
AP008218_126	12600001	12710000
AP008218_127	12700001	12810000
AP008218_128	12800001	12910000
AP008218_129	12900001	13010000
AP008218_130	13000001	13110000
AP008218_131	13100001	13210000
AP008218_132	13200001	13310000
AP008218_133	13300001	13410000
AP008218_134	13400001	13510000
AP008218_135	13500001	13610000
AP008218_136	13600001	13710000
AP008218_137	13700001	13810000
AP008218_138	13800001	13910000
AP008218_139	13900001	14010000
AP008218_140	14000001	14110000

```
AP008218_141      14100001  14210000
AP008218_142      14200001  14310000
AP008218_143      14300001  14410000
AP008218_144      14400001  14510000
AP008218_145      14500001  14610000
AP008218_146      14600001  14710000
AP008218_147      14700001  14810000
AP008218_148      14800001  14910000
AP008218_149      14900001  15010000
AP008218_150      15000001  15110000
AP008218_151      15100001  15210000
AP008218_152      15200001  15310000
AP008218_153      15300001  15410000
AP008218_154      15400001  15510000
AP008218_155      15500001  15610000
AP008218_156      15600001  15710000
AP008218_157      15700001  15810000
AP008218_158      15800001  15910000
AP008218_159      15900001  16010000
AP008218_160      16000001  16110000
AP008218_161      16100001  16210000
AP008218_162      16200001  16310000
AP008218_163      16300001  16410000
AP008218_164      16400001  16510000
AP008218_165      16500001  16610000
AP008218_166      16600001  16710000
AP008218_167      16700001  16810000
AP008218_168      16800001  16910000
AP008218_169      16900001  17010000
AP008218_170      17000001  17110000
AP008218_171      17100001  17210000
AP008218_172      17200001  17310000
AP008218_173      17300001  17410000
AP008218_174      17400001  17510000
AP008218_175      17500001  17610000
AP008218_176      17600001  17710000
AP008218_177      17700001  17810000
AP008218_178      17800001  17910000
AP008218_179      17900001  18010000
AP008218_180      18000001  18110000
AP008218_181      18100001  18210000
AP008218_182      18200001  18310000
AP008218_183      18300001  18410000
AP008218_184      18400001  18510000
AP008218_185      18500001  18610000
AP008218_186      18600001  18710000
AP008218_187      18700001  18810000
AP008218_188      18800001  18910000
AP008218_189      18900001  19010000
AP008218_190      19000001  19110000
AP008218_191      19100001  19210000
AP008218_192      19200001  19310000
AP008218_193      19300001  19410000
AP008218_194      19400001  19510000
AP008218_195      19500001  19610000
AP008218_196      19600001  19710000

Alignment Scores:
Pred. No.:                0.0347          Length:         110000
Score:                    137.00          Matches:        48
Percent Similarity:       45.1%           Conservative:   21
Best Local Similarity:    31.4%           Mismatches:     62
Query Match:              11.5%           Indels:         22
DB:                       15              Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x AP008218_128 (1-110000)

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 83
              ::::::   |||::::::   ::::::|||   :::      ||||||   ::::::
Db      27848 ATGCAGATGAGCAAATTGCCTGTTGAGTCTTTTGACTGCACTCCATTTTCCAACTCAATA 27907

Qy         84 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn---GluLeuCysLeu 102
              |||   |||   |||         ||||||         |||      :::|||   |||

Db      27908 ATGGGAAGAGAACTGTGTGTAGCACATGTCAAAACTGGAGGTGCAGTTAAGCTAGTGCTG 27967

Qy        103 MetThrSerHisLeuGluSer----------------------ThrArgGlyHisAla 114
              ||||||||||||||||||                        :::   ::::::
Db      27968 GCCACTAGCCACCTTGAGAGTCCAATGCCTGGGCCTCCTACTTGGGATCAGATGTACAGC 28027

Qy        115 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 134
              ||||||:::   |||             ||||||   :::
Db      28028 ACGGAACGAGTTGCTCAGGCAAACAAATCTCTCAAAATTTTG---------GGATCCTTC 28078

Qy        135 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 154
              |||||||||   ||||||   |||      |||:::               |||
Db      28079 CGCAATGTAATATTTTGTGGCGATATGAACTGGGATGACAAA----------GGA 28123

Qy        155 Gly------------LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 170
                          ||||||          |||  |||      |||   ||||||
Db      28124 GATGGGCCCTTCCCCCTACCAGCTGGCTGGACCGATGCTTGGATTGAGTTG---AAGCCG 28180

Qy        171 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 190
                     :::|||:::||||||:::   ||||||   |||
Db      28181 GGTGAAGATGGCTGGACTTATGACACAAAAGCTAACAGTATGCTATCAGCCAACCGTAAG 28240

Qy        191 CysLysLeuArgPheAspArgIlePhePheArgAlaAla 203
                 :::   |||   ||||||        :::   |||
Db      28241 TTACAGAAGAGATTGGACCGATTCGTGTGCAAGCTGGCA 28279

RESULT 54
CNS08CB2/c
LOCUS       CNS08CB2              156724 bp    DNA     linear   PLN 21-NOV-2003
DEFINITION  Oryza sativa chromosome 12, . BAC OSJNBa0036H15 of library OSJNBa
            from chromosome 12 of cultivar Nipponbare of ssp. japonica of Oryza
            sativa (rice), complete sequence.
ACCESSION   AL844879
VERSION     AL844879.5  GI:28625295
KEYWORDS    HTG.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 156724)
  AUTHORS   Choisne,N., Orjeda,G., Cattolico,L., Demange,N., Wincker,P.,
            Segurens,B., Pelletier,E., Scarpelli,C., Salanoubat,M.,
            Weissenbach,J. and Quetier,F.
  TITLE     Oryza sativa chromosome 12 sequencing
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 156724)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2003) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Feb 28, 2003 this sequence version replaced gi:28460625.
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            ---------------
            The following sequence is oriented from the T7 to the SP6 end. The
            nucleotide sequence of this BAC clone was generated by combining
            Syngenta and Genoscope sequencing data.
            Upstream BAC (overlapping the T7 end) : OSJNBa0017G24 (AC=AL928777)
            Downstream BAC (overlapping the SP6 end) : OSJNBa0032G05
            (AC=AL831800) ---------- Finishing boundaries
            FINISHED SEGMENT STARTS AT BASE 5830
            FINISHED SEGMENT ENDS AT BASE 100667
            ----------.
FEATURES             Location/Qualifiers
     source          1. .156724
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /sub_species="japonica"
```

```
                     /db_xref="taxon:39947"
                     /chromosome="12"
                     /clone="OSJNBa0036H15"
                     /clone_lib="OSJNBa"
ORIGIN

Alignment Scores:
Pred. No.:                0.0524          Length:         156724
Score:                    137.00          Matches:        48
Percent Similarity:       45.1%           Conservative:   21
Best Local Similarity:    31.4%           Mismatches:     62
Query Match:              11.5%           Indels:         22
DB:                       15              Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x CNS08CB2 (1-156724)

Qy          64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 83
               :::::: |||:::::: ::::::||| ::: ||||||  ::::::
Db       91379 ATGCAGATGAGCAAATTGCCTGTTGAGTCTTTTGACTGCACTCCATTTTCCAACTCAATA 91320

Qy          84 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn---GluLeuCysLeu 102
               ||| ||| ||| |||||| ||| :::||| |||
Db       91319 ATGGGAAGAGAACTGTGTGTAGCACATGTCAAAACTGGAGGTGCAGTTAAGCTAGTGCTG 91260

Qy         103 MetThrSerHisLeuGluSer----------------------ThrArgGlyHisAla 114
               ||||||||||||||||||                     ::: ::::::
Db       91259 GCCACTAGCCACCTTGAGAGTCCAATGCCTGGGCCTCCTACTTGGGATCAGATGTACAGC 91200

Qy         115 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 134
               ||||||::: ||| |||||| :::
Db       91199 ACGGAACGAGTTGCTCAGGCAAACAAATCTCTCAAAATTTTG---------GGATCCTTC 91149

Qy         135 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 154
               ||||||||| |||||| ||| ||:::  |||
Db       91148 CGCAATGTAATATTTTGTGGCGATATGAACTGGGATGACAAA--------------GGA 91104

Qy         155 Gly------------LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 170
                           ||||||              ||| ||| ||| ||||||
Db       91103 GATGGGCCCTTCCCCCTACCAGCTGGCTGGACCGATGCTTGGATTGAGTTG---AAGCCG 91047

Qy         171 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 190
                        :::|||:::||||||::: ||||||  |||
Db       91046 GGTGAAGATGGCTGGACTTATGACACAAAAGCTAACAGTATGCTATCAGCCAACCGTAAG 90987

Qy         191 CysLysLeuArgPheAspArgIlePhePheArgAlaAla 203
               :::  |||  |||||| ::: |||
Db       90986 TTACAGAAGAGATTGGACCGATTCGTGTGCAAGCTGGCA 90948

RESULT 55
AC149575/c
LOCUS       AC149575               111678 bp    DNA     linear   PLN 09-JUN-2004
DEFINITION  Populus trichocarpa clone Pop1-41G18, complete sequence.
ACCESSION   AC149575
VERSION     AC149575.1  GI:48475260
KEYWORDS    HTG.
SOURCE      Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
  ORGANISM  Populus trichocarpa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 111678)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 111678)
  AUTHORS   Stanford Human Genome Center.
  CONSRTM   DOE Joint Genome Institute
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUN-2004) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
FEATURES             Location/Qualifiers
     source          1. .111678
                     /organism="Populus trichocarpa"
                     /mol_type="genomic DNA"
                     /sub_species="trichocarpa"
                     /db_xref="taxon:3694"
                     /clone="Pop1-41G18"
ORIGIN

Alignment Scores:
Pred. No.:                0.0571          Length:         111678
Score:                    135.00          Matches:        46
Percent Similarity:       48.6%           Conservative:   25
Best Local Similarity:    31.5%           Mismatches:     55
Query Match:              11.3%           Indels:         20
DB:                       15              Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x AC149575 (1-111678)

Qy          64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 83
               |||::: |||:::::: ::: |||||| |||||| ::::::
Db       61185 TTGCAGTTAAGCAAACTGCCAGTAAAATCCTTCAGCACTAAACCCTTTATGAATTCAATA 61126

Qy          84 MetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu---LeuCysLeu 102
               ||| ||| ||| ::: ||| ||| ::: ||| :::
Db       61125 ATGGGGAGAGAACTTTGCATTGCTGAGCTTGAAGTTCCAGGAAAAAAGTCATTAGTTGTT 61066

Qy         103 MetThrSerHisLeuGluSer----------------------ThrArgGlyHisAla 114
               |||||||||||||||||||                    ::: :::
Db       61065 GCCACTAGCCATCTAGAGAGTCCATGTCCAGCACCTCCAAAATGGGATCAGATGTTCAGC 61006

Qy         115 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 134
               ||||||:::::|||  |||  ::: ::::::        ::::::
Db       61005 AAGGAACGTGTGGATCAGGCAAAGGAGGCCATCAACCTTCTCAAG---------AAGAAC 60955

Qy         135 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 154
               :::  |||||||||  |||||  |||  |||                ::::
Db       60954 TCAAATGTGATTTTTTGTGGTGACATGAACTGGGATGAT------------AAGTTGGAT 60907

Qy         155 Gly---------LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 171
               |||                ||:::  ||||||  |||  |||  ||||||
Db       60906 GGTCAGTTTCCTTTCCCTGATGGATGGGTTGATGCCTGGGTAGAGTTG---AAGCCAGGG 60850

Qy         172 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 191
                  :::|||:::||||||:::  |||  |||  |||
Db       60849 GACAATGGATGGACATATGATACCAAATCCAATCAAATGTTGTCTGGTAATCGTGCATTG 60790

Qy         192 LysLeuArgPheAspArg 197
               ::: |||  ||||||
Db       60789 CAAAAGAGACTGGATCGA 60772

RESULT 56
AR240556/c
LOCUS       AR240556                268 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 312 from patent US 6468758.
ACCESSION   AR240556
VERSION     AR240556.1  GI:27285636
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 268)
  AUTHORS   Benson,D.R., Lodes,M.J., Mitcham,J.L. and King,G.E.
  TITLE     Compositions and methods for ovarian cancer therapy and diagnosis
  JOURNAL   Patent: US 6468758-A 312 22-OCT-2002;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .268
                     /organism="unknown"
```

```
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                6.41e-05        Length:          268
Score:                    134.00          Matches:         26
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              11.2%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x AR240556 (1-268)

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAA 208

Qy        222 LeuAspCysGlyArgPhe 227
              |||||||||||||||||
Db        207 CTGGACTGTGGTAGATTT 190

RESULT 57
AP004494/c
LOCUS       AP004494               103129 bp    DNA     linear   PLN 22-JUL-2003
DEFINITION  Lotus corniculatus var. japonicus genomic DNA, chromosome 4,
            clone:LjT09I17, TM0026, complete sequence.
ACCESSION   AP004494
VERSION     AP004494.1  GI:17736861
KEYWORDS    HTG.
SOURCE      Lotus corniculatus var. japonicus (Lotus japonicus)
  ORGANISM  Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1
  AUTHORS   Sato,S., Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T. and Tabata,S.
  TITLE     Structural Analysis of a Lotus japonicus Genome. I. Sequence
            Features and Mapping of Fifty-six TAC clones which cover the 5.4 Mb
            Regions of the Genome
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 103129)
  AUTHORS   Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-DEC-2001) Yasukazu Nakamura, Kazusa DNA Research
            Institute, Department of Plant Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
            URL:http://www.kazusa.or.jp, Tel:81-438-52-3935,
            Fax:81-438-52-3934)
FEATURES             Location/Qualifiers
     source          1. .103129
                     /organism="Lotus corniculatus var. japonicus"
                     /mol_type="genomic DNA"
                     /variety="japonicus"
                     /db_xref="taxon:34305"
                     /chromosome="4"
                     /clone="LjT09I17"
                     /clone_lib="LjT library"
                     /note="TAC clone:TM0026
                     synonym: Lotus japonicus"
ORIGIN

Alignment Scores:
Pred. No.:                0.173           Length:          103129
Score:                    130.00          Matches:         45
Percent Similarity:       45.9%           Conservative:    33
Best Local Similarity:    26.5%           Mismatches:      70
Query Match:              10.9%           Indels:          22
DB:                       15              Gaps:            7

US-10-757-745-2_COPY_121_347 (1-227) x AP004494 (1-103129)

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 83
              |||:::   |||::::::   :::||||||                  |||   ::::::
Db      32779 TTGCAGTTAAGCAAACTACCTGTGAAATCTTTCAGCAACAAGTCCTTTAGCAATTCTATA 32720

Qy         84 MetMetArgAsnLeuLeuCysValHisValAsn---ValSerGlyAsnGluLeuCysLeu 102
              |||   |||   |||           |||         :::|||       |||   :::
Db      32719 ATGGGAAGGGAACTTTGCATTGCTGAGGTTGAAGCTGCAAATGGCAAGTCATTGGTTATT 32660

Qy        103 MetThrSerHisLeuGluSer----------------------ThrArgGlyHisAla 114
              ||||||||||||||||||                        :::   ::::::
Db      32659 GCCACTAGCCATCTTGAGAGTCCCTGTCCAGGCCCTCCGAAATGGGATCAGATGTACAGC 32600

Qy        115 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 134
              ||||||:::   |||           :::     :::::::::   |||
Db      32599 AAGGAACGAGTAGAGCAGGCGAACGAGGCTGTAAACCTGCTCAAGAAGCACCCT------ 32546

Qy        135 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 154
              |||:::|||   ||||||  |||     |||:::                  |||
Db      32545 ---AATGTTGTTTTTGGAGGTGACATGAACTGGGATGACAAA---------CTGGATGGT 32498

Qy        155 GlyLeuPro------AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              |||     :::     :::|||  |||     |||   :::|||
Db      32497 CAATATCCTTCACAAGATGGGTGGATCGATGCCTGGTCTCAGCTA---AGGCCAGATCAA 32441

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              :::|||:::|||||:::   |||     |||           :::
Db      32440 AGTGGGTGGACGTATGATACTAAGTCGAACCAGATGTTGTCAGGCAACCGTACTCTCCAA 32381

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              |||  ||||||        |||   :::           :::
Db      32380 AAGCGATTAGATCGTTTCATTTGCCGTTTAAGT-----------GATCTTAAGATAAGC 32333

Qy        213 SerLeuAspLeuLeuGlyLeuGluLysLeu 222
              |||:::||:::::::|||::::::   :::
Db      32332 AGCATTGACATGATTGGGATGGATGCAATA 32303

RESULT 58
AC151708
LOCUS       AC151708               86100 bp    DNA     linear   PLN 16-DEC-2004
DEFINITION  Medicago truncatula clone mth2-15h5, complete sequence.
ACCESSION   AC151708
VERSION     AC151708.6  GI:56461838
KEYWORDS    HTG.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 86100)
  AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
  TITLE     Medicago truncatula BAC Clone mth2-15h5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 86100)
  AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-SEP-2004) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   3  (bases 1 to 86100)
  AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2004) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   4  (bases 1 to 86100)
  AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-2004) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
```

```
            OK 73019, USA
REFERENCE   5  (bases 1 to 86100)
  AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
            and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-2004) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Dec 9, 2004 this sequence version replaced gi:54306117.
            -------------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            ----------------------.
FEATURES             Location/Qualifiers
     source          1. .86100
                     /organism="Medicago truncatula"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3880"
                     /clone="mth2-15h5"
                     /clone_lib="Medicago truncatula BAC library H2"
ORIGIN

Alignment Scores:
Pred. No.:                0.366          Length:          86100
Score:                    126.00         Matches:         46
Percent Similarity:       43.8%          Conservative:    35
Best Local Similarity:    24.9%          Mismatches:      76
Query Match:              10.6%          Indels:          28
DB:                       15             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x AC151708 (1-86100)

Qy          55 GluGluGlyTyrPheThrAlaIleMet-----------------LeuLysLysSerArg 68
               |||   |||   ::::::   ::::::                  :::   |||:::
Db       64637 GAAAGAGGGTGTTATTCACATGTACTGAAATATCTGTCAAATTAATTTCAGTTAAGCAAA 64696

Qy          69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 88
               :::   :::||||||              |||   ::::::   |||   |||   |||
Db       64697 CTGCCGGTGAAATCTTTCTCCGCCAAATCATTTAGCAACTCTATAATGGGAAGAGAACTT 64756

Qy          89 LeuCysValHisVal---AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 107
                  |||   :::|||   |||          :::   ||||||||||||
Db       64757 TGCATTGCTGAAGTGGAGGATGTGGGCGGCAAGTCGTTTGTTGTGGCCACTAGCCATCTC 64816

Qy         108 GluSer-----------------------ThrArgGlyHisAlaAlaGluArgMetAsn 119
               ||||||                  :::   :::   ||||||:::
Db       64817 GAGAGTCCTTGTCCGGCCCCTCCAAAATGGGATCAGATGTTCAGCAAGGAACGAGTGGAG 64876

Qy         120 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 139
               |||           |||      ::::::      |||          |||:::|||
Db       64877 CAGGCTAATGAGGCTCTGAACATTCTCAAAAGGCACCCT---------AATGTTGTTTTT 64927

Qy         140 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro------ 157
               ||||||   |||      |||:::         :::   |||     |||
Db       64928 GGGGGTGATATGAACTGGGATGACAAA---------AAGGATGGTCAATATCCTTTGCAA 64978

Qy         158 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
               :::      :::|||   |||      |||   :::|||          :::|||:::
Db       64979 GACGGGTGGCTTGATGCCTGGTCTGTACTA---AGACCAAATGAAGCTGGTTGGACATAT 65035

Qy         178 AspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArg 197
               ||||||:::   |||      |||                 :::   |||   ||||||
Db       65036 GATACCAAGTCAAACCAAATGTTGACAGGCAATCGTACTCTCCAAAAGCGATTGGATCGT 65095

Qy         198 IlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu 217
                  |||                                 ::::::|||::::::
Db       65096 TTTGTTTGTCGTTTACGT-----------GATTTTAAGATAAGCAATATTGACATGATT 65143

Qy         218 GlyLeuGluLysLeu 222
               |||::::::::::::
Db       65144 GGGATGGATGAAATA 65158
```

```
RESULT 59
AC134967/c
LOCUS       AC134967               133028 bp    DNA     linear   PLN 15-AUG-2003
DEFINITION  Medicago truncatula clone mth2-26b8, complete sequence.
ACCESSION   AC134967
VERSION     AC134967.19  GI:33589882
KEYWORDS    HTG.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 133028)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Medicago truncatula BAC Clone mth2-26b8
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 133028)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-OCT-2002) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   3  (bases 1 to 133028)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-AUG-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
REFERENCE   4  (bases 1 to 133028)
  AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
            Cook,D., Kim,D. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2003) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Aug 12, 2003 this sequence version replaced gi:33342303.
            -------------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            ----------------------.
FEATURES             Location/Qualifiers
     source          1. .133028
                     /organism="Medicago truncatula"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3880"
                     /clone="mth2-26b8"
                     /clone_lib="Medicago truncatula BAC library H2"
ORIGIN

Alignment Scores:
Pred. No.:                0.608          Length:          133028
Score:                    126.00         Matches:         46
Percent Similarity:       43.8%          Conservative:    35
Best Local Similarity:    24.9%          Mismatches:      76
Query Match:              10.6%          Indels:          28
DB:                       15             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x AC134967 (1-133028)

Qy          55 GluGluGlyTyrPheThrAlaIleMet-----------------LeuLysLysSerArg 68
               |||   |||   ::::::   ::::::                  :::   |||:::
Db       68387 GAAAGAGGGTGTTATTCACATGTACTGAAATATCTGTCAAATTAATTTCAGTTAAGCAAA 68328

Qy          69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 88
               :::   :::||||||              |||   ::::::   |||   |||   |||
Db       68327 CTGCCGGTGAAATCTTTCTCCGCCAAATCATTTAGCAACTCTATAATGGGAAGAGAACTT 68268

Qy          89 LeuCysValHisVal---AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 107
```

```
Db      68267 TGCATTGCTGAAGTGGAGGATGTGGGCGGCAAGTCGTTTGTTGTGGCCACTAGCCATCTC 68208

Qy        108 GluSer-----------------------ThrArgGlyHisAlaAlaGluArgMetAsn 119

Db      68207 GAGAGTCCTTGTCCGGCCCCTCCAAAATGGGATCAGATGTTCAGCAAGGAACGAGTGGAG 68148

Qy        120 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 139

Db      68147 CAGGCTAATGAGGCTCTGAACATTCTCAAAAGGCACCCT---------AATGTTGTTTTT 68097

Qy        140 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro------ 157

Db      68096 GGGGGTGATATGAACTGGGATGACAAA---------AAGGATGGTCAATATCCTTTGCAA 68046

Qy        158 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177

Db      68045 GACGGGTGGCTTGATGCCTGGTCTGTACTA---AGACCAAATGAAGCTGGTTGGACATAT 67989

Qy        178 AspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArg 197

Db      67988 GATACCAAGTCAAACCAAATGTTGACAGGCAATCGTACTCTCCAAAAGCGATTGGATCGT 67929

Qy        198 IlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu 217

Db      67928 TTTGTTTGTCGTTTACGT-----------GATTTTAAGATAAGCAATATTGACATGATT 67881

Qy        218 GlyLeuGluLysLeu 222

Db      67880 GGGATGGATGAAATA 67866

RESULT 60
CR626927_15
WPCOMMENT
  Sequence split into 52 fragments  LOCUS CR626927 Accession CR626927
     Fragment Name      Begin     End
     CR626927_00            1    110000
     CR626927_01       100001    210000
     CR626927_02       200001    310000
     CR626927_03       300001    410000
     CR626927_04       400001    510000
     CR626927_05       500001    610000
     CR626927_06       600001    710000
     CR626927_07       700001    810000
     CR626927_08       800001    910000
     CR626927_09       900001   1010000
     CR626927_10      1000001   1110000
     CR626927_11      1100001   1210000
     CR626927_12      1200001   1310000
     CR626927_13      1300001   1410000
     CR626927_14      1400001   1510000
     CR626927_15      1500001   1610000
     CR626927_16      1600001   1710000
     CR626927_17      1700001   1810000
     CR626927_18      1800001   1910000
     CR626927_19      1900001   2010000
     CR626927_20      2000001   2110000
     CR626927_21      2100001   2210000
     CR626927_22      2200001   2310000
     CR626927_23      2300001   2410000
     CR626927_24      2400001   2510000
     CR626927_25      2500001   2610000
     CR626927_26      2600001   2710000
     CR626927_27      2700001   2810000
     CR626927_28      2800001   2910000
     CR626927_29      2900001   3010000
     CR626927_30      3000001   3110000
     CR626927_31      3100001   3210000
     CR626927_32      3200001   3310000
     CR626927_33      3300001   3410000
     CR626927_34      3400001   3510000
     CR626927_35      3500001   3610000
     CR626927_36      3600001   3710000
     CR626927_37      3700001   3810000
     CR626927_38      3800001   3910000
     CR626927_39      3900001   4010000
     CR626927_40      4000001   4110000
     CR626927_41      4100001   4210000
     CR626927_42      4200001   4310000
     CR626927_43      4300001   4410000
     CR626927_44      4400001   4510000
     CR626927_45      4500001   4610000
     CR626927_46      4600001   4710000
     CR626927_47      4700001   4810000
     CR626927_48      4800001   4910000
     CR626927_49      4900001   5010000
     CR626927_50      5000001   5110000
     CR626927_51      5100001   5205140
Continuation (16 of 52) of CR626927 from base 1500001 (CR626927 Bacteroides fragilis NCT.

Alignment Scores:
Pred. No.:                2.06          Length:          110000
Score:                    120.00        Matches:         55
Percent Similarity:       36.2%         Conservative:    37
Best Local Similarity:    21.7%         Mismatches:      102
Query Match:              10.1%         Indels:          61
DB:                       1             Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x CR626927_15 (1-110000)

Qy          3 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 22

Db      65296 GGCGACAGTTTGAATTCCTGGAAGTATCGTAAAGAGAAGGTCTGTGAATTTATCCGTGAG 65355

Qy         23 TyrSerProAspValIlePheLeuGlnGluValIlePro---------------------- 35

Db      65356 AAACATCCCGATGTGCTGGGAATGCAGGAAGTACTGAATCATCAGTTGAAAGACCTGTTG 65415

Qy         36 ---------ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52

Db      65416 TCCGGCCTGCCCGATTATGCCTATGTAGGTGTGGGACGCGAGGAT--------------- 65460

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72

Db      65461 GGAAAGACTCAGGGGGAATATGCTCCTGTATTCTATCGGAAAGATAAATATGATTTGTTA 65520

Qy         73 SerGlnGluIleIle---------------ProPheProSerThrLysMetMetArgAsn 87

Db      65521 GACAGTAATACATTCTGGCTTTCCGAACATCCGGATAGTATCGGCAAACTTGGTTGGGAT 65580

Qy         88 LeuLeuCysValHisVal-----------------------AsnValSerGlyAsnGlu 99

Db      65581 GCCGCGTGTACCCGCGTGGCTACTTGGGCTAAATTAAAGGATAAAACTACCGGAAAAGAG 65640

Qy        100 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 119

Db      65641 TTTCTGATGTTAAACACCCATTTTGACCACGTTGGTACGGAAGCT------CGTCGTAAC 65694

Qy        120 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 139

Db      65695 AGTGCCCTGCTTATTATTGATAAGATAAAAGAAATAGCCGGTACACATCCTGCTATGATG 65754

Qy        140 AlaGlyAspThrAsnLeuArgAsp--------------ArgGluValThrArgCysGly 154

Db      65755 ACCGGAGACTTTAATGTTTCTGAAGAGTGGGAAGCTTACAAAACGATCACA--------- 65805

Qy        155 GlyLeuProAsnAsnIleVal------AspValTrpGluPheLeuGlyLysProLysHis 172

Db      65806 ------TCCAATGAGTTTGTCTTGAAAGATGCATGGAAGATAGCCGGGAAACAGTCCGGT 65859

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192

Db      65860 GAGAACTATACTTTCCATGATTT-CGGAAGAGTACCTGTGGGCGAACGGGAAAA------ 65912

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly--------------- 207

Db      65913 ------TCCAATGAGTTTGTCTTGAAAGATGCATGGAAGATAGCCGGGAAACAGTCCGGT 65966
```

```
Db      65913 ---------GATCGATTTCATTTTCGTAACTCCGCAAATCAAGGTGGCAGATGCAGAGAT 65963

Qy        208 ------------------------HisIleIleProArgSer 213
                                      |||:::      ||||||
Db      65964 TATTTCTTCTGCGATTACAGATTCCACCTATTTGTCCGATCA 66005
```

RESULT 61
AP006841_15
WPCOMMENT
Sequence split into 53 fragments LOCUS AP006841 Accession AP006841

Fragment Name	Begin	End
AP006841_00	1	110000
AP006841_01	100001	210000
AP006841_02	200001	310000
AP006841_03	300001	410000
AP006841_04	400001	510000
AP006841_05	500001	610000
AP006841_06	600001	710000
AP006841_07	700001	810000
AP006841_08	800001	910000
AP006841_09	900001	1010000
AP006841_10	1000001	1110000
AP006841_11	1100001	1210000
AP006841_12	1200001	1310000
AP006841_13	1300001	1410000
AP006841_14	1400001	1510000
AP006841_15	1500001	1610000
AP006841_16	1600001	1710000
AP006841_17	1700001	1810000
AP006841_18	1800001	1910000
AP006841_19	1900001	2010000
AP006841_20	2000001	2110000
AP006841_21	2100001	2210000
AP006841_22	2200001	2310000
AP006841_23	2300001	2410000
AP006841_24	2400001	2510000
AP006841_25	2500001	2610000
AP006841_26	2600001	2710000
AP006841_27	2700001	2810000
AP006841_28	2800001	2910000
AP006841_29	2900001	3010000
AP006841_30	3000001	3110000
AP006841_31	3100001	3210000
AP006841_32	3200001	3310000
AP006841_33	3300001	3410000
AP006841_34	3400001	3510000
AP006841_35	3500001	3610000
AP006841_36	3600001	3710000
AP006841_37	3700001	3810000
AP006841_38	3800001	3910000
AP006841_39	3900001	4010000
AP006841_40	4000001	4110000
AP006841_41	4100001	4210000
AP006841_42	4200001	4310000
AP006841_43	4300001	4410000
AP006841_44	4400001	4510000
AP006841_45	4500001	4610000
AP006841_46	4600001	4710000
AP006841_47	4700001	4810000
AP006841_48	4800001	4910000
AP006841_49	4900001	5010000
AP006841_50	5000001	5110000
AP006841_51	5100001	5210000
AP006841_52	5200001	5277274

Continuation (16 of 53) of AP006841 from base 1500001 (AP006841 Bacteroides fragilis YCH

Alignment Scores:

Pred. No.:	2.06	Length:	110000
Score:	120.00	Matches:	55
Percent Similarity:	36.2%	Conservative:	37
Best Local Similarity:	21.7%	Mismatches:	102
Query Match:	10.1%	Indels:	61
DB:	1	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x AP006841_15 (1-110000)

```
Qy          3 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 22
              |||      ||||||:::        |||         ||||||   ::::::
Db      61293 GGCGACAGTTTGAATTCCTGGAAGTATCGTAAAGAGAAGGTCTGTGAATTTATCCGTGAG 61352

Qy         23 TyrSerProAspValIlePheLeuGlnGluValIlePro---------------------- 35
              |||||||||:::   :::||||||||||:::
Db      61353 AAACATCCCGATGTGCTGGGAATGCAGGAAGTACTGAATCATCAGTTGAAAGACCTGTTG 61412

Qy         36 ---------ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
                      |||   |||:::|||:::                  :::
Db      61413 TCCGGCCTGCCTGATTATGCCTATGTAGGTGTGGGACGCGAGGAT--------------- 61457

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              |||      :::|||   :::      :::      :::|||   :::      |||
Db      61458 GGAAAGACTCAGGGGGAATATGCTCCTGTATTCTATCGGAAAGATAAATATGATTTGTTA 61517

Qy         73 SerGlnGluIleIle--------------ProPheProSerThrLysMetMetArgAsn 87
                                           |||             |||:::     :::
Db      61518 GACAGTAATACATTCTGGCTTTCCGAACATCCGGATAGTATCGGCAAACTTGGTTGGGAT 61577

Qy         88 LeuLeuCysValHisVal----------------------AsnValSerGlyAsnGlu 99
              |||      |||                           :::|||   |||
Db      61578 GCCGCGTGTACCCGCGTGGCTACTTGGGCTAAATTAAAGGATAAAACTACCGGAAAAGAG 61637

Qy        100 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 119
              ::::::   :::|||   :::                |||      |||   |||
Db      61638 TTTCTGATGTTAAACACCCATTTTGACCACGTTGGTACGGAAGCT------CGTCGTAAC 61691

Qy        120 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 139
              :::::::::   |||::::::|||          :::         :::
Db      61692 AGTGCCCTGCTTATTATTGATAAGATAAAAGAAATAGCCGGTACACATCCTGCTATGATG 61751

Qy        140 AlaGlyAspThrAsnLeuArgAsp--------------ArgGluValThrArgCysGly 154
              ||||||   |||:::   :::              :::   :::|||
Db      61752 ACCGGAGACTTTAATGTTTCTGAAGAGTGGGAAGCTTACAAAACGATCACA--------- 61802

Qy        155 GlyLeuProAsnAsnIleVal------AspValTrpGluPheLeuGlyLysProLysHis 172
              |||      |||      |||   |||:::      ||||||
Db      61803 ------TCCAATGAGTTTGTCTTGAAAGATGCATGGAAGATAGCCGGGAAACAGTCCGGT 61856

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||:::                |||      |||   |||
Db      61857 GAGAACTATACTTTCCATGATTT-CGGAAGAGTACCTGTGGCCGAACGGGAAAA------ 61909

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly--------------- 207
              ||||||     ||||||   :::|||   :::|||
Db      61910 ---------GATCGATTTCATTTTCGTAACTCCGCAAATCAAGGTGGCAGATGCAGAGAT 61960

Qy        208 ------------------------HisIleIleProArgSer 213
                                      |||:::      ||||||
Db      61961 TATTTCTTCTGCGATTACAGATTCCACCTATTTGTCCGATCA 62002
```

RESULT 62
AP008216_130
WPCOMMENT
Sequence split into 227 fragments LOCUS AP008216 Accession AP008216

Fragment Name	Begin	End
AP008216_000	1	110000
AP008216_001	100001	210000
AP008216_002	200001	310000
AP008216_003	300001	410000
AP008216_004	400001	510000
AP008216_005	500001	610000
AP008216_006	600001	710000
AP008216_007	700001	810000
AP008216_008	800001	910000
AP008216_009	900001	1010000
AP008216_010	1000001	1110000
AP008216_011	1100001	1210000
AP008216_012	1200001	1310000

AP008216_013	1300001	1410000
AP008216_014	1400001	1510000
AP008216_015	1500001	1610000
AP008216_016	1600001	1710000
AP008216_017	1700001	1810000
AP008216_018	1800001	1910000
AP008216_019	1900001	2010000
AP008216_020	2000001	2110000
AP008216_021	2100001	2210000
AP008216_022	2200001	2310000
AP008216_023	2300001	2410000
AP008216_024	2400001	2510000
AP008216_025	2500001	2610000
AP008216_026	2600001	2710000
AP008216_027	2700001	2810000
AP008216_028	2800001	2910000
AP008216_029	2900001	3010000
AP008216_030	3000001	3110000
AP008216_031	3100001	3210000
AP008216_032	3200001	3310000
AP008216_033	3300001	3410000
AP008216_034	3400001	3510000
AP008216_035	3500001	3610000
AP008216_036	3600001	3710000
AP008216_037	3700001	3810000
AP008216_038	3800001	3910000
AP008216_039	3900001	4010000
AP008216_040	4000001	4110000
AP008216_041	4100001	4210000
AP008216_042	4200001	4310000
AP008216_043	4300001	4410000
AP008216_044	4400001	4510000
AP008216_045	4500001	4610000
AP008216_046	4600001	4710000
AP008216_047	4700001	4810000
AP008216_048	4800001	4910000
AP008216_049	4900001	5010000
AP008216_050	5000001	5110000
AP008216_051	5100001	5210000
AP008216_052	5200001	5310000
AP008216_053	5300001	5410000
AP008216_054	5400001	5510000
AP008216_055	5500001	5610000
AP008216_056	5600001	5710000
AP008216_057	5700001	5810000
AP008216_058	5800001	5910000
AP008216_059	5900001	6010000
AP008216_060	6000001	6110000
AP008216_061	6100001	6210000
AP008216_062	6200001	6310000
AP008216_063	6300001	6410000
AP008216_064	6400001	6510000
AP008216_065	6500001	6610000
AP008216_066	6600001	6710000
AP008216_067	6700001	6810000
AP008216_068	6800001	6910000
AP008216_069	6900001	7010000
AP008216_070	7000001	7110000
AP008216_071	7100001	7210000
AP008216_072	7200001	7310000
AP008216_073	7300001	7410000
AP008216_074	7400001	7510000
AP008216_075	7500001	7610000
AP008216_076	7600001	7710000
AP008216_077	7700001	7810000
AP008216_078	7800001	7910000
AP008216_079	7900001	8010000
AP008216_080	8000001	8110000
AP008216_081	8100001	8210000
AP008216_082	8200001	8310000
AP008216_083	8300001	8410000
AP008216_084	8400001	8510000
AP008216_085	8500001	8610000
AP008216_086	8600001	8710000
AP008216_087	8700001	8810000
AP008216_088	8800001	8910000
AP008216_089	8900001	9010000
AP008216_090	9000001	9110000
AP008216_091	9100001	9210000
AP008216_092	9200001	9310000
AP008216_093	9300001	9410000
AP008216_094	9400001	9510000
AP008216_095	9500001	9610000
AP008216_096	9600001	9710000
AP008216_097	9700001	9810000
AP008216_098	9800001	9910000
AP008216_099	9900001	10010000
AP008216_100	10000001	10110000
AP008216_101	10100001	10210000
AP008216_102	10200001	10310000
AP008216_103	10300001	10410000
AP008216_104	10400001	10510000
AP008216_105	10500001	10610000
AP008216_106	10600001	10710000
AP008216_107	10700001	10810000
AP008216_108	10800001	10910000
AP008216_109	10900001	11010000
AP008216_110	11000001	11110000
AP008216_111	11100001	11210000
AP008216_112	11200001	11310000
AP008216_113	11300001	11410000
AP008216_114	11400001	11510000
AP008216_115	11500001	11610000
AP008216_116	11600001	11710000
AP008216_117	11700001	11810000
AP008216_118	11800001	11910000
AP008216_119	11900001	12010000
AP008216_120	12000001	12110000
AP008216_121	12100001	12210000
AP008216_122	12200001	12310000
AP008216_123	12300001	12410000
AP008216_124	12400001	12510000
AP008216_125	12500001	12610000
AP008216_126	12600001	12710000
AP008216_127	12700001	12810000
AP008216_128	12800001	12910000
AP008216_129	12900001	13010000
AP008216_130	13000001	13110000
AP008216_131	13100001	13210000
AP008216_132	13200001	13310000
AP008216_133	13300001	13410000
AP008216_134	13400001	13510000
AP008216_135	13500001	13610000
AP008216_136	13600001	13710000
AP008216_137	13700001	13810000
AP008216_138	13800001	13910000
AP008216_139	13900001	14010000
AP008216_140	14000001	14110000
AP008216_141	14100001	14210000
AP008216_142	14200001	14310000
AP008216_143	14300001	14410000
AP008216_144	14400001	14510000
AP008216_145	14500001	14610000
AP008216_146	14600001	14710000
AP008216_147	14700001	14810000
AP008216_148	14800001	14910000
AP008216_149	14900001	15010000
AP008216_150	15000001	15110000
AP008216_151	15100001	15210000
AP008216_152	15200001	15310000
AP008216_153	15300001	15410000
AP008216_154	15400001	15510000
AP008216_155	15500001	15610000
AP008216_156	15600001	15710000
AP008216_157	15700001	15810000
AP008216_158	15800001	15910000

AP008216_159	15900001	16010000
AP008216_160	16000001	16110000
AP008216_161	16100001	16210000
AP008216_162	16200001	16310000
AP008216_163	16300001	16410000
AP008216_164	16400001	16510000
AP008216_165	16500001	16610000
AP008216_166	16600001	16710000
AP008216_167	16700001	16810000
AP008216_168	16800001	16910000
AP008216_169	16900001	17010000
AP008216_170	17000001	17110000
AP008216_171	17100001	17210000
AP008216_172	17200001	17310000
AP008216_173	17300001	17410000
AP008216_174	17400001	17510000
AP008216_175	17500001	17610000
AP008216_176	17600001	17710000
AP008216_177	17700001	17810000
AP008216_178	17800001	17910000
AP008216_179	17900001	18010000
AP008216_180	18000001	18110000
AP008216_181	18100001	18210000
AP008216_182	18200001	18310000
AP008216_183	18300001	18410000
AP008216_184	18400001	18510000
AP008216_185	18500001	18610000
AP008216_186	18600001	18710000
AP008216_187	18700001	18810000
AP008216_188	18800001	18910000
AP008216_189	18900001	19010000
AP008216_190	19000001	19110000
AP008216_191	19100001	19210000
AP008216_192	19200001	19310000
AP008216_193	19300001	19410000
AP008216_194	19400001	19510000
AP008216_195	19500001	19610000
AP008216_196	19600001	19710000

Alignment Scores:

Pred. No.:	2.62	Length:	110000
Score:	119.00	Matches:	46
Percent Similarity:	44.7%	Conservative:	22
Best Local Similarity:	30.3%	Mismatches:	62
Query Match:	10.0%	Indels:	22
DB:	15	Gaps:	7

US-10-757-745-2_COPY_121_347 (1-227) x AP008216_130 (1-110000)

```
Qy         59 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 78
              |||           ::::::   |||::::::   :::                ||||||
Db      46083 TTCAATTGCAAACAAATGCAGATGAGCAAATTACCGGTGAGCGAATCGAACCCCATCCCA 46142

Qy         79 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 98
              |||   |||      :::|||:::      |||||||||   |||      :::|||
Db      46143 TTT---TCGAAATCAATAATGAAGAGGGAGCTATGCGTAGCGGTCGTCAAAACGGGAGAG 46199

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 |||  :::  ||||||||||||||||||
Db      46200 ATCCATTTGGCTGTGGGCACAAGTCACCTAGAGAGCCCC------TGCCCTCTACCTCCT 46253

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
              :::      ||||||         |||::::::   :::|||   :::|||      :::
Db      46254 CTGTGGGATCTGAAGTACAGCGAGAAGCGAGTAGCTCAGGCAAAGCAGTCTCTGGAGATT 46313

Qy        138 --------------------IlePheAlaGlyAspThrAsnLeuArgAspArgGluVal 150
                                 ||||||   ||||||   |||      |||:::   |||
Db      46314 CTTGGACAATTGCGCAATGCAATCTTCTGCGGCGACATGAACTGGGAAGACAAA---GTA 46370

Qy        151 ThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 170
                    ||||||:::      :::|||   |||      |||   ||||||
Db      46371 GATGGTCCATTCCCTCTACCTGACGGCTGGATTGATGCTTGGGTGGAGCTG---AAGCCA 46427
```

```
Qy        171 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr------ 188
                       :::|||:::|||||||:::   |||:::   |||
Db      46428 GGTGACAATGGTTGGACCTATGATACGAAAGCCAATGCCATGCTGTCAGCCAACTTCAAG 46487

Qy        189 --------------------AlaAlaCysLysLeu 193
                                 |||||||||
Db      46488 CAGCAGAAGAGGCCAGACCAGTTCGTGTGCAAGTTG 46523
```

RESULT 63
AC083943

```
LOCUS       AC083943               169270 bp    DNA     linear   PLN 14-JUL-2002
DEFINITION  Genomic sequence for Oryza sativa, Nipponbare strain, clone
            OSJNBa0044A10, from chromosome 10, complete sequence.
ACCESSION   AC083943
VERSION     AC083943.8  GI:21328304
KEYWORDS    HTG.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 169270)
  AUTHORS   McCombie,W.R., de la Bastide,M., Spiegel,L., Preston,R.,
            Ferraro,K., Kuit,K., Nascimento,L., Zutavern,T., Balija,V.,
            Bell,M., Baker,J., Miller,B., Katzenberger,F., Muller,S., King,L.,
            Sullivan,P., Yang,C., Dike,S., O'Shaughnessy,A., Palmer,L. and
            Dedhia,N.
  TITLE     Genomic sequence for Oryza sativa, Nipponbare strain, clone
            OSJNBa0044A10, from chromosome 10, complete sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 169270)
  AUTHORS   McCombie,W.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-OCT-2000) Lita Annenberg Hazen Genome Center, Cold
            Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,
            NY 11724, USA
REFERENCE   3  (bases 1 to 169270)
  AUTHORS   McCombie,W.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Lita Annenberg Hazen Genome Center, Cold
            Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,
            NY 11724, USA
  REMARK    Updated sequence
REFERENCE   4  (bases 1 to 169270)
  AUTHORS   McCombie,W.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Lita Annenberg Hazen Genome Center, Cold
            Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor,
            NY 11724, USA
REFERENCE   5  (bases 1 to 169270)
  AUTHORS   Palmer,L.E., de la Bastide,M., Spiegel,L., Preston,R., Ferraro,K.,
            Kuit,K., Nascimento,L., Zutavern,T., Balija,V., Bell,M., Baker,J.,
            Miller,B., Katzenberger,F., Muller,S., King,L., Sullivan,P.,
            Yang,C., Dike,S., O'Shaughnessy,A., Dedhia,N. and McCombie,W.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) Lita Annenberg Hazen Genome Center, Cold
            Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY
            11724, USA
  REMARK    Genomic sequence for Oryza sativa (japonica cultivar-group) clone
            OSJNBa0044A10, from chromosome 10, complete sequence.
REFERENCE   6  (bases 1 to 169270)
  AUTHORS   McCombie,W.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2002) Genome Research Center, Cold Spring Harbor
            Laboratory, 500 Sunnyside Blvd, Woodbury, NY 11797, USA
COMMENT     On Jun 7, 2002 this sequence version replaced gi:20303537.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest.  The nucleotide
```

```
          sequence of this BAC clone was generated by combining Syngenta and
          Cold Spring Harbor Laboratory Genome Center sequencing data.
          Clone OSJNBa0044A10 overlaps clone OSJNBa0050N08 (AC021891) from
          base 1 to base 1196.  The overlap is from base 143398 to base
          144593 on OSJNBa0050N08.
FEATURES             Location/Qualifiers
     source          1. .169270
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="10"
                     /clone="OSJNBa0044A10"
                     /clone_lib="HindIII"
     repeat_region   complement(30. .72)
                     /note="Simple_repeat (GA)n"
     gene            866. .2216
                     /gene="OSJNBa0044A10.1"
                     /note="Hypothetical protein"
     CDS             join(866. .997,1623. .2216)
                     /gene="OSJNBa0044A10.1"
                     /codon_start=1
                     /product="Hypothetical protein"
                     /protein_id="AAM74261.1"
                     /db_xref="GI:21671899"
                     /translation="MSVPFQPQVPGNKENEKIEPLPVVVLHPPARGLPPNLEAKLAGR
                     MSKLPVSESNPIPFSKSIMKRELCVAVVKTGEIHLAVGTSHLESPCPLPPLWDLKYSE
                     KRVAQAKQSLEILGQLRNAIFCGDMNWEDKVDGPFPLPDGWIDAWVELKPGDNGWTYD
                     TKANAMLSANFKQQKRPDQFVCKLSDFKIDDIEMIGKEAIPGVVYYKEKIVRKEFHKL
                     ELPVLPSKHFGLVLTITLQDDIL"
     gene            3104. .3905
                     /gene="OSJNBa0044A10.2"
                     /note="Hypothetical protein"
     CDS             join(3104. .3351,3848. .3905)
                     /gene="OSJNBa0044A10.2"
                     /codon_start=1
                     /product="Hypothetical protein"
                     /protein_id="AAM74262.1"
                     /db_xref="GI:21671900"
                     /translation="MVVEGGNGEKEARDEDGTWEIGGGDCAVEIGSEGGEHAEDTGGS
                     GEHERKGERDQELCFTVASSNLDLNSITHWRRRAEDGSGKNSYLENSNMFFAVCSHA"
     gene            4504. .4863
                     /gene="OSJNBa0044A10.3"
                     /note="Hypothetical protein"
     CDS             4504. .4863
                     /gene="OSJNBa0044A10.3"
                     /codon_start=1
                     /product="Hypothetical protein"
                     /protein_id="AAM74263.1"
                     /db_xref="GI:21671901"
                     /translation="MARDVEAVAENSRREPEWWALPPSPPLSSPFVALVIMVLAATTV
                     VTVLECTVGGCPMAECTKWDAKGRRMLVSNFFGASVSELYARRGGGEGGGGAGCIWRV
                     RAPTPSEGSHYGSLSTC"
     repeat_region   complement(5046. .5094)
                     /note="Simple_repeat (GA)n"
     repeat_region   complement(5591. .6298)
                     /note="Similar to Oryza latifolia dispersed repeat"
     repeat_region   complement(6662. .6825)
                     /note="Similar to Oryza latifolia dispersed repeat"
     gene            6898. .7743
                     /gene="OSJNBa0044A10.4"
                     /note="Hypothetical protein similar to putative
                     retroelements"
     CDS             join(6898. .6990,7093. .7743)
                     /gene="OSJNBa0044A10.4"
                     /note="Similar to gb|AAM08793.1|AC090486_3 (AC090486)
                     Hypothetical protein with similarity to putative
                     retroelement [Oryza sativa] [Oryza sativa (japonica
                     cultivar-group)] . E-value=3e-60"
                     /codon_start=1
                     /product="Hypothetical protein similar to putative
                     retroelements"
                     /protein_id="AAM74264.1"
                     /db_xref="GI:21671902"
                     /translation="MAEKQPPPLSSPGSVKGKALKQGLTDIIDDHTRKGILVQKYKIK
                     VVADSPGTGSSKDGEVKQAPGGSTQPSTKGATNGSSSDQGDNSQGVHGVQGDDAQEPQ
                     GGNLNQNSELAQDFFNNFQDRVDYAVHHALINQSGVLVNTLSNMMKTIADGSVAEHQA
                     AGPVYLQGGVFPNYQSLITDVQPSVQAVPSVASTTQPTAPASTPLPATLAMMLGQLVN
                     PWLLTRDQPQHGGQATNRLTCNAPIFVRD"
     repeat_region   7714. .16477
                     /note="Similar to putative retrotransposon,
                     Ty3-gypsy-like"
     gene            16501. .21022
                     /gene="OSJNBa0044A10.5"
                     /note="Putative retroelement"
     CDS             join(16501. .17607,18210. .18401,18649. .19067,19183. .19589,
                     19635. .21022)
                     /gene="OSJNBa0044A10.5"
                     /note="Similar to sp|Q05654|RDPO_SCHPO Retrotransposable
                     element Tf2 155 kDa protein . E-value=3e-45; Contains pfam
                     domain: rvt Reverse transcriptase (RNA-dependent DN
                     e-value=7.9e-27; Contains pfam domain: rnaseH RNase H
                     e-value=0.0016"
                     /codon_start=1
                     /product="Putative retroelement"
                     /protein_id="AAM74265.1"
                     /db_xref="GI:21671903"
                     /translation="MFLPPPPIFDPVQQQPIQQTPPIQPIVQSIQQQVVQPAQQTPPR
                     QQPQQPIQQTPLRQQVVQPIQHPGSVNASAGFAVPGGQPVQQFVNQVIPEHLVHHAQP
                     DGTVMPQMVPEHLVRNIQPNLHNYQGGNLSYQYQPPSLQAQYQLGGLAQPQFASQYGQ
                     FESMQQQPQGAAQQRPWADVIADVMREQFGLKPKETGSLYRQPYPEWFERVPLPNRFK
                     VPDFSKFSGQEGVSTYEHISRFLAQCHEASAVDALKVRLFRLSLSGSAFTWFSSLPYG
                     SINSWADLEKQFHSYFYSGVHEMKLSDLTAIKQRHDEPVHEYIQRFREMRNKCFSLSL
                     TNAQLADLAFQGMIPPIREKFSSEDFDSLSHLIQKGCEFFRFCWNEGMRLPSIENCPG
                     CSDFIESSSRSYSRGNRLRQARLPVHQRLGPMNQDCGQEDEAIFEKPEGTENRHLKPL
                     YINGYVNGKTMSKMMVDGGAAVNLMPYATFRKLGRNTEALIKTNMVLKDFGGNPSETR
                     GVLNVELTVGSKTIPTTFFVINGNGSYSLLLGRDWIHANCCIPSTMHQCLIQWQGDKT
                     EIVPADSQLKMENPCYYFEGVVEGSNVYAKDTVDDLDDKQGQDFMSADELEEIDIGPG
                     DRPRPTFIRKNLSPEFRTKLIELLKEFRDCFAWEYYEMPGLSRSIVEHRLPIKPGVRP
                     CQQPPRRCKANMLEPVKAEIKRLYDAVPVIKKNGKVRVCIDFRDLNKATPKDEYPMPV
                     ADQLVDAASGNKILSFVDGNAGYNQIFMAEEDIHKMAFRCPSAIGLFEWVVMTFGLKS
                     AGATYQRAMNYIYHDLIGWLVEVYIDNVVVKSKEVEDHIADLRKVFERTRKYGLKMNP
                     TKCAFGVSAGQFLGFLVHERGIEVTQRSVNAIKKIQPPENKTELQEMIGKINFVRRFI
                     SNLSGRLEPFTPLLRLKANQQFTWGAEQQKALDNIKEYLSSPPVLIPPQKGIPFRLYL
                     SAGEKSIGSVLIQELEGKERVVFYLSRRLLDAETQYSPVEKLCLCLYFSCTRLRHYLL
                     SNECTVICKADVVRYMLLAPILKRRVGKWIFSLTEFDLRYESPKAIKGQAIADFIVEH
                     RDDSIGSVEIVPWTLFFDGSVCTHGCGISLVIISPRAACFEFAYTIKPYATNNQAEYE
                     AVLKGLQLLKEVEADTLKSWGILC"
     repeat_region   complement(17064. .17507)
                     /note="Similar to putative retrotransposon,
                     Ty3-gypsy-like"
     repeat_region   complement(19634. .19747)
                     /note="Similar to putative retrotransposon,
                     Ty3-gypsy-like"
     repeat_region   complement(19772. .20040)
                     /note="Similar to putative retrotransposon,
                     Ty3-gypsy-like"
     repeat_region   complement(19878. .20173)
                     /note="Similar to putative retrotransposon,
                     Ty3-gypsy-like"
     repeat_region   21530. .21874
                     /note="Similar to putative retrotransposon,
                     Ty3-gypsy-like"
     repeat_region   22441. .22489
                     /note="Similar to putative retrotransposon,
                     Ty3-gypsy-like"
     gene            complement(22828. .25666)

Alignment Scores:
Pred. No.:                4.33          Length:          169270
Score:                    119.00        Matches:         46
Percent Similarity:       44.7%         Conservative:    22
Best Local Similarity:    30.3%         Mismatches:      62
Query Match:              10.0%         Indels:          22
DB:                       15            Gaps:            7
```

```
US-10-757-745-2_COPY_121_347 (1-227) x AC083943 (1-169270)

Qy          59 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 78
               |||              ::::::   |||::::::   :::                 ||||||
Db        1602 TTCAATTGCAAACAAATGCAGATGAGCAAATTACCGGTGAGCGAATCGAACCCCATCCCA 1661

Qy          79 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 98
               |||   |||      :::|||:::      |||||||||   |||      :::|||
Db        1662 TTT---TCGAAATCAATAATGAAGAGGGAGCTATGCGTAGCGGTCGTCAAAACGGGAGAG 1718

Qy          99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                  |||   :::   ||||||||||||||||||
Db        1719 ATCCATTTGGCTGTGGGCACAAGTCACCTAGAGAGCCCC------TGCCCTCTACCTCCT 1772

Qy         118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
               :::      ||||||         |||::::::   :::|||   :::|||      :::
Db        1773 CTGTGGGATCTGAAGTACAGCGAGAAGCGAGTAGCTCAGGCAAAGCAGTCTCTGGAGATT 1832

Qy         138 --------------------IlePheAlaGlyAspThrAsnLeuArgAspArgGluVal 150
                                 ||||||   ||||||   |||      |||:::   |||
Db        1833 CTTGGACAATTGCGCAATGCAATCTTCTGCGGCGACATGAACTGGGAAGACAAA---GTA 1889

Qy         151 ThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 170
                        ||||||:::      :::|||   |||      |||   ||||||
Db        1890 GATGGTCCATTCCCTCTACCTGACGGCTGGATTGATGCTTGGGTGGAGCTG---AAGCCA 1946

Qy         171 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr------ 188
                  :::|||:::||||||:::   |||:::   |||
Db        1947 GGTGACAATGGTTGGACCTATGATACGAAAGCCAATGCCATGCTGTCAGCCAACTTCAAG 2006

Qy         189 --------------------AlaAlaCysLysLeu 193
                                   |||||||||
Db        2007 CAGCAGAAGAGGCCAGACCAGTTCGTGTGCAAGTTG 2042

RESULT 64
AE017091
LOCUS       AE017091               324367 bp    DNA     linear   PLN 06-JUN-2003
DEFINITION  Oryza sativa (japonica cultivar-group) chromosome 10, section 45 of
            77 of the complete sequence.
ACCESSION   AE017091 AE016959
VERSION     AE017091.1  GI:31431918
KEYWORDS    .
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 324367)
  AUTHORS   .
  CONSRTM   The Rice Chromosome 10 Sequencing Consortium
  TITLE     In-depth view of structure, activity, and evolution of rice
            chromosome 10
  JOURNAL   Science 300, 1566-1569 (2003)
REFERENCE   2  (bases 1 to 324367)
  AUTHORS   Buell,C.R., Wing,R.A., McCombie,W.R., Messing,J. and Yuan,Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2003) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT     This is the pseudomolecule for rice chromosome 10, which was
            constructed by resolving discrepancies between overlapping BACs,
            trimming the overlap regions, and linking the unique sequences to
            form a contigous sequence. Genes in individual BAC clone were
            identified by a combination of several methods: Gene prediction
            programs, searches of the complete sequence against a peptide
            database and EST databases. Genes with similarity to other proteins
            are named after the database hits. Genes without significant
            peptide similarity but with EST similarity are named as  unknown
            proteins. Genes without protein or EST similarity, that are
            predicted by more than two gene prediction programs over most of
            their length are annotated as  hypothetical  proteins. Genes
            encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy,
            http://genome.wustl.edu/eddy/tRNAscan-SE/).
FEATURES             Location/Qualifiers
     source          1. .324367
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="10"
     misc_feature    <1. .31465
                     /note="Chromosome Sequence Derivation: nucleotide sequence
                     in this region was derived from BAC clone OSJNBb0058B20
                     (GB:AC108884)."
     gene            complement(<555. .>2788)
                     /locus_tag="OSJNBb0058B20.8"
                     /note="hypothetical protein"
     mRNA            complement(join(<555. .724,2335. .>2788))
                     /locus_tag="OSJNBb0058B20.8"
     CDS             complement(join(555. .724,2335. .2788))
                     /locus_tag="OSJNBb0058B20.8"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAP53625.1"
                     /db_xref="GI:31431919"
                     /translation="MPVRSPPLLSLPTRGRRLRRGEARRGAGEARVRPWRRAGQRRQL
                     PAPLPSPPLPDPARGRPGEGGGGEGRRGAAMTTHGVAAAAPARLPSPPLPDLAGARPG
                     QGGGGEARRGATTTTRGAAAAAVRLPSPPRFGRREAGAGRRQRRRGEAQGLSAAEAVS
                     MAGVMAIRASSTSAAGHPGELRLSHRSPERTREIEQGLGRDWVEGDG"
     gene            complement(<4256. .>6116)
                     /locus_tag="OSJNBb0058B20.7"
                     /note="Similar to  (AP003208) hypothetical protein similar
                     to Oryza sativa chromosome 1, P0416G11 19 E value 8e 15
                     Identities 52 267 (of subject)"
     mRNA            complement(join(<4256. .4458,4711. .5342,6109. .>6116))
                     /locus_tag="OSJNBb0058B20.7"
     CDS             complement(join(4256. .4458,4711. .5342,6109. .6116))
                     /locus_tag="OSJNBb0058B20.7"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="AAP53626.1"
                     /db_xref="GI:31431920"
                     /translation="MWKAYCVSRGEGARGPGPREWGPRGSLTVHGGPGAPGLTPTGAV
                     GPTHQSHPRARAADGRAPRGSHSARPRAAKPAQPAGGGARAPMVAGGGHRHGGAAAER
                     GEGKGEKKRRSTAHPRSTATTKKAAGAEEGGGAARVDGVDGVPAVGERNGGVDEVGED
                     AAKPKEAAPRWEVVRGDDGGGPELDGDGGERERRRELDSGEEEGRREAETDGGGRGRG
                     GGSGGRKGMTGGPHPSARVAGGPARQRCARGERPDGPRGEEREGEGDGPSRPKRERES
                     KGTF"
     gene            <6199. .>10984
                     /locus_tag="OSJNBb0058B20.6"
                     /note="hypothetical protein"
     mRNA            join(<6199. .6342,6458. .6535,7355. .7504,7737. .8256,
                     8513. .8849,8903. .9326,9491. .9862,10013. .10252,
                     10346. .10387,10547. .>10984)
                     /locus_tag="OSJNBb0058B20.6"
     CDS             join(6199. .6342,6458. .6535,7355. .7504,7737. .8256,
                     8513. .8849,8903. .9326,9491. .9862,10013. .10252,
                     10346. .10387,10547. .10984)
                     /locus_tag="OSJNBb0058B20.6"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAP53627.1"
                     /db_xref="GI:31431921"
                     /translation="MDISRTNEGYTSCGPVVEMSWHRAGVLLLGAQSCLPVPGVPAVG
                     GVGSMYTSDQSIIYSNSRALGYGQAEQCTQHLLTPQAFRPFSYFAAANAPITMAPRKP
                     NPASAKGPDPGRVDDDNTAFLGFGSLSSTGRAGLPGLSLTPSSSAPYFRHRELEAVIT
                     PAGTKKTVFRSVNFNVRPERSDLWPVNAAMSKWDRHWIARWFYHTIPFEAGSDPAKAL
                     RCWRRAIAPNRKPKIAVDGAMEARFVLLRKVCSHISCHDLVEEFCMLRIFPLSQSWQV
                     TVDQGEEVDGLPNLILPEGANRPSRKRMRGQVKLAPRKRRAPASSDSDADDEDDAEER
                     DGEEEGEEEEEEVVADKAANEAVEDRIDTPGYTPTPSPGHNETGVESNSSPLRRTDLE
                     GAKALVAFSSGKEAKEVRDDESSDGTPTSPAGRSLDLSTAPLVPLGVAGAGGSIAAGA
                     SASAERIVSAVARVFGSPLRELAISPLVKAKGKCAGAEASASEYSLAAPHFAPGDFET
                     RADLVPFVEGVSNLVLPAGTPSLFTELNEFDEGCSAIKSLAVRPSSGQGRRNKPQELG
                     GGGPGQTLKEAKAEVERLQSELDKGKEARAEVDRLKAELEKEKAHSAVLVDYYNLTEP
                     KMEALRLKASRAEASAVEESQRFLREMAKTAESARTACQTLRLALTDMGTKFGCEHVA
                     EFPTYAKGDWEISAQNISPALRAWRIQFWQKDGRSAAKARLLEQLAKAEAADRGEGED
                     AACEEGGGNAQDHPEAVGNPPYCSRRRPLGVHRKSCIRPPSGSWRSQNEKFVALKGLP
```

```
                KVRCPFRPVYREMEINGRSYKVFKVADWLRAHPGRTLEDYDLVHSRRLEDMARFWRNH
                QKTKRQETIFQRRYSLVCVSPPTPLKVVYDISSEDEPSSPDHSLGRDSDYGGEDVIYL
                "
     gene       complement(<11934. .>16576)
                /locus_tag="OSJNBb0058B20.5"
                /note="Similar to  (AC093181) putative polyprotein (Oryza
                sativa) E value 0 0 Identities 1089 1931 (of subject);
                contains pfam domain rve Integrase core domain e value 3
                9e 33; contains pfam domain rvt Reverse transcriptase (RNA
                dependent e value 6 5e 29; Contai"
     mRNA       complement(join(<11934. .12425,12546. .12736,12924. .13222,
                13262. .13807,13983. .15376,15620. .>16576))
                /locus_tag="OSJNBb0058B20.5"
     CDS        complement(join(11934. .12425,12546. .12736,12924. .13222,
                13262. .13807,13983. .15376,15620. .16576))
                /locus_tag="OSJNBb0058B20.5"
                /codon_start=1
                /product="putative retroelement"
                /protein_id="AAP53628.1"
                /db_xref="GI:31431922"
                /translation="MAAQQVPPVGVGQPNAMAQPHVQAAISPFATPYPQQGAVNRAGG
                EKGLPLSGGIKTRPIPPQFKFPPVPRYSGETDPKEFLSIYESAIEAAHGDENTKAKVI
                HLALDGIARSWYFNLPDNSIYSWEQLRDVPVLNFRGTYEEPKTQQHLLGIRQRPGESI
                REYMRRFSQARCQVQDITEASVINAASAGLLEGELTRKIANKEPLTLEHLLCIIDGFA
                RGEEDSKRRQAIQAEYDKASITAAQAQAQVQVAEPPPLSVRQSQPAIQGQPPRQSQAP
                ITWRKFRTDRAGKAVMAVEEVQALRKEFDAQQASNHQQPAHKKKKMQTRMVHSITSAG
                EGAPQYLNQLISFGPEDAEGVIFPHQDPLVISAEIAGFEVRRILVDGGSSADVIFAEA
                YAKMGLPTQALTPAPVSLRGFGGEAVQVLGQTLLLIAFGSGENRREEQILFDVVDIPY
                NYNAIFGRATLNKFEAISHHNYLKLKMPGPTAVIVVKGLQPSATSKGDLAIIHRAVHS
                VETEPHERLKHMPKPTPHSKMTTVQIDDVDPTKLVSLGGDMGEGELESILEVLKKNID
                IFAWSPDEVGGVPANLIMHHLAVKPDIKPRKQKLRKMSADRQEAAKAEVQKLLRAGVI
                QEINHPEWLANPVLVRKSNGKWRMCVDFTDLNKACPKDDFPLPRIDQLVDSTARCELM
                SFLDAYSGYHQIHMNPLDIPKTSFITPFGTFCHLRMPFGLRNAGATFATLVYKVLGKQ
                LGRNVEAYIDDIVVKSRKAFDHAIDLQETFDSLRAAGIKLNPENAAERGLPFFKTLRG
                AGKFNWTPECQAAFDELKQYLQSSPALISPPPGSELLLYLAASPVAVSAALVQETEFG
                QKPVYFISEVLQGAKTRYIEMEKLAYALVMVSHKLKHYFQAHKVIVPSQYPLGEILRG
                KEVTGRLSKWAAELSPFDLHFIARFAIKSQVLADFVAEWTPALATDREPTEQRAGIAA
                ALFSPNGVPIRYAARLQFDTTNNATEYEAVLLGLRKAKALGVQRLLIRTDSKLVAGHV
                DKSFEAKEERMKRYLEAVRSMEKCFTGITVEHLPRDEAEAKRLQLRATKYKMVSGQLY
                RSGVLQPLLRCISFAEGEEMAKEIHQGLCGAHQAARTVASKPIAPIWPFPRWGLDIIG
                PFPVARNGYKFAIVSVEYFSGWIEAEPLGAITSAAVQKFVWKNIVCRFGVPKEFIPDN
                GKQFDSDKFREMCEGLNLEIRFASVAHPQSNGAAERTNGKILEALKKRLEGAAKGKWP
                EELLSVLWALWTTPTRPSLARLCFFMAMRQ"
     gene       complement(<18432. .>18917)
                /locus_tag="OSJNBb0058B20.4"
                /note="Similar to  (AC021891) putative disease resistance
                protein (Oryza sativa) E value 6e 48 Identities 101 161
                (of subject); contains pfam domain disease_resp Plant
                disease resistance response protei e value 9 1e 05"
     mRNA       complement(<18432. .>18917)
                /locus_tag="OSJNBb0058B20.4"
     CDS        complement(18432. .18917)
                /locus_tag="OSJNBb0058B20.4"
                /codon_start=1
                /product="putative disease resistance protein"
                /protein_id="AAP53629.1"
                /db_xref="GI:31431923"
                /translation="MAASHRAILALILLFSSSAATIAAAQPAAADGGRTTHLHFYMHE
                FFNGGPNGSMARVAPPARSGGSLFGPVSVVDDALREGADPASRLVGRAQGLAAGTSLA
                DGSVTTMLDFVFTDGRTRAARWRRSASGCRCPAARRWSGRWSAARARSGWREGTRSAG
                R"
     gene       complement(<20841. .>21135)
                /locus_tag="OSJNBb0058B20.3"
                /note="unknown protein"
     mRNA       complement(join(<20841. .20929,20991. .>21135))
                /locus_tag="OSJNBb0058B20.3"
     CDS        complement(join(20841. .20929,20991. .21135))
                /locus_tag="OSJNBb0058B20.3"
                /codon_start=1
                /product="unknown protein"
                /protein_id="AAP53630.1"
                /db_xref="GI:31431924"
                /translation="MAAGPPPPGAAPPPWPAAACHGRRPVTGHGPAGRQMGRGGWAGK
                ERGGGRGEEEVEDRGEDRMRKGRRWRGWIRMRM"
     gene       complement(<23158. .>24838)
                /locus_tag="OSJNBb0058B20.2"
                /note="Similar to  (AC021891) putative retroelement (Oryza
                sativa) E value e 102 Identities 207 406 (of subject)"
     mRNA       complement(join(<23158. .23553,23818. .24151,24768. .>24838))
                /locus_tag="OSJNBb0058B20.2"
     CDS        complement(join(23158. .23553,23818. .24151,24768. .24838))
                /locus_tag="OSJNBb0058B20.2"
                /codon_start=1
                /product="putative retroelement"
                /protein_id="AAP53631.1"
                /db_xref="GI:31431925"
                /translation="MLALIARAIEMGYLFRRTRQKRERIPPALHEGRLIGKGTVGKKR
                FYDDDYKAVAEAHSSVLQQLAILELYIDEKHKNEIRACNPRRSATWISKEQKHKFPEL
                LKDQNLSFGESLEEKNIAEAGKWPNLPRDLMASIVEKEKIGEFVARREHDELTAALGT
                AEHSGRVRGKSSRTSWNVGFPKESKSYKKRDTYNTKLREKKVTQQVTQQFYRLAAQHP
```

```
Alignment Scores:
Pred. No.:              9.24            Length:         324367
Score:                  119.00          Matches:        46
Percent Similarity:     44.7%           Conservative:   22
Best Local Similarity:  30.3%           Mismatches:     62
Query Match:            10.0%           Indels:         22
DB:                     15              Gaps:           7

US-10-757-745-2_COPY_121_347 (1-227) x AE017091 (1-324367)

Qy         59 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 78
              |||                 ::::::   |||::::::   :::                |||||
Db     156699 TTCAATTGCAAACAAATGCAGATGAGCAAATTACCGGTGAGCGAATCGAACCCCATCCCA 156758

Qy         79 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 98
              |||   |||      :::|||:::      |||||||||   |||      :::|||
Db     156759 TTT---TCGAAATCAATAATGAAGAGGGAGCTATGCGTAGCGGTCGTCAAAACGGGAGAG 156815

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 |||   :::   ||||||||||||||||||
Db     156816 ATCCATTTGGCTGTGGGCACAAGTCACCTAGAGAGCCCC------TGCCCTCTACCTCCT 156869

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
              :::       ||||||      |||::::::   :::|||   :::|||      :::
Db     156870 CTGTGGGATCTGAAGTACAGCGAGAAGCGAGTAGCTCAGGCAAAGCAGTCTCTGGAGATT 156929

Qy        138 ----------------------IlePheAlaGlyAspThrAsnLeuArgAspArgGluVal 150
                                   ||||||   ||||||   |||      |||:::   |||
Db     156930 CTTGGACAATTGCGCAATGCAATCTTCTGCGGCGACATGAACTGGGAAGACAAA---GTA 156986

Qy        151 ThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 170
                        ||||||:::      :::|||   |||      |||   ||||||
Db     156987 GATGGTCCATTCCCTCTACCTGACGGCTGGATTGATGCTTGGGTGGAGCTG---AAGCCA 157043

Qy        171 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr------ 188
                    :::|||:::||||||:::   |||:::   |||
Db     157044 GGTGACAATGGTTGGACCTATGATACGAAAGCCAATGCCATGCTGTCAGCCAACTTCAAG 157103

Qy        189 ----------------------AlaAlaCysLysLeu 193
                                   |||||||||
Db     157104 CAGCAGAAGAGGCCAGACCAGTTCGTGTGCAAGTTG 157139

RESULT 65
AE016789
LOCUS       AE016789               304517 bp    DNA     linear   BCT 12-DEC-2002
DEFINITION  Pseudomonas putida KT2440 section 16 of 21 of the complete genome.
ACCESSION   AE016789 AE015451
VERSION     AE016789.1  GI:26557034
KEYWORDS    .
SOURCE      Pseudomonas putida KT2440
  ORGANISM  Pseudomonas putida KT2440
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 304517)
  AUTHORS   Nelson,K., Paulsen,I., Weinel,C., Dodson,R., Hilbert,H., Fouts,D.,
```

```
            Gill,S., Pop,M., Martins Dos Santos,V., Holmes,M., Brinkac,L.,
            Beanan,M., DeBoy,R., Daugherty,S., Kolonay,J., Madupu,R.,
            Nelson,W., White,O., Peterson,J., Khouri,H., Hance,I., Lee,P.,
            Holtzapple,E., Scanlan,D., Tran,K., Moazzez,A., Utterback,T.,
            Rizzo,M., Lee,K., Kosack,D., Moestl,D., Wedler,H., Lauber,J.,
            Hoheisel,J., Straetz,M., Heim,S., Kiewitz,C., Eisen,J., Timmis,K.,
            Duesterhoft,A., Tummler,B. and Fraser,C.
  TITLE     Complete genome sequence and comparative analysis of the
            metabolically versatile Pseudomonas putida KT2440
  JOURNAL   Environ. Microbiol. 4 (12), 799-808 (2002)
REFERENCE   2  (bases 1 to 304517)
  AUTHORS   Nelson,K., Paulsen,I., Weinel,C., Dodson,R., Hilbert,H., Fouts,D.,
            Gill,S., Pop,M., Martins Dos Santos,V., Holmes,M., Brinkac,L.,
            Beanan,M., DeBoy,R., Daugherty,S., Kolonay,J., Madupu,R.,
            Nelson,W., White,O., Peterson,J., Khouri,H., Hance,I., Lee,P.,
            Holtzapple,E., Scanlan,D., Tran,K., Moazzez,A., Utterback,T.,
            Rizzo,M., Lee,K., Kosack,D., Moestl,D., Wedler,H., Lauber,J.,
            Hoheisel,J., Straetz,M., Heim,S., Kiewitz,C., Eisen,J., Timmis,K.,
            Duesterhoft,A., Tummler,B. and Fraser,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .304517
                     /organism="Pseudomonas putida KT2440"
                     /mol_type="genomic DNA"
                     /strain="KT2440"
                     /db_xref="taxon:160488"
     gene            100. .1470
                     /gene="purB"
                     /locus_tag="PP4016"
     CDS             100. .1470
                     /gene="purB"
                     /locus_tag="PP4016"
                     /note="similar to GB:X06233, GB:M21064, GB:A12031,
                     GB:A12029, SP:P06702, PID:34771, PID:386958, PID:490104,
                     PID:490108, and PID:516621; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="adenylosuccinate lyase"
                     /protein_id="AAN69610.1"
                     /db_xref="GI:24985715"
                     /translation="MQLSSLTAVSPVDGRYAGKTQALRPIFSEFGLIRFRALVEVRWL
                     QRLAAHPQIGEVPAFSAEANALLDSLATDFKLEHAERVKEIERTTNHDVKAIEYLLKE
                     QAAKLPELAKVSEFIHFACTSEDINNLSHALMLRAGRDEVLLPLMRQVADAIRALAHA
                     HAEVPMLSRTHGQPASPTTLGKELANVVYRLERQIAQVAAVPLLGKINGAVGNYNAHL
                     SAYSQIDWEENARAFIEDELGLQFNPYTTQIEPHDYIAELFDAIARFNTILIDFDRDV
                     WGYISLGYFKQKTVAGEIGSSTMPHKVNPIDFENSEGNLGIANALFQHLASKLPISRW
                     QRDLTDSTVLRNLGVGFAHSVIAFEASLKGIGKLEVNEARIAADLDACWEVLAEPIQT
                     VMRRFNIENPYEKLKELTRGKGITPEALLTFIDGLDMPAEAKAELKQLTPATYIGNAA
                     AQAKRI"
     gene            1546. .2718
                     /locus_tag="PP4017"
     CDS             1546. .2718
                     /locus_tag="PP4017"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAN69611.1"
                     /db_xref="GI:24985716"
                     /translation="MNMNPDTPLQLLGGISAREFMRDYWQKKPLLVRQAFPDFISPID
                     PDELAGLALEDEVESRIVLEHGAHPWELRRGPFNEDTFAELPEKDWTLLVQAVDQFVP
                     EVAELLENFRFLPSWRIDDVMISFAAPGGSVGPHFDNYDVFLLQGHGQRNWKIGQMCS
                     SDSPLLEHADLRILAEFEQSGEWTLEPGDMLYLPPRLAHYGVAVDDCLTYSVGFRAPS
                     AAEVLTHFTDFLGQFLPDEERYSDADAQPVSDPHQIQHDALDRLKALLDKHMGDKDLL
                     LTWFGQFMTEPRYPEQIVGEELSEEELINALEQGAILIRNPSARLAWSEFEDDLLLFA
                     SGRSCPLPGKLRELLKLVCAADALHIDNLAEWLQDEDGLMLVQQLVKQGSLGFANE"
     gene            2711. .3139
                     /locus_tag="PP4018"
     CDS             2711. .3139
                     /locus_tag="PP4018"
                     /note="identified by match to PFAM protein family HMM
                     PF00583"
                     /codon_start=1
                     /transl_table=11
                     /product="acetyltransferase, GNAT family"
                     /protein_id="AAN69612.1"
                     /db_xref="GI:24985717"
                     /translation="MNKISVRLADWHKDNADIRRIREAVFVAEQHIPPELEFDSEDQD
                     ALHFLALEGDYPIGTARLLADGTIGRISVLKDWRGLKVGDALVNAVIVEAQNRDLKQQ
                     MLSAQVHATPFYERLGFRVVSEEFLEAGIPHVDMVRDSRA"
     gene            3307. .5268
                     /gene="topB"
                     /locus_tag="PP4019"
     CDS             3307. .5268
                     /gene="topB"
                     /locus_tag="PP4019"
                     /note="similar to GP:9106851, and SP:P14294; identified by
                     sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA topoisomerase III"
                     /protein_id="AAN69613.1"
                     /db_xref="GI:24985719"
                     /translation="MRLFLCEKPSQAKDIAKVLGANRKGDGCWQGTDICVTWCIGHLL
                     ETAPPDSYDERYKRWNLADLPIIPDKWKMLVKPKTASQFKAVKRLLGEARELVIATDA
                     DREGEMIARELVEHCRYRGPVQRLWLSALDDASIRKALARLLPGHETFNLYHSALGRS
                     RADWLIGMNMSRLFTLLGRQSGYQGVLPVGRVQTPTLRLVVDRDRSIANFVPVPFWAI
                     DVQLEHAGQRFNAQWRAPEDACDDQGRCLKQSLAQQAAADIGHAGTARAVKVATERVR
                     EAAPLPFDLGTLQELCSKKFGLGAQETLDIAQALYETHKLITYPRSDCGYLPQSQHAE
                     APAILAALQRADASLAPLQPYLEPQRRSRAWNDAKVSAHHGIIPTAAASDPSRLPAKH
                     KAVYTLIRARYLAQFLPNHEYDRTQADFDCAGHALRAVGKQIVEPGWRRALPEALTPA
                     KGREAPPAQVLPALREGQDCNVHGLQLKDLWTQPPKPFTEGDLIKAMKNVAKLVDDPR
                     LKQKLKETTGIGTEATRASIIQGLLDRGYLVKNGKALSATPAAFSLIDAVPRAIADPG
                     TTAIWEQALDMVQSGEMPLEEFVARQSAWMGKLVERCSGMRMTISGPAAGAAPPWKKK
                     RRGGGKSKATGSKATPSKPRQARKKRAPL"
     gene            5414. .6703
                     /locus_tag="PP4020"
     CDS             5414. .6703
                     /locus_tag="PP4020"
                     /note="similar to GP:9951625; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="oxidoreductase, putative"
                     /protein_id="AAN69614.1"
                     /db_xref="GI:24985720"
                     /translation="MFKQSAQHVASYYAQTYPTNIPLRPTLQGTQDADVLIIGAGFSG
                     LHTALRLTQAGKRVTLLEASRVAWAASGRNGGQALLGWSCDMPPLEQALGIERTQRLW
                     ASMCWAAEEMRELPQRHGFDIDYRQGSLWTAVLPRRVKMLEEALHDAEHKWGYDALRL
                     IGRDELPQWIDSPRYQAALYDAKGAHLNPLKLAQGLASTIEAAGGQIYEQSQVLSYQQ
                     ARDGYVARTDRGEVRCQVLVLACNAYIDRLDRGLSRRLLPVGSYQVATAPLDADFAHS
                     LLPRNSCVIDNQFVPDYFRLTPDHRLLFGGGCTYLGGIPKDVASATRPYLERVFPQLA
                     GVAIDYAWGGHIDCSIKRTPDVGREGQRYWLQGFSGHGVLPTLAAARAVSDAILGDED
                     LLALYQGIDNGRFPGGDLLAAPLEAAAKAWYRMRDRV"
     gene            6876. .7706
                     /gene="est"
                     /locus_tag="PP4021"
     CDS             6876. .7706
                     /gene="est"
                     /locus_tag="PP4021"
                     /note="similar to GB:U07989, SP:P01596, SP:P01602,
                     SP:P01604, SP:P01834, PID:1335148, PID:1619960,
                     PID:185945, PID:185947, PID:185949, PID:468243,
                     PID:470602, PID:470608, PID:470612, PID:561696,
                     PID:561705, PID:563650, PID:619427, PID:619601,
                     PID:619602, PID:619613, PID:619615, PID:619616,
                     PID:632988, PID:642582, PID:663000, PID:663001,
                     PID:722576, PID:758085, PID:762937, PID:790811,
                     PID:860994, PID:896278, GB:U07989, SP:P01596, SP:P01602,
                     SP:P01604, SP:P01834, PID:1335148, PID:1619960,
                     PID:185945, PID:185947, PID:185949, PID:468243,
                     PID:470602, PID:470608, PID:470612, PID:561696,
                     PID:561705, PID:563650, PID:619427, PID:619601,
```

```
                     PID:619602, PID:619613, PID:619615, PID:619616,
                     PID:632988, PID:642582, PID:663000, PID:663001,
                     PID:722576, PID:758085, PID:762937, PID:790811,
                     PID:860994, and PID:896278; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="esterase"
                     /protein_id="AAN69615.1"
                     /db_xref="GI:24985721"
                     /translation="MSYVTTKDGVQIFYKDWGPRDAPVIHFHHGWPLSADDWDAQMLF
                     FLAEGYRVVAHDRRGHGRSSQVWNGHDMDHYADDVAAVVAHLGTQGAVHVGHSTGGGE
                     VVRYMARYPEDKVAKAVLIAAVPPLMVQTPGNPGGLPKSVFDGFQAQVASNRAQFYRD
                     VPTGPFYGYNRPGVDASEGIIGNWWRQGMIGSAKAHYDGIVAFSQTDFTEDLKGIQQP
                     VLVMHGDDDQIVPYENSGVLSAKLLPNGTLKTYEGYPHGMPTTHADVINADLLAFIQS
                     "
     gene            7826. .8617
                     /locus_tag="PP4022"
     CDS             7826. .8617
                     /locus_tag="PP4022"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved domain protein"
                     /protein_id="AAN69616.1"
                     /db_xref="GI:24985722"
                     /translation="MLIRAKYDEITQLILNVLTAMLTFKQTPDSAASLTNALRHFAYS
                     CYGYKPQAQFLLATGFNGQYIDLFSGHYQLGSYRTYNTGLMRFHSPDSDSPFDGGGLN
                     AYAYCSGDPINYHDPSGHGPVLSKKSFLPKKKTIMQRREAALKAINENPVLEAFQGIA
                     ENSVIKKELKTLSDTGKALNQFTDLMENKLISIDKQGQLKYKLMDDFESLRFQILEDE
                     VMKNDKTLNSLLSAYPRANEPEPSLSSASLSPATKNKLTRTQKGN"
     gene            8693. .9337
                     /locus_tag="PP4023"
                     /note="This region contains an authentic frame shift and
                     is not the result of a sequencing artifact; This gene is
                     interrupted by an IS element.; membrane protein, putative,
                     interruption-N; identified by match to PFAM protein family
                     HMM PF01925"
     repeat_region   9347. .11387
                     /note="IS0024"
                     /insertion_seq="ISPpu15"
     gene            9432. .9791
                     /locus_tag="PP4024"
     CDS             9432. .9791

Alignment Scores:
Pred. No.:              15.6          Length:         304517
Score:                  116.50        Matches:        55
Percent Similarity:     41.6%         Conservative:   36
Best Local Similarity:  25.1%         Mismatches:     69
Query Match:            9.8%          Indels:         59
DB:                     1             Gaps:           11

US-10-757-745-2_COPY_121_347 (1-227) x AE016789 (1-304517)

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys------ 43
              |||:::::::|||||||||||||||||:::      :::      :::      :::
Db      48910 GACATTGTCTTCCTTCAGGAAGTGCTCGGCAGCCACGATCGCCACGCCGCGCGCTACCCC 48969

Qy         44 ---------ArgSerSerAsnTyrGluIleIleThr----------------------- 52
                    ::::::|||   ||||||   :::
Db      48970 GGCTGGCCGCAGACCTCGCAGTACGAATTCCTCGCCGACAGCATGTGGAGCGACTTTGCC 49029

Qy         53 --------------------------GlyHisGluGluGlyTyrPheThrAlaIleMet 63
                                        ||||||        :::      |||::::::
Db      49030 TATGGCCGCAACGCGGTCTACCCTGACGGCCAC---------CACGGCAATGCGTTGCTG 49080

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIle---IleProPheProSerThr 82
              |||          ::::::   ::::::   ::::::   |||      |||
Db      49081 TCCAAATACCCGATCATCGAGCACCGCAACCTGGACGTGTCGATCACCGGCCCCGAA--- 49137

Qy         83 LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly------------Asn 98
                 |||   ||||||          :::   |||   |||              :::
Db      49138 ------CGCCGCGGCCTGTTGCACTGTGTGCTGGAGGTGCCGGGGCAAGGCCAGGTGCAT 49191

Qy         99 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 118
                 :::|||:::      |||   |||||||||                      |||
Db      49192 GCGGTCTGCGTGCACCTGTCGTTGCTGGAAAGC------------------CACCGGCAA 49233

Qy        119 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 138
                |||||:::::::::   |||   :::          |||    :::|||   ||||||
Db      49234 AAGCAGCTGCAGTTGCTGCGTAAACTGCTGGACGCCCTGCCGGCCAACGCCCCGGTGATC 49293

Qy        139 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 158
              |||||||||   |||                  |||       ||||||
Db      49294 ATCGCCGGCGACTTCAACGACTGGAAAAGCCATGGCAACCGCACCCTTGGCCTGCAACGC 49353

Qy        159 AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAsp 178
              ::::::   :::   :::|||                :::|||
Db      49354 GACTTGCACGAAGCCTTCGAG--------------CGCCAC------------------ 49380

Qy        179 ThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys---------LeuArgPhe 195
                    :::   :::|||      |||          :::         ||||||
Db      49381 ---------CACGGCCACCTCGCGCGCACCTACCCGGCCCGCCTGCCGCTGCTACGCCTG 49431

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeu 214
              |||||||||:::   |||   |||          |||         ||||||   |||
Db      49432 GACCGCATATACCTGCGTAACGCCGAAAGCCATGGG---------CCGCGGATTCTG 49479

RESULT 66
AC145862/c
LOCUS       AC145862              182686 bp    DNA     linear   HTG 01-AUG-2003
DEFINITION  Pan troglodytes chromosome UNK clone RP43-39B14, *** SEQUENCING IN
            PROGRESS ***, 34 unordered pieces.
ACCESSION   AC145862
VERSION     AC145862.1  GI:33386792
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 182686)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Pan troglodytes clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 182686)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT
            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information --------------
            Center project name: C_PT039B14
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 34 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1      1233: contig of 1233 bp in length
            *    1234      1333: gap of unknown length
            *    1334      2405: contig of 1072 bp in length
            *    2406      2505: gap of unknown length
            *    2506      4956: contig of 2451 bp in length
            *    4957      5056: gap of unknown length
            *    5057      6720: contig of 1664 bp in length
            *    6721      6820: gap of unknown length
            *    6821      9202: contig of 2382 bp in length
```

```
            *     9203     9302: gap of unknown length
            *     9303    12425: contig of 3123 bp in length
            *    12426    12525: gap of unknown length
            *    12526    14496: contig of 1971 bp in length
            *    14497    14596: gap of unknown length
            *    14597    16234: contig of 1638 bp in length
            *    16235    16334: gap of unknown length
            *    16335    19053: contig of 2719 bp in length
            *    19054    19153: gap of unknown length
            *    19154    21198: contig of 2045 bp in length
            *    21199    21298: gap of unknown length
            *    21299    23598: contig of 2300 bp in length
            *    23599    23698: gap of unknown length
            *    23699    26628: contig of 2930 bp in length
            *    26629    26728: gap of unknown length
            *    26729    28984: contig of 2256 bp in length
            *    28985    29084: gap of unknown length
            *    29085    31641: contig of 2557 bp in length
            *    31642    31741: gap of unknown length
            *    31742    34105: contig of 2364 bp in length
            *    34106    34205: gap of unknown length
            *    34206    37534: contig of 3329 bp in length
            *    37535    37634: gap of unknown length
            *    37635    40872: contig of 3238 bp in length
            *    40873    40972: gap of unknown length
            *    40973    44778: contig of 3806 bp in length
            *    44779    44878: gap of unknown length
            *    44879    49432: contig of 4554 bp in length
            *    49433    49532: gap of unknown length
            *    49533    54283: contig of 4751 bp in length
            *    54284    54383: gap of unknown length
            *    54384    58526: contig of 4143 bp in length
            *    58527    58626: gap of unknown length
            *    58627    62047: contig of 3421 bp in length
            *    62048    62147: gap of unknown length
            *    62148    65590: contig of 3443 bp in length
            *    65591    65690: gap of unknown length
            *    65691    68906: contig of 3216 bp in length
            *    68907    69006: gap of unknown length
            *    69007    72340: contig of 3334 bp in length
            *    72341    72440: gap of unknown length
            *    72441    78465: contig of 6025 bp in length
            *    78466    78565: gap of unknown length
            *    78566    85703: contig of 7138 bp in length
            *    85704    85803: gap of unknown length
            *    85804    94817: contig of 9014 bp in length
            *    94818    94917: gap of unknown length
            *    94918   105263: contig of 10346 bp in length
            *   105264   105363: gap of unknown length
            *   105364   114262: contig of 8899 bp in length
            *   114263   114362: gap of unknown length
            *   114363   125451: contig of 11089 bp in length
            *   125452   125551: gap of unknown length
            *   125552   140596: contig of 15045 bp in length
            *   140597   140696: gap of unknown length
            *   140697   160689: contig of 19993 bp in length
            *   160690   160789: gap of unknown length
            *   160790   182686: contig of 21897 bp in length.
FEATURES             Location/Qualifiers
     source          1. .182686
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="UNK"
                     /clone="RP43-39B14"
     misc_feature    1. .1233
                     /note="assembly_name:Contig47"
     gap             1234. .1333
                     /estimated_length=unknown
     misc_feature    1334. .2405
                     /note="assembly_name:Contig48"
     gap             2406. .2505
                     /estimated_length=unknown
     misc_feature    2506. .4956
                     /note="assembly_name:Contig49"
     gap             4957. .5056
                     /estimated_length=unknown
     misc_feature    5057. .6720
                     /note="assembly_name:Contig50"
     gap             6721. .6820
                     /estimated_length=unknown
     misc_feature    6821. .9202
                     /note="assembly_name:Contig52"
     gap             9203. .9302
                     /estimated_length=unknown
     misc_feature    9303. .12425
                     /note="assembly_name:Contig53"
     gap             12426. .12525
                     /estimated_length=unknown
     misc_feature    12526. .14496
                     /note="assembly_name:Contig54"
     gap             14497. .14596
                     /estimated_length=unknown
     misc_feature    14597. .16234
                     /note="assembly_name:Contig55"
     gap             16235. .16334
                     /estimated_length=unknown
     misc_feature    16335. .19053
                     /note="assembly_name:Contig56"
     gap             19054. .19153
                     /estimated_length=unknown
     misc_feature    19154. .21198
                     /note="assembly_name:Contig57"
     gap             21199. .21298
                     /estimated_length=unknown
     misc_feature    21299. .23598
                     /note="assembly_name:Contig58"
     gap             23599. .23698
                     /estimated_length=unknown
     misc_feature    23699. .26628
                     /note="assembly_name:Contig59"
     gap             26629. .26728
                     /estimated_length=unknown
     misc_feature    26729. .28984
                     /note="assembly_name:Contig60"
     gap             28985. .29084
                     /estimated_length=unknown
     misc_feature    29085. .31641
                     /note="assembly_name:Contig61"
     gap             31642. .31741
                     /estimated_length=unknown
     misc_feature    31742. .34105
                     /note="assembly_name:Contig62"
     gap             34106. .34205
                     /estimated_length=unknown
     misc_feature    34206. .37534
                     /note="assembly_name:Contig63"
     gap             37535. .37634
                     /estimated_length=unknown
     misc_feature    37635. .40872
                     /note="assembly_name:Contig64"
     gap             40873. .40972
                     /estimated_length=unknown
     misc_feature    40973. .44778
                     /note="assembly_name:Contig65"
     gap             44779. .44878
                     /estimated_length=unknown
     misc_feature    44879. .49432
                     /note="assembly_name:Contig66"
     gap             49433. .49532
                     /estimated_length=unknown
     misc_feature    49533. .54283
                     /note="assembly_name:Contig67"
     gap             54284. .54383
                     /estimated_length=unknown
     misc_feature    54384. .58526
```

```
                 /note="assembly_name:Contig68"
     gap         58527. .58626
                 /estimated_length=unknown
     misc_feature 58627. .62047
                 /note="assembly_name:Contig69"
     gap         62048. .62147
                 /estimated_length=unknown
     misc_feature 62148. .65590

Alignment Scores:
Pred. No.:              9.72          Length:        182686
Score:                116.00          Matches:       52
Percent Similarity:    48.2%          Conservative:  42
Best Local Similarity: 26.7%          Mismatches:    67
Query Match:           9.7%           Indels:        34
DB:                    14             Gaps:          9

US-10-757-745-2_COPY_121_347 (1-227) x AC145862 (1-182686)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::         :::         :::   :::::::::
Db     143391 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAAGACACAGACTGGTAAATTGGATAAAG 143332

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||     :::||||||                  ::::::|||
Db     143331 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 143290

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
              |||   |||   ::::::          |||   |||       :::   |||
Db     143289 AAGTGCAGAGACACACATAGGCTCAAAATAAAGGGACGGAGGAAGATCTACCAAGCAAAT 143230

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                 ||||||:::|||||||   :::         :::         |||
Db     143229 GGAAAACAAAAAAAGGCAAGGGTTGCAATCCTAGTCTCTGATAAAACAGATTTTAAACCA 143170

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::         :::::   |||   |||:::
Db     143169 ACAAAGATCAAAAGAGACAAAGAAAGCCATTACATAATGGTAAAGGGATCAATTCAACAA 143110

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||             |||   |||
Db     143109 GAAGAGCTAACTATTCTAAATATGTATGCACCCACTACA---------GGAGCACCCAGA 143059

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                :::|||   ||||||      :::|||         :::|||   |||:::
Db     143058 TTC------ATAAAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 143005

Qy        138 IlePheAlaGlyAsp----------------------ThrAsnLeuArgAspArgGlu 149
              |||      ||||||                     ||||||   |||||||:::
Db     143004 ATAATG---GGAGATTTAACACCCCACTGTCAACATTAGACAAATCAACGAGACAGAAAG 142948

Qy        150 ValThrArgCysGly-GlyLeu---------------ProAsn-AsnIleValAspValT 164
              :::||||||      |||:::             |||::: :::::::::|||::::
Db     142947 TTAACAAGGATATCCAGGAATTGAACTCAACTCTGCACCAAGTGGACCTAATAGACATCT 142888

Qy        164 rpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::    |||   :::||||||      :::||||||:::
Db     142887 ACAGAACTCTC---CAGCCCAAATCAAGAGAATATACATTC 142850

RESULT 67
AC148939/c
LOCUS       AC148939              187205 bp    DNA     linear   HTG 19-APR-2005
DEFINITION  Pan troglodytes chromosome 7 clone CH251-51M11, WORKING DRAFT
            SEQUENCE, 4 unordered pieces.
ACCESSION   AC148939
VERSION     AC148939.7  GI:62734900
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 187205)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Pan troglodytes clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 187205)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2004) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 187205)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2005) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT     On Apr 19, 2005 this sequence version replaced gi:59933409.

            ---------------- Genome Center ----------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information ------------------
            Center project name: C_AB0051M11
            ----------------------------------------------------
            ------------- Summary Statistics  -------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 184666 bases at least Q40
            Consensus quality: 185209 bases at least Q30
            Consensus quality: 185512 bases at least Q20
            ----------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1132: contig of 1132 bp in length
            *     1133     1232: gap of unknown length
            *     1233     2610: contig of 1378 bp in length
            *     2611     2710: gap of unknown length
            *     2711    97398: contig of 94688 bp in length
            *    97399    97498: gap of unknown length
            *    97499   187205: contig of 89707 bp in length.
FEATURES            Location/Qualifiers
     source         1. .187205
                    /organism="Pan troglodytes"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9598"
                    /chromosome="7"
                    /clone="CH251-51M11"
     misc_feature   1. .1132
                    /note="assembly_name:Contig49"
     gap            1133. .1232
                    /estimated_length=unknown
     misc_feature   1233. .2610
                    /note="assembly_name:Contig61"
     gap            2611. .2710
                    /estimated_length=unknown
     misc_feature   2711. .97398
                    /note="assembly_name:Contig64"
     gap            97399. .97498
                    /estimated_length=unknown
     misc_feature   97499. .187205
                    /note="assembly_name:Contig65"
ORIGIN

Alignment Scores:
Pred. No.:              10            Length:        187205
```

```
Score:                 116.00        Matches:       52
Percent Similarity:    48.2%         Conservative:  42
Best Local Similarity: 26.7%         Mismatches:    67
Query Match:           9.7%          Indels:        34
DB:                    14            Gaps:          9

US-10-757-745-2_COPY_121_347 (1-227) x AC148939 (1-187205)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db     159561 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAAGACACAGACTGGTAAATTGGATAAAG 159502

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::||||||                    :::::::|||
Db     159501 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 159460

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
              |||   |||   ::::::         |||   |||       :::   |||
Db     159459 AAGTGCAGAGACACACATAGGCTCAAAATAAAGGGACGGAGGAAGATCTACCAAGCAAAT 159400

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||::||||||   :::         :::         |||
Db     159399 GGAAAACAAAAAAAGGCAAGGGTTGCAATCCTAGTCTCTGATAAAACAGATTTTAAACCA 159340

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::             ::::::   |||   |||:::
Db     159339 ACAAAGATCAAAAGAGACAAAGAAAGCCATTACATAATGGTAAAGGGATCAATTCAACAA 159280

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||             |||   |||
Db     159279 GAAGAGCTAACTATTCTAAATATGTATGCACCCACTACA---------GGAGCACCCAGA 159229

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   ||||||      :::|||          :::|||   |||:::
Db     159228 TTC------ATAAAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 159175

Qy        138 IlePheAlaGlyAsp-----------------------ThrAsnLeuArgAspArgGlu 149
              |||      ||||||                     ||||||   |||||||||:::
Db     159174 ATAATG---GGAGATTTAACACCCCACTGTCAACATTAGACAAATCAACGAGACAGAAAG 159118

Qy        150 ValThrArgCysGly-GlyLeu---------------ProAsn-AsnIleValAspValT 164
              :::||||||       |||:::               |||::: ::::::::::|||::::
Db     159117 TTAACAAGGATATCCAGGAATTGAACTCAACTCTGCACCAAGTGGACCTAATAGACATCT 159058

Qy        164 rpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::      |||   :::||||||      :::||||||:::
Db     159057 ACAGAACTCTC---CAGCCCAAATCAAGAGAATATACATTC 159020

RESULT 68
AC150749
LOCUS       AC150749               237211 bp    DNA     linear   PRI 27-APR-2005
DEFINITION  Pan troglodytes BAC clone CH251-553P23 from chromosome 7, complete
            sequence.
ACCESSION   AC150749
VERSION     AC150749.3  GI:62198463
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 237211)
  AUTHORS   Tomlinson,C., Kozlowicz,A. and Haglund,K.
  TITLE     The sequence of Pan troglodytes BAC clone CH251-553P23
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 237211)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-AUG-2004) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 237211)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAR-2005) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 237211)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-APR-2005) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   5  (bases 1 to 237211)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2005) Washington University School of Medicine,
            Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO
            63108, USA
COMMENT     On Apr 5, 2005 this sequence version replaced gi:59933396.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: C_AB0553P23


            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The CHORI-251 Chimpanzee BAC library has been constructed at the
            Children's Hospital Oakland Research Institute, BACPAC Resources,
            by Dr. Baoli Zhu.  DNA was isolated from white blood cells obtained
            from a male chimpanzee (Pan troglodytes, 'Clint', Yerkes #C0471;
            birthdate:6-6-80).  The clone and detailed information can be
            obtained from Pieter de Jong and co-workers at
            http://www.bacpac.chori.org.

            This sequence is the entire insert of the clone.
FEATURES              Location/Qualifiers
     source           1. .237211
                      /organism="Pan troglodytes"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9598"
                      /chromosome="7"
                      /clone="CH251-553P23"
                      /clone_lib="CHORI251"
     unsure           15317. .15318
                      /note="Sequence derived from one plasmid subclone."
     unsure           15552. .15585
                      /note="Sequence derived from one plasmid subclone."
     unsure           34920. .35171
                      /note="Sequence derived from one plasmid subclone."
     unsure           35260. .35316
                      /note="Sequence derived from one plasmid subclone."
     unsure           48319. .48350
                      /note="Sequence derived from one plasmid subclone."
     unsure           86305. .86793
                      /note="Sequence derived from one plasmid subclone."
     unsure           86833. .86969
                      /note="Unresolved simple sequence repeat."
     unsure           89495. .89499
                      /note="Sequence derived from one plasmid subclone."
```

```
     unsure          89702. .89797
                     /note="Sequence derived from one plasmid subclone."
     unsure          89907. .89969
                     /note="Sequence derived from one plasmid subclone."
     unsure          117320. .117327
                     /note="Sequence derived from one plasmid subclone."
     misc_feature    136745. .136813
                     /note="Sequence derived from PCR product of project DNA"
     unsure          146723
                     /note="Unresolved bases"
     unsure          146742
                     /note="Unresolved bases"
     unsure          146747
                     /note="Unresolved bases"
     misc_feature    185394. .185590
                     /note="Sequence derived from PCR product of project DNA"
     unsure          186584. .186603
                     /note="Sequence derived from one plasmid subclone."
     unsure          188541. .188554
                     /note="Sequence derived from one plasmid subclone."
     misc_feature    189006. .189512
                     /note="Sequence derived from PCR product of project DNA"
ORIGIN

Alignment Scores:
Pred. No.:              13.2            Length:         237211
Score:                  116.00          Matches:        52
Percent Similarity:     48.2%           Conservative:   42
Best Local Similarity:  26.7%           Mismatches:     67
Query Match:            9.7%            Indels:         34
DB:                     8               Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x AC150749 (1-237211)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db     113785 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAAGACACAGACTGGTAAATTGGATAAAG 113844

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::||||||                  ::::::|||
Db     113845 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 113886

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
              |||   |||   ::::::          |||   |||       :::   |||
Db     113887 AAGTGCAGAGACACACATAGGCTCAAAATAAAGGGACGGAGGAAGATCTACCAAGCAAAT 113946

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::||||||    :::         :::         |||
Db     113947 GGAAAACAAAAAAAGGCAAGGGTTGCAATCCTAGTCTCTGATAAAACAGATTTTAAACCA 114006

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::         ::::::   |||   ||:::
Db     114007 ACAAAGATCAAAAGAGACAAAGAAAGCCATTACATAATGGTAAAGGGATCAATTCAACAA 114066

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||            |||   |||
Db     114067 GAAGAGCTAACTATTCTAAATATGTATGCACCCACTACA---------GGAGCACCCAGA 114117

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                 :::|||  ||||||      :::|||         :::||||   |||:::
Db     114118 TTC------ATAAAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 114171

Qy        138 IlePheAlaGlyAsp----------------------ThrAsnLeuArgAspArgGlu 149
              |||      ||||||                     ||||||   |||||||||:::
Db     114172 ATAATG---GGAGATTTAACACCCCACTGTCAACATTAGACAAATCAACGAGACAGAAAG 114228

Qy        150 ValThrArgCysGly-GlyLeu---------------ProAsn-AsnIleValAspValT 164
              :::||||||       |||:::               |||::: ::::::::::|||::::
Db     114229 TTAACAAGGATATCCAGGAATTGAACTCAACTCTGCACCAAGTGGACCTAATAGACATCT 114288

Qy        164 rpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::      |||   :::||||||      :::||||||:::
Db     114289 ACAGAACTCTC---CAGCCCAAATCAAGAGAATATACATTC 114326
```

```
RESULT 69
BX640415/c
LOCUS       BX640415               347071 bp    DNA     linear   BCT 17-APR-2005
DEFINITION  Bordetella pertussis strain Tohama I, complete genome; segment
            5/12.
ACCESSION   BX640415 BX470248
VERSION     BX640415.1  GI:33572084
KEYWORDS    complete genome.
SOURCE      Bordetella pertussis Tohama I
  ORGANISM  Bordetella pertussis Tohama I
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Bordetella.
REFERENCE   1
  AUTHORS   Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
            Harris,D.E., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,
            Mungall,K.L., Cerdeno-Tarraga,A.M., Temple,L., James,K., Harris,B.,
            Quail,M.A., Achtman,M., Atkin,R., Baker,S., Basham,D., Bason,N.,
            Cherevach,I., Chillingworth,T., Collins,M., Cronin,A., Davis,P.,
            Doggett,J., Feltwell,T., Goble,A., Hamlin,N., Hauser,H.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Norberczak,H.,
            O'Neil,S., Ormond,D., Price,C., Rabbinowitsch,E., Rutter,S.,
            Sanders,M., Saunders,D., Seeger,K., Sharp,S., Simmonds,M.,
            Skelton,J., Squares,R., Squares,S., Stevens,K., Unwin,L.,
            Whitehead,S., Barrell,B.G. and Maskell,D.J.
  TITLE     Comparative analysis of the genome sequences of Bordetella
            pertussis, Bordetella parapertussis and Bordetella bronchiseptica
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 347071)
  AUTHORS   Sebaihia,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-2003) Submitted on behalf of the Pathogen
            Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA, E-mail: ms5@sanger.ac.uk
FEATURES             Location/Qualifiers
     source          1. .347071
                     /organism="Bordetella pertussis Tohama I"
                     /mol_type="genomic DNA"
                     /strain="Tohama I"
                     /db_xref="taxon:257313"
     gene            100. .3071
                     /locus_tag="BP1322"
                     /pseudo
     CDS             join(100. .1953,1953. .3071)
                     /locus_tag="BP1322"
                     /note="Pseudogene. This CDS appears to have a frameshift
                     mutation following codon 618. The sequence has been
                     checked and believed to be correct. Similar, over its
                     C-terminal region, to Pseudomonas syringae histidine
                     protein kinase-response regulator hybrid protein CvgsY
                     TR:Q9WWH5 (EMBL:AF133263) (671 aa) fasta scores: E():
                     2.6e-60, 38.48% id in 569 aa, and to Rhizobium loti
                     two-component sensor histidine kinase mll8067 TR:Q984C1
                     (EMBL:AP003013) (1386 aa) fasta scores: E(): 7.7e-49,
                     41.6% id in 387 aa"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
                     /product="two-component sensor histidine kinase
                     (Pseudogene)"
                     /db_xref="PSEUDO:CAE41616.1"
     misc_feature    226. .645
                     /locus_tag="BP1322"
                     /note="HMMPfam hit to PF01590, GAF domain"
                     /pseudo
     misc_feature    766. .849
                     /locus_tag="BP1322"
                     /note="HMMPfam hit to PF00989, PAS domain"
                     /pseudo
     misc_feature    1546. .1740
                     /locus_tag="BP1322"
                     /note="HMMPfam hit to PF00989, PAS domain"
                     /pseudo
```

misc_feature join(1942. .1953,1953. .2147)
/locus_tag="BP1322"
/note="HMMPfam hit to PF00512, His Kinase A (phosphoacceptor) domain"
/pseudo
variation 1954
/locus_tag="BP1322"
/note="C(-)G in pertussis; C(T)G in parapertussis and bronchiseptica"
misc_feature 2268. .2639
/locus_tag="BP1322"
/note="HMMPfam hit to PF02518, Histidine kinase-, DNA gyrase B-, phytochrome-like ATPase"
/pseudo
misc_feature 2703. .3041
/locus_tag="BP1322"
/note="HMMPfam hit to PF00072, Response regulator receiver domain"
/pseudo
gene complement(3062. .5605)
/locus_tag="BP1323"
CDS complement(3062. .5605)
/locus_tag="BP1323"
/note="Similar to Streptomyces coelicolor putative alpha amylase Scbac1a6.05C TR:Q9ADI2 (EMBL:AL589708) (805 aa) fasta scores: E(): 8.8e-23, 38.42% id in 864 aa, and to Pseudomonas aeruginosa probable glycosyl hydrolase Pa2162 TR:Q9I1V3 (EMBL:AE004643) (926 aa) fasta scores: E(): 4.1e-28, 45.05% id in 930 aa"
/codon_start=1
/transl_table=11
/product="probable alpha amylase"
/protein_id="CAE41617.1"
/db_xref="GI:33572085"
/db_xref="GOA:Q7VYK3"
/db_xref="InterPro:IPR006047"
/db_xref="UniProt/TrEMBL:Q7VYK3"
/translation="MIRATVRLQLHAGFTLDAAGAQLDYYDALGISHLYLSPVARARC
GSTHGYDVIDHAQVNPELGGLAALRRLAGAARARGMGLILDIVPNHMAAHPDNRWWRE
VLRDGPASAYADWFDIDWRPPGRTLRGRVLLPVLTCGARAALASGEIGLRAGSDGTPA
LHANGLDLPLAPASLPMAALLLHHGGTPAQWHALLRAQAYRLAWWRSAAQAINWRRFF
EIGDLVGVRVERPEVFDAVHALPLALYRAGLIDGLRIDHVDGLAEPGAYCRRLRAELA
QAGQARAARGLSAEPYVVVEKILGDKETLRTDWPVQGTTGYDFMNEASAWLHDARGAA
PLAAFWRDNAGTWPHWREQLRRIRLRLLRRHFGAERRALARSLADLPELRASAAAIDP
VLTQWLACFPVYRTYAEGGAGQAADQAWRQTAQTRANAHLDPAQQALLAALCLALDGR
MADAPHRLALRRLQQLTPPLAAKALEDTWFYRDGTLLSRNEVGPAPAQFALGGGALRR
SVAARARRHPQALLATASHDHKRGEDVRARLAVLSEMPQQWCDWARQWLADAGGSAAP
PAWADRYMLLQTLVGAWPADLAPDDQAGMAALLRRVAAWQAKALREARQRSSWTCPDL
AYETLCARYLRALADDGAALPAVAALARRLAPAGLVNSLAQTALRLGLPGVPDLYQGC
ETWDLSLVDPDNRRAVDYAARRGALASLGNGALALPELWRDGRAKQALVARALQWRRE
QPALPEHGACRWLRVSGARARHVLACLRTRGPQQALIVVPRASALALAGQTGPDRDAA
AHWRDTCVHLPADAPPWRDILTHAGHQGRALALSELLARWPVALCRVQS"
misc_feature complement(4154. .5593)
/locus_tag="BP1323"
/note="HMMPfam hit to PF00128, Alpha amylase, catalytic domain"
gene complement(join(5602. .6896,6898. .6975))
/gene="malQ"
/locus_tag="BP1324"
/note="synonym: malA"
/pseudo
CDS complement(join(5602. .6576,6576. .6896,6898. .6975))
/gene="malQ"
/locus_tag="BP1324"
/EC_number="2.4.1.25"
/note="Pseudogene. This CDS appears to contain a deletion that has lead to a shift in reading frame around codon 133, and also appears to have a frameshift mutation following codon 26. The sequence has been checked and believed to be correct. Similar to Escherichia coli 4-alpha-glucanotransferase MalQ or MalA or B3416 SWALL:MALQ_ECOLI (SWALL:P15977) (694 aa) fasta scores: E(): 2.1e-25, 38.31% id in 327 aa"
/pseudo
/codon_start=1
/transl_table=11
/product="putative 4-alpha-glucanotransferase (pseudogene)"
/db_xref="PSEUDO:CAE41618.1"
variation complement(6576. .6578)
/gene="malQ"
/locus_tag="BP1324"
/note="large deletion relative to parapertussis and bronchiseptica"
variation complement(6898. .6900)
/gene="malQ"
/locus_tag="BP1324"
/note="CGC(-) in pertussis; CGC(A) in parapertussis and bronchiseptica"
gene complement(6972. .8726)
/locus_tag="BP1326"
CDS complement(6972. .8726)
/locus_tag="BP1326"
/note="Similar to Sulfolobus shibatae maltooligosyl trehalose trehalohydrolase TreZ TR:Q9UWN9 (EMBL:AF201335) (559 aa) fasta scores: E(): 3.4e-50, 34.79% id in 500 aa, and to Pseudomonas aeruginosa probable glycosyl hydrolase Pa2164 TR:Q9I1V1 (EMBL:AE004643) (583 aa) fasta scores: E(): 3.1e-101, 53.79% id in 554 aa"
/codon_start=1
/transl_table=11
/product="probable alpha amylase"
/protein_id="CAE41619.1"
/db_xref="GI:33572086"
/db_xref="GOA:Q7VYK2"
/db_xref="InterPro:IPR004193"
/db_xref="InterPro:IPR006047"
/db_xref="UniProt/TrEMBL:Q7VYK2"
/translation="MNEPAPLGATVCAPGKIRFRLWAPSAPPALALEIEGRAPIALRP
GPDGYAQAIVDGAPGLRYRYRLAADCVVPDPASRLQDGDVHGHSVVVGPDRYPWRHTG
WRGRPWQDSVIYEAHAGLCGGYAGLRRRLPEVARLGVTLLELMPIADFPGPRNWGYDG
VLPYAPDTAYGTPDELKALIDAAHGLEMGVMLDVVYNHFGPDGNYLHRYAADFFDARA
DTPWGAAIDYRQQAVQRYFIDNALYWLRKYRFDGLRLDAVHAMGRTDWLAGLAREARA
ACPDRRVHLVVENDANQASLLEQGFDAQWNDDAHHVLHHLLTGERQGYYADYAQAPAA
LLARALGEGYVYQGQAAPTRGGRARGEPSAHLPPWAFVLFLHNHDQVGNRADGLRLTS
LLAPGSPALRAAIALQLLAPHIPLLFMGEEYGSTAPFFYFTSHGPELAAAVRAGRARE
FAASMHDCDPPPDPNDPDTYRRSCPWPPPGAERQAWFEYYRELLRLRARLLGERLTGA
RADGAAALGPAAVRAAWRLGDGARLTLYANLLPGSVIVPAPPAGQLAYATLAGAGQDL
VAGVLRGGCALALWEPAP"
misc_feature complement(7005. .8393)
/locus_tag="BP1326"
/note="HMMPfam hit to PF00128, Alpha amylase, catalytic domain"
gene complement(8723. .10843)
/gene="glgX"
/locus_tag="BP1327"
CDS complement(8723. .10843)
/gene="glgX"
/locus_tag="BP1327"
/note="Similar to Escherichia coli glycogen operon protein

Alignment Scores:

Pred. No.:	20.5	Length:	347071
Score:	116.00	Matches:	53
Percent Similarity:	38.8%	Conservative:	27
Best Local Similarity:	25.7%	Mismatches:	80
Query Match:	9.7%	Indels:	46
DB:	1	Gaps:	11

US-10-757-745-2_COPY_121_347 (1-227) x BX640415 (1-347071)

```
Qy         25 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 44
                ||||||::::::|||||||||||||||:::      :::      :::      :::
Db      26269 CCGGATCTCGTGTTCCTGCAGGAAGTGCTGGGCGAACACCAGCGGCACGCGGCGCGCCAC 26210

Qy         45 Ser--------------SerAsnTyrGluIleIleThrGly----------------- 53
```

```
            :::                 |||   ||||||   :::
Db      26209 GCGCGCTGGCCGGCGGTCTCGCAATACGAGTTCCTGGCCGACACCCTGTGGAGCGCCTAT 26150

Qy         54 --------------------HisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 66
                                  :::   :::|||:::          :::|||   |||
Db      26149 GCCTATGGCCGCAACGCCGTCTACCCCGATGGCCACCATGGCAACGCCGTATTGTCCAAG 26090

Qy         67 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 86
                 :::                 |||   :::         :::             |||
Db      26089 TATCCGATCGTCTCGCACCGCAACGAGGACCTGAGCGTGGGCGCGGCGCACGAGCCGCGC 26030

Qy         87 AsnLeuLeu---CysVal---HisValAsnValSerGlyAsnGluLeuCysLeuMetThr 104
              :::|||   ||||||      |||      :::          |||      :::
Db      26029 GGCATGCTGCACTGCGTGCTGGACGTGCGCGACGCGCCCGCGCCCCTGCATGCGCTGTGC 25970

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
               ||||||           |||      |||::::::   :::|||
Db      25969 GTGCACCTGGGCCTGCGCGAGGCGCATCGCCAGGAGCAGCTGCGGCGCCTGTGCCAGTAC 25910

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
               |||:::      |||      |||      |||      :::   |||||||||   |||
Db      25909 CTGCGC------CAGCACGTGCCGGCGCAGGCCGCGGCCCTGCTGGCGGGGGACTTCAAC 25856

Qy        145 ------------LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
                          |||      :::   |||||||||                  :::
Db      25855 GACTGGCGCCAGCGGGCCGACGCAACGCTGGCGCGCTGCGGG--------------CTG 25811

Qy        161 ValAspValTrpGluPhe---LeuGlyLysProLysHisCysGlnTyrThrThrTrpAspThr 179
                 ||||||           |||||      |||      |||::::
Db      25810 GCCGACGTCCACCGGCTGGCGCTGGGG---------CACGCGGCGCGTACCTTT------ 25766

Qy        180 GlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys---LeuArgPheAspArgIle 198
                                      |||||||||      ||||||   |||||||||
Db      25765 ------------------------CCGGCCGCCTGCCCGCTGCTGCGGCTGGACCGCATC 25730

Qy        199 PhePheArgAlaAlaAla 204
               :::   |||      |||
Db      25729 TACGTGCGCAACGTCGCC 25712
```

```
RESULT 70
BX640445
LOCUS       BX640445               348706 bp    DNA     linear   BCT 17-APR-2005
DEFINITION  Bordetella bronchiseptica strain RB50, complete genome; segment
            9/16.
ACCESSION   BX640445 BX470250
VERSION     BX640445.1  GI:33576035
KEYWORDS    complete genome.
SOURCE      Bordetella bronchiseptica RB50
  ORGANISM  Bordetella bronchiseptica RB50
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Bordetella.
REFERENCE   1
  AUTHORS   Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
            Harris,D.E., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,
            Mungall,K.L., Cerdeno-Tarraga,A.M., Temple,L., James,K., Harris,B.,
            Quail,M.A., Achtman,M., Atkin,R., Baker,S., Basham,D., Bason,N.,
            Cherevach,I., Chillingworth,T., Collins,M., Cronin,A., Davis,P.,
            Doggett,J., Feltwell,T., Goble,A., Hamlin,N., Hauser,H.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Norberczak,H.,
            O'Neil,S., Ormond,D., Price,C., Rabbinowitsch,E., Rutter,S.,
            Sanders,M., Saunders,D., Seeger,K., Sharp,S., Simmonds,M.,
            Skelton,J., Squares,R., Squares,S., Stevens,K., Unwin,L.,
            Whitehead,S., Barrell,B.G. and Maskell,D.J.
  TITLE     Comparative analysis of the genome sequences of Bordetella
            pertussis, Bordetella parapertussis and Bordetella bronchiseptica
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 348706)
  AUTHORS   Sebaihia,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-2003) Submitted on behalf of the Pathogen
            Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA, E-mail: ms5@sanger.ac.uk
FEATURES             Location/Qualifiers
```

```
source          1. .348706
                /organism="Bordetella bronchiseptica RB50"
                /mol_type="genomic DNA"
                /strain="RB50"
                /db_xref="taxon:257310"
gene            160. .2523
                /gene="ppsA"
                /locus_tag="BB2622"
                /note="synonym: pps"
CDS             160. .2523
                /gene="ppsA"
                /locus_tag="BB2622"
                /EC_number="2.7.9.2"
                /note="ortholog of Bordetella pertussis (BX470248) BP1436"
                /codon_start=1
                /transl_table=11
                /product="phosphoenolpyruvate synthase"
                /protein_id="CAE33115.1"
                /db_xref="GI:33576036"
                /db_xref="GOA:Q7WJ77"
                /db_xref="InterPro:IPR000121"
                /db_xref="InterPro:IPR002192"
                /db_xref="InterPro:IPR006319"
                /db_xref="InterPro:IPR008279"
                /db_xref="UniProt/TrEMBL:Q7WJ77"
                /translation="MSYVVSFEQLRMTDVDSVGGKNASLGEMISQLAGAGVRVPGGFA
                TTAQAFRDFLQASGLDKRIADRLATLNPEDVRELASAGAEIRQWVIDAPFSPEFEQAI
                RTAFAELDADGKGSFAVRSSATAEDLPDASFAGQQETFLNVVGVDDVLDKIRHVFASL
                YNDRAISYRVHKGYAHAEVALSAGIQRMVRSDKGSAGVMFTIDTESGFQDVVFITSSY
                GLGETVVQGAVNPDEFYVFKPSLERGNYPIVSRRIGSKLIKMEFDPERPEGRAVRTVD
                VPVSERNRYSLTDDEVIELARYAVIIEKHYQRPMDIEWGRDGVDGKIYILQARPETVK
                SQQGVNDVQQRYRLKATGQVLVTGRAIGQKIGAGKVRVVADISEMDKVQAGDVLVTDM
                TDPNWEPVMKRASAIITNRGGRTCHAAIIARELGIPAVVGCGDATDLLKDGQAVTASC
                AEGDEGRIYDGLIETEVEEVRRGEMPPIDVKIMMNVGNPQLAFDFAQIPNAGVGLARL
                EFIINNNIGIHPKAVLDYPNVDGELKKAVESAARGHASPRAFFVEKLAEGVATIGAAF
                WPKPVIVRMSDFKSNEYRKLVGGSRYEPEEENPMLGFRGASRYIAADFEECFRMECEA
                LKKVRDEMGLTNVEIMVPFVRTLGQARKVVELLAAHGLARGENGLKLIMMCEVPSNAI
                LAEQFLEYFDGFSIGSNDMTQLTLGLDRDSGMELLAADFDERDDAVKFMLRRAIKACL
                AANKYVGICGQGPSDHPDFAQWLKDEGILSMSLNPDTVVETWQQLAK"
misc_feature    163. .1245
                /gene="ppsA"
                /locus_tag="BB2622"
                /note="HMMPfam hit to PF01326, DE Pyruvate phosphate
                dikinase, PEP/pyruvate binding domain, score 1.9e-129"
misc_feature    1246. .1527
                /gene="ppsA"
                /locus_tag="BB2622"
                /note="HMMPfam hit to PF00391, DE PEP-utilizing enzyme,
                mobile domain, score 3.4e-40"
misc_feature    1393. .1428
                /gene="ppsA"
                /locus_tag="BB2622"
                /note="PS00370 PEP-utilizing enzymes phosphorylation site
                signature."
misc_feature    1549. .2517
                /gene="ppsA"
                /locus_tag="BB2622"
                /note="This hit extended beyond the end of the feature by
                1 aa and was clipped.
                HMMPfam hit to PF02896, DE PEP-utilizing enzyme, TIM
                barrel domain, score 7e-104"
misc_feature    2233. .2289
                /gene="ppsA"
                /locus_tag="BB2622"
                /note="PS00742 PEP-utilizing enzymes signature 2."
gene            complement(2591. .3310)
                /locus_tag="BB2623"
CDS             complement(2591. .3310)
                /locus_tag="BB2623"
                /note="ortholog of Bordetella pertussis (BX470248) BP1437"
                /codon_start=1
                /transl_table=11
                /product="probable class-I glutamine amidotransferase)"
```

```
                /protein_id="CAE33116.1"
                /db_xref="GI:33576037"
                /db_xref="GOA:Q7U3A5"
                /db_xref="InterPro:IPR000991"
                /db_xref="UniProt/TrEMBL:Q7U3A5"
                /translation="MTAALPVLILHTGNPAAALRAAHGSYAEMLRQAAGLDAAGTHVV
                ASYKDEAPRAPGAYRAALITGSPAMVTDRAPWSEAAADWLRQAAGAGLPMFGICYGHQ
                LLAHALGGVVGDNPAGRELGTLPVELLPAAATDPLLAGLPAAFDAQMMHEQAVLAAPP
                GAAVLARSAQDAHQMLRLAPRIYTAQFHPEFTPGFVAAHLRHHAALYAAHGLDVERLA
                RAAGPTPLAASLVRRFLACCA"
misc_feature    complement(2693. .3259)
                /locus_tag="BB2623"
                /note="HMMPfam hit to PF00117, DE Glutamine
                amidotransferase class-I, score 3.2e-09"
misc_feature    complement(2999. .3034)
                /locus_tag="BB2623"
                /note="PS00442 Glutamine amidotransferases class-I active
                site."
gene            3434. .3859
                /locus_tag="BB2624"
CDS             3434. .3859
                /locus_tag="BB2624"
                /note="ortholog of Bordetella pertussis (BX470248) BP1438"
                /codon_start=1
                /transl_table=11
                /product="putative membrane protein"
                /protein_id="CAE33117.1"
                /db_xref="GI:33576038"
                /db_xref="UniProt/TrEMBL:Q7WJ76"
                /translation="MWIWFGLAAVALIGEVGSGTFYLLLVALGLAAGGIAAWLGADLA
                WQLVACGLVALLGLLALRRTGVLKKREVNAARNADVNLDIGQSVMVDAWSDAGTTQVY
                YRGAQWQAELAHGQPRQAGRHIIAEIRGTRLVLAPAPAA"
misc_feature    3434. .3526
                /locus_tag="BB2624"
                /note="Signal peptide predicted for BB2624 by SignalP 2.0
                HMM (Signal peptide probabilty 0.622) with cleavage site
                probability 0.235 between residues 31 and 32"
misc_feature    join(3452. .3520,3563. .3616)
                /locus_tag="BB2624"
                /note="2 probable transmembrane helices predicted for
                BB2624 by TMHMM2.0 at aa 7-29 and 44-61"
gene            3913. .4839
                /locus_tag="BB2625"
CDS             3913. .4839
                /locus_tag="BB2625"
                /note="ortholog of Bordetella pertussis (BX470248) BP1440"
                /codon_start=1
                /transl_table=11
                /product="putative membrane protein"
                /protein_id="CAE33118.1"
                /db_xref="GI:33576039"
                /db_xref="GOA:Q7WJ75"
                /db_xref="InterPro:IPR001107"
                /db_xref="UniProt/TrEMBL:Q7WJ75"
                /translation="MIDVSTVVLIVIVILALMIVVKAIAIVPQQHAWVVERLGKFDRV
                LSPGAGFVIPFIERVSYKHSLKEIPLDVPSQVCITRDNTQLQVDGVLYFQVTDPMRAS
                YGSSNYISAITQLAQTTLRSVIGKLELDRTFEEREFINSTIVASLDEAALNWGVKVLR
                YEIKDLTPPNEILRAMQAQITAERERKRALIAASEGRRQEQINIATGEREAAIARSEGE
                KQAQINQAQGEAAAVLAIAEATAKAIEQVGEAVRQPGGMEAVNLKVAERYVDAFSNVA
                KEGNTLILPSNLSDVGGLIASAMTIVKSTKSA"
misc_feature    3913. .3987
                /locus_tag="BB2625"
                /note="Signal peptide predicted for BB2625 by SignalP 2.0
                HMM (Signal peptide probabilty 0.709) with cleavage site
                probability 0.611 between residues 25 and 26"
misc_feature    3931. .3999
                /locus_tag="BB2625"
                /note="1 probable transmembrane helix predicted for BB2625
                by TMHMM2.0 at aa 7-29"
misc_feature    3979. .4503
                /locus_tag="BB2625"
                /note="HMMPfam hit to PF01145, DE SPFH domain / Band 7
                family, score 1.1e-70"
gene            4905. .6062
                /locus_tag="BB2626"
CDS             4905. .6062
                /locus_tag="BB2626"
                /note="ortholog of Bordetella pertussis (BX470248) BP1441"
                /codon_start=1
                /transl_table=11
                /product="probable membrane protein"
                /protein_id="CAE33119.1"
                /db_xref="GI:33576040"
                /db_xref="GOA:Q7WJ74"
                /db_xref="InterPro:IPR000160"
                /db_xref="UniProt/TrEMBL:Q7WJ74"
                /translation="MLAPVTLLVISTLSSLVTLGVLGSLLRSGIAGIRELMWANVLVL
                CSLLLFTLQANPAMPPWLAIQAPNVLISSGVVLFCAGVFRFMQQRPPWRLLLLGLALS
                IGANIWFHYVDPSVNARVVAASGLHALLWGGSAWTIYRNMPLRRSRYSYWFAGVTATV
                AALGHALRTAVYGLQIEQTQGLLQSSVWNVAFLAIGVLVMPSMTLGMIMMIHDRMLAE
                RREREANQDFLTGLLSRKAWWREGERLCARALRTGRPLTLLALDIDRFKQVNDQHGHAA
                GDVVLRHFGSLATALLRSGDVIGRLGGEEFVALLPDTDGDTGVSVAERLLASVRATPC
                SHGGKALTYTFSGGVAQYRAGDTLPVLVERADAALYAAKQAGRDRIRNRAG"
misc_feature    4905. .4973
                /locus_tag="BB2626"
```

```
Alignment Scores:
Pred. No.:              20.7          Length:          348706
Score:                  116.00        Matches:         53
Percent Similarity:     38.8%         Conservative:    27
Best Local Similarity:  25.7%         Mismatches:      80
Query Match:            9.7%          Indels:          46
DB:                     1             Gaps:            11

US-10-757-745-2_COPY_121_347 (1-227) x BX640445 (1-348706)

Qy         25 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 44
              ||||||::::::||||||||||||||||:::      :::      :::      :::
Db     215820 CCGGATCTCGTGTTCCTGCAGGAAGTGCTGGGCGAACACCAGCGGCACGCGGCGCGCCAC 215879

Qy         45 Ser---------------SerAsnTyrGluIleIleThrGly----------------- 53
              :::               |||   ||||||   :::
Db     215880 GCGCGCTGGCCGGCGGTCTCGCAATACGAGTTCCTGGCCGACACCCTGTGGAGCGCCTAT 215939

Qy         54 --------------------HisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 66
                                  :::   :::|||:::          :::|||   |||
Db     215940 GCCTATGGCCGCAACGCCGTCTACCCCGATGGCCACCATGGCAACGCCGTATTGTCCAAG 215999

Qy         67 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 86
              :::          |||   :::          :::          |||
Db     216000 TATCCGATCGTCTCGCACCGCAACGAGGACCTGAGCGTGGGCGCGGCGCACGAGCCGCGC 216059

Qy         87 AsnLeuLeu---CysVal---HisValAsnValSerGlyAsnGluLeuCysLeuMetThr 104
                 :::|||   ||||||      |||      :::          |||      :::
Db     216060 GGCATGCTGCACTGCGTGCTGGACGTGCGCGACGCGCCCGCGCCCCTGCATGCGCTGTGC 216119

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              ||||||                |||      |||::::::   :::|||
Db     216120 GTGCACCTGGGCCTGCGCGAGGCGCATCGCCAGGAGCAGCTGCGGCGCCTGTGCCAGTAC 216179

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              |||:::      |||      |||      |||      :::   |||||||||   |||
Db     216180 CTGCGC------CAGCACGTGCCGGCGCAGGCCGCGGCCCTGCTGGCGGGGGACTTCAAC 216233

Qy        145 -----------LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
                         |||      :::   |||||||||                  :::
Db     216234 GACTGGCGCCAGCGGGCCGACGCAACGCTGGCGCGCTGCGGG--------------CTG 216278

Qy        161 ValAspValTrpGluPhe---LeuGlyLysProLysHisCysGlnTyrThrTrpAspThr 179
              ||||||               |||||         |||          |||:::
Db     216279 GCCGACGTCCACCGGCTGGCGCTGGGG---------CACGCGGCGCGTACCTTT------ 216323

Qy        180 GlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys---LeuArgPheAspArgIle 198
                                  |||||||||      ||||||   |||||||||
Db     216324 ----------------------CCGGCCGCCTGCCCGCTGCTGCGGCTGGACCGCATC 216359
```

```
Qy        199 PhePheArgAlaAlaAla 204
              :::   |||     |||
Db     216360 TACGTGCGCAACGTCGCC 216377

RESULT 71
HS133N16S/c
LOCUS       HS133N16S               411 bp    DNA     linear   STS 02-SEP-1999
DEFINITION  H.sapiens STS from genomic clone 133N16, sequence tagged site.
ACCESSION   AL110544
VERSION     AL110544.1  GI:5830260
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   Mungall,A.J., Huckle,E., Taylor,R. and Hunt,S.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-1999) E-mail contact: humquery@sanger.ac.uk
COMMENT     Marker stdJ133N16SP6 (Primer A : TGTTCCCTCTTTGCTGCAG; Primer B :
            AACAGTTCAAGGTTGTGGGC; amplimer size : 127 bp) is from sequence
            generated from the SP6 end of PAC 133N16. 133N16 is part of the
            bacterial clone contigs constructed by the Chromosome 6 Mapping
            Group.(http://www.sanger.ac.uk/HGP/Chr6/) 133N16 is from the
            library  constructed at the Roswell Park Cancer Institute by the
            group of Pieter de Jong.
            For further details see http://bacpac.med.buffalo.edu/.
FEATURES             Location/Qualifiers
     source          1. .411
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="133N16"
ORIGIN

Alignment Scores:
Pred. No.:                0.0114          Length:         411
Score:                    114.50          Matches:        27
Percent Similarity:       69.0%           Conservative:   2
Best Local Similarity:    64.3%           Mismatches:     10
Query Match:              9.6%            Indels:         3
DB:                       10              Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x HS133N16S (1-411)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 256

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||      ::::::|||          |||          |||   |||
Db        255 GCTTTGTAAGTATTATCACTTCTATTTAATAGTAATGTG---------TCTTATGTATAT 205

Qy         41 LeuLys 42
              |||
Db        204 ACAAAG 199

RESULT 72
AE016853_39/c
WPCOMMENT
  Sequence split into 64 fragments  LOCUS AE016853 Accession AE016853
     Fragment Name      Begin     End
     AE016853_00            1    110000
     AE016853_01       100001    210000
     AE016853_02       200001    310000
     AE016853_03       300001    410000
     AE016853_04       400001    510000
     AE016853_05       500001    610000
     AE016853_06       600001    710000
     AE016853_07       700001    810000
     AE016853_08       800001    910000
     AE016853_09       900001   1010000
     AE016853_10      1000001   1110000
     AE016853_11      1100001   1210000
     AE016853_12      1200001   1310000
     AE016853_13      1300001   1410000
     AE016853_14      1400001   1510000
     AE016853_15      1500001   1610000
     AE016853_16      1600001   1710000
     AE016853_17      1700001   1810000
     AE016853_18      1800001   1910000
     AE016853_19      1900001   2010000
     AE016853_20      2000001   2110000
     AE016853_21      2100001   2210000
     AE016853_22      2200001   2310000
     AE016853_23      2300001   2410000
     AE016853_24      2400001   2510000
     AE016853_25      2500001   2610000
     AE016853_26      2600001   2710000
     AE016853_27      2700001   2810000
     AE016853_28      2800001   2910000
     AE016853_29      2900001   3010000
     AE016853_30      3000001   3110000
     AE016853_31      3100001   3210000
     AE016853_32      3200001   3310000
     AE016853_33      3300001   3410000
     AE016853_34      3400001   3510000
     AE016853_35      3500001   3610000
     AE016853_36      3600001   3710000
     AE016853_37      3700001   3810000
     AE016853_38      3800001   3910000
     AE016853_39      3900001   4010000
     AE016853_40      4000001   4110000
     AE016853_41      4100001   4210000
     AE016853_42      4200001   4310000
     AE016853_43      4300001   4410000
     AE016853_44      4400001   4510000
     AE016853_45      4500001   4610000
     AE016853_46      4600001   4710000
     AE016853_47      4700001   4810000
     AE016853_48      4800001   4910000
     AE016853_49      4900001   5010000
     AE016853_50      5000001   5110000
     AE016853_51      5100001   5210000
     AE016853_52      5200001   5310000
     AE016853_53      5300001   5410000
     AE016853_54      5400001   5510000
     AE016853_55      5500001   5610000
     AE016853_56      5600001   5710000
     AE016853_57      5700001   5810000
     AE016853_58      5800001   5910000
     AE016853_59      5900001   6010000
     AE016853_60      6000001   6110000
     AE016853_61      6100001   6210000
     AE016853_62      6200001   6310000
     AE016853_63      6300001   6397126
Continuation (40 of 64) of AE016853 from base 3900001 (AE016853 Pseudomonas syringae pv

Alignment Scores:
Pred. No.:                11.1            Length:         110000
Score:                    113.00          Matches:        54
Percent Similarity:       35.5%           Conservative:   23
Best Local Similarity:    24.9%           Mismatches:     98
Query Match:              9.5%            Indels:         42
DB:                       1               Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x AE016853_39 (1-110000)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              |||       |||   |||       |||   |||                     |||
Db      57655 CTTCCGGAGTTGCGTGAGGCCGTACGCAGTGTA-------------------TCGGCC 57617

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys------ 43
              |||::::::|||||||||||||||||         ::::::|||:::   ::::::
```

```
Db      57616 GATATCGTATTTCTTCAGGAAGTGCACGGTGAACACCATAGCCATGCCCGCAGGGTCAAG 57557

Qy         44 ---------ArgSerSerAsnTyrGluIleIleThrGly---------------------- 53

Db      57556 GACTGGCCGACGATTTCGCAGTACGAGTTTCTCGCCGACAGCATGTGGAGCGACTTTGCC 57497

Qy         54 ------------------HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 67

Db      57496 TACGGCCGCAACGCGGTGTACCCGGACGGCGATCACGGTAATGCGCTGCTCTCCAAGTAC 57437

Qy         68 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 87

Db      57436 CCCATCATCCAGCATGAAAACCTCGACGTTTCCATTGATGGCACCGAGCAGCGCGGTTTG 57377

Qy         88 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 107

Db      57376 CTGCACTGCATTCTGCAAGTGCCTGATCATGGCCAGGTGCATGCGGTCTGCGTGCACCTG 57317

Qy        108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127

Db      57316 GGGTTGTTTGAAAGCCAT---------CGGCGCCAACAGCTGAAACTGCTTGCCCAGCTC 57266

Qy        128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn---LeuArg 146

Db      57265 ATGGACCGCCTGCCTGAAGGTGCGCCGGTCATTGTGGCAGGTGATTTCAATGACTGGCGC 57206

Qy        147 AspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPhe 166

Db      57205 CAACGTGCCGATGCGTTGTTGCAGGGCACCGGGCTGCACGAGGTGTTTGTACAGCGATTC 57146

Qy        167 LeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGly 186

Db      57145 GGCGCTCCGGCGAAAAGTTTTCCGGCCCGCTGGCCGCTG---------------------- 57107

Qy        187 IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAla 203

Db      57106 ------------------CTCCGGCTGGACCGGATTTATGTGCGCAACGCG 57074
```

RESULT 73
CR855361/c

```
LOCUS       CR855361              225886 bp    DNA     linear   HTG 12-JUN-2005
DEFINITION  Danio rerio chromosome 14 clone CH211-199E8, WORKING DRAFT
            SEQUENCE, 9 unordered pieces.
ACCESSION   CR855361
VERSION     CR855361.3  GI:67514755
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 225886)
  AUTHORS   McLaren,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests:
            http://www.sanger.ac.uk/Projects/D_rerio/faqs.shtml#dataeight
COMMENT     On Jun 12, 2005 this sequence version replaced gi:55294804.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zC199E8
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 221894 bases at least Q40
            Consensus quality: 222929 bases at least Q30
            Consensus quality: 223766 bases at least Q20
            Insert size: 225086; sum-of-contigs
            Insert size: 211979; 3.3% error; agarose-fp
            Quality coverage: 9.91x in Q20 bases; sum-of-contigs Quality
            coverage: 10.65x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 9 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1    48263: contig of 48263 bp in length
            *   48264    48363: gap of 100 bp
            *   48364    63904: contig of 15541 bp in length
            *   63905    64004: gap of 100 bp
            *   64005    78065: contig of 14061 bp in length
            *   78066    78165: gap of 100 bp
            *   78166    81998: contig of 3833 bp in length
            *   81999    82098: gap of 100 bp
            *   82099    93026: contig of 10928 bp in length
            *   93027    93126: gap of 100 bp
            *   93127   111025: contig of 17899 bp in length
            *  111026   111125: gap of 100 bp
            *  111126   117571: contig of 6446 bp in length
            *  117572   117671: gap of 100 bp
            *  117672   198161: contig of 80490 bp in length
            *  198162   198261: gap of 100 bp
            *  198262   225886: contig of 27625 bp in length.
FEATURES             Location/Qualifiers
     source          1. .225886
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /chromosome="14"
                     /clone="CH211-199E8"
                     /clone_lib="CHORI-211"
     misc_feature    1. .48263
                     /note="assembly_fragment:01320
                     clone_end:SP6
                     vector_side:left"
     misc_feature    48364. .63904
                     /note="assembly_fragment:00249
                     fragment_chain:1"
     misc_feature    64005. .78065
                     /note="assembly_fragment:00452
                     fragment_chain:1"
     misc_feature    78166. .81998
                     /note="assembly_fragment:00015
                     fragment_chain:1"
     misc_feature    82099. .93026
                     /note="assembly_fragment:00039
                     fragment_chain:2"
     misc_feature    93127. .111025
                     /note="assembly_fragment:00666
                     fragment_chain:2"
     misc_feature    111126. .117571
                     /note="assembly_fragment:00132
                     fragment_chain:2"
     misc_feature    117672. .198161
                     /note="assembly_fragment:02086
                     fragment_chain:2"
     misc_feature    198262. .225886
                     /note="assembly_fragment:00907
                     fragment_chain:2
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:              25.6          Length:         225886
Score:                  113.00        Matches:        58
Percent Similarity:     34.4%         Conservative:   25
```

```
Best Local Similarity:  24.1%          Mismatches:     67
Query Match:            9.5%           Indels:         91
DB:                     14             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x CR855361 (1-225886)

Qy         19 TyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu------------------ 32
              ||||||  |||  |||||||||||||||||||||
Db     115222 TATTTGCGGCTCAAAAAGGCTGATGTTATTTTCTTACAAGAAACGCATAGAGACCAACAA 115163

Qy         32 ------------------------------------------------------------ 32

Db     115162 AATCAGGCGCAGTGGCATAGTGGGTGGAAAGGAAATGTTCTACTAAGCCATAGCACAAAC 115103

Qy         33 ------------ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 48
                          |||     :::      |||:::     :::|||:::::::
Db     115102 CTAAGTGTTGGTGTTGCCTTTCTTTTTTCCCCATATGTCAGTATGCAATCAAATTCAACA 115043

Qy         49 GluIleIleThrGlyHisGlu------------------------GluGlyTyrPheThr 60
              |||||||||   |||   :::                      :::   ||||||:::
Db     115042 GAAATAATACCCGGTCGGAAATTGAGGGCTGATATTGTCATAGGAGATGCCTATTTCTCT 114983

Qy         61 AlaIleMetLeu------------------------LysLysSerArgValLysLeuLys 72
                 |||   :::                      :::   |||     |||:::
Db     114982 TTTATTAATATATATGCACCAACTGTTGGTCATGACCGTGTTTCTTTCTTTAAAATACTT 114923

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              |||     |||:::     |||      |||
Db     114922 TCTGATGCCATATTACAGTGTCCTCGGGGGAATATG------------------------ 114887

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              :::   :::   |||:::     |||   :::   |||||||||:::     |||
Db     114886 ATTGTACTTGGAGGGGATTTTAATTGCACTATTGATTCACATTTGGAC------AGAAAT 114833

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              |||   |||     |||          |||||||||:::   |||
Db     114832 CAC---AACGAACCTCATAATTCTTCTGCTGAGAGTTTAAAAAAAATTAATGAA------ 114782

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152

Db     114782 ------------------------------------------------------------ 114782

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
                        :::   :::||||||||||||          |||
Db     114781 --------------CATAAGTTAGTTGATGTGTGGAGG---GAAGCTTTCCTGGAGTC 114740

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||                ::::::||||||
Db     114739 AGACAGTACACATGGTTAAAAAATAATTCAAATAATTTGTCA--------------GGG 114695

Qy        193 LeuArgPheAspArgIlePhe-PheArgAlaAlaAlaGluGluGlyHisIleIleProAr 212
                 |||   ||||||   :::   :::   |||   |||   |||   :::|||   ||
Db     114694 GCAAGGCTTGATCGGTTTTATGTTGAAAAGGGCACTAGAGGTAGGTTTTTTAATAGCTCG 114635

Qy        212 g 212
              |
Db     114634 A 114634

RESULT 74
CR382132_15
WPCOMMENT
  Sequence split into 40 fragments  LOCUS CR382132 Accession CR382132
     Fragment Name      Begin     End
     CR382132_00            1    110000
     CR382132_01       100001    210000
     CR382132_02       200001    310000
     CR382132_03       300001    410000
     CR382132_04       400001    510000
     CR382132_05       500001    610000
     CR382132_06       600001    710000
     CR382132_07       700001    810000
     CR382132_08       800001    910000
     CR382132_09       900001   1010000
     CR382132_10      1000001   1110000
     CR382132_11      1100001   1210000
     CR382132_12      1200001   1310000
     CR382132_13      1300001   1410000
     CR382132_14      1400001   1510000
     CR382132_15      1500001   1610000
     CR382132_16      1600001   1710000
     CR382132_17      1700001   1810000
     CR382132_18      1800001   1910000
     CR382132_19      1900001   2010000
     CR382132_20      2000001   2110000
     CR382132_21      2100001   2210000
     CR382132_22      2200001   2310000
     CR382132_23      2300001   2410000
     CR382132_24      2400001   2510000
     CR382132_25      2500001   2610000
     CR382132_26      2600001   2710000
     CR382132_27      2700001   2810000
     CR382132_28      2800001   2910000
     CR382132_29      2900001   3010000
     CR382132_30      3000001   3110000
     CR382132_31      3100001   3210000
     CR382132_32      3200001   3310000
     CR382132_33      3300001   3410000
     CR382132_34      3400001   3510000
     CR382132_35      3500001   3610000
     CR382132_36      3600001   3710000
     CR382132_37      3700001   3810000
     CR382132_38      3800001   3910000
     CR382132_39      3900001   4003362
Continuation (16 of 40) of CR382132 from base 1500001 (CR382132 Yarrowia lipolytica chrc

Alignment Scores:
Pred. No.:              15.9           Length:         110000
Score:                  111.50         Matches:        39
Percent Similarity:     44.8%          Conservative:   30
Best Local Similarity:  25.3%          Mismatches:     50
Query Match:            9.4%           Indels:         35
DB:                     15             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x CR382132_15 (1-110000)

Qy         85 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr 104
              ::::::      :::   :::   :::|||||||    ||||||   :::      :::
Db     104632 CTCAAGGGCGCCGTGGGTCTCCGACTCAATGTTGATGGCAATCCCGTTGTGTGTGGCT 104691

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              :::||||||   :::        |||          :::|||   :::          :::
Db     104692 GCGCACCTGGCGGCCAATGAGGGCTTTTTGGAGCAGCGCAACGACGATTTCGCCACCATT 104751

Qy        125 LeuLysLysMetGlnGluAlaPro------------------GluSerAlaThrValIle 138
                 |||:::|||:::                          :::   |||||||||
Db     104752 GCCAAGGAAATGAAGTTCCTCGACCCACGCACAGGCATCTACGATGACGCCACGGTATTC 104811

Qy        139 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 158
              |||   ||||||   |||   |||   |||      :::      ||||||   :::
Db     104812 TTTTTGGGTGATCTCAACTACCGAGCGTCGGAAAACGAGCAG------GGACTCGCTGAC 104865

Qy        159 ---AsnIleValAsp--------------------------------ValTrpGluPhe 166
                 :::||||||                                   |||:::|||
Db     104866 TACAAGGTGGTAGACGAGCTGTTTTACGCCATGAGGTCGCTCGATGCCTTCTGGGACTTC 104925

Qy        167 ------------LeuGly---LysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
                         :::|||   ||||||  :::      :::|||          :::::
Db     104926 CAGGAGGCCGAAATCGGATTCAAACCAACCTACAAGTTCCACACCGGAACGTCCGAATAC 104985

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
              |||:::     |||   :::   |||            ||||||:::   :::|||
Db     104986 AACACCAAG---CGAATTCCGTCCTGGTGC------------GACCGAGTGCTGTATAGA 105030

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
                       |||   ::::::||||||       |||
Db     105031 --------------GGAGCAGTGGTGCCCCGGAAATACGAC 105057
```

```
RESULT 75
CR382132_16
WPCOMMENT
  Sequence split into 40 fragments  LOCUS CR382132 Accession CR382132
     Fragment Name      Begin     End
     CR382132_00            1    110000
     CR382132_01       100001    210000
     CR382132_02       200001    310000
     CR382132_03       300001    410000
     CR382132_04       400001    510000
     CR382132_05       500001    610000
     CR382132_06       600001    710000
     CR382132_07       700001    810000
     CR382132_08       800001    910000
     CR382132_09       900001   1010000
     CR382132_10      1000001   1110000
     CR382132_11      1100001   1210000
     CR382132_12      1200001   1310000
     CR382132_13      1300001   1410000
     CR382132_14      1400001   1510000
     CR382132_15      1500001   1610000
     CR382132_16      1600001   1710000
     CR382132_17      1700001   1810000
     CR382132_18      1800001   1910000
     CR382132_19      1900001   2010000
     CR382132_20      2000001   2110000
     CR382132_21      2100001   2210000
     CR382132_22      2200001   2310000
     CR382132_23      2300001   2410000
     CR382132_24      2400001   2510000
     CR382132_25      2500001   2610000
     CR382132_26      2600001   2710000
     CR382132_27      2700001   2810000
     CR382132_28      2800001   2910000
     CR382132_29      2900001   3010000
     CR382132_30      3000001   3110000
     CR382132_31      3100001   3210000
     CR382132_32      3200001   3310000
     CR382132_33      3300001   3410000
     CR382132_34      3400001   3510000
     CR382132_35      3500001   3610000
     CR382132_36      3600001   3710000
     CR382132_37      3700001   3810000
     CR382132_38      3800001   3910000
     CR382132_39      3900001   4003362
Continuation (17 of 40) of CR382132 from base 1600001 (CR382132 Yarrowia lipolytica chrc

Alignment Scores:
Pred. No.:                15.9          Length:         110000
Score:                    111.50        Matches:        39
Percent Similarity:       44.8%         Conservative:   30
Best Local Similarity:    25.3%         Mismatches:     50
Query Match:              9.4%          Indels:         35
DB:                       15            Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x CR382132_16 (1-110000)

Qy         85 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr 104
              ::::::      :::   :::   :::||||||   ||||||   :::      :::
Db       4632 CTCAAGGGCGCCGTGGGTCTCCGACTCAATGTTGATGGCAATCCCGTTGTGTGTGTGGCT 4691

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              :::||||||   :::       |||          :::|||   :::          :::
Db       4692 GCGCACCTGGCGGCCAATGAGGGCTTTTTGGAGCAGCGCAACGACGATTTCGCCACCATT 4751

Qy        125 LeuLysLysMetGlnGluAlaPro-----------------GluSerAlaThrValIle 138
                 |||:::|||:::          :::   |||||||||
Db       4752 GCCAAGGAAATGAAGTTCCTCGACCCACGCACAGGCATCTACGATGACGCCACGGTATTC 4811

Qy        139 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsn 158
              |||   ||||||   |||   |||       |||      :::      ||||||  :::
Db       4812 TTTTTGGGTGATCTCAACTACCGAGCGTCGGAAAACGAGCAG------GGACTCGCTGAC 4865

Qy        159 ---AsnIleValAsp--------------------------------ValTrpGluPhe 166
                    :::||||||                                  |||:::|||
Db       4866 TACAAGGTGGTAGACGAGCTGTTTTACGCCATGAGGTCGCTCGATGCCTTCTGGGACTTC 4925

Qy        167 ------------LeuGly---LysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
                          :::|||   ||||||   :::      :::|||      ::::::
Db       4926 CAGGAGGCCGAAATCGGATTCAAACCAACCTACAAGTTCCACACCGGAACGTCCGAATAC 4985

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
              |||:::         |||   :::   |||            ||||||:::   :::|||
Db       4986 AACACCAAG---CGAATTCCGTCCTGGTGC-----------GACCGAGTGCTGTATAGA 5030

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
                          |||   ::::::||||||      |||
Db       5031 --------------GGAGCAGTGGTGCCCCGGAAATACGAC 5057


Search completed: March 10, 2006, 22:41:51
Job time : 8816.38 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 15:40:34 ; Search time 187.478 Seconds
(without alignments)
1746.920 Million cell updates/sec

Title: US-10-757-745-2_COPY_115_121
Perfect score: 38
Sequence: 1 SLITWNI 7

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170329_4058/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_12397_@runat_09032006_170329_4058 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	38	100.0	216	11	CR196611	CR196611 Forward s
	2	38	100.0	221	11	CR404572	CR404572 Arabidops
c	3	38	100.0	227	11	CR094068	CR094068 Forward s
c	4	38	100.0	246	10	CNS00VF5	AL092207 Arabidops
c	5	38	100.0	289	1	BB125347	BB125347 BB125347
	6	38	100.0	310	8	T26956	T26956 11d155proj
c	7	38	100.0	347	5	BU070378	BU070378 im09c09.x
	8	38	100.0	352	5	BQ787387	BQ787387 im09c09.y
c	9	38	100.0	368	2	BE694706	BE694706 PM0-BT075
	10	38	100.0	394	1	AV661333	AV661333 AV661333
	11	38	100.0	429	5	BQ582059	BQ582059 il12c11.y
	12	38	100.0	457	11	CR404571	CR404571 Arabidops
	13	38	100.0	463	8	DN851694	DN851694 4147331 B
	14	38	100.0	471	7	CR543841	CR543841 DKFZp459O
	15	38	100.0	473	6	CB120234	CB120234 K-EST0167
	16	38	100.0	480	1	AI750554	AI750554 cn04a01.y
	17	38	100.0	485	1	AA486032	AA486032 ab40b10.r
c	18	38	100.0	511	5	BU659432	BU659432 cl47f06.z
	19	38	100.0	518	9	CC153114	CC153114 CSU-K34.1
	20	38	100.0	530	3	BM846221	BM846221 K-EST0125
	21	38	100.0	531	7	CN233559	CN233559 RJA113F02
	22	38	100.0	554	3	BP220985	BP220985 BP220985
c	23	38	100.0	558	10	BX997515	BX997515 Forward s
	24	38	100.0	565	3	BP226620	BP226620 BP226620
	25	38	100.0	566	3	BP220508	BP220508 BP220508
	26	38	100.0	568	3	BP257211	BP257211 BP257211
	27	38	100.0	570	3	BM172060	BM172060 imageqc_3
	28	38	100.0	570	3	BP221518	BP221518 BP221518
	29	38	100.0	570	5	BU397611	BU397611 603534313
	30	38	100.0	572	3	BP262485	BP262485 BP262485
	31	38	100.0	573	1	AU279894	AU279894 AU279894
	32	38	100.0	573	3	BP221260	BP221260 BP221260
	33	38	100.0	573	3	BP261500	BP261500 BP261500
	34	38	100.0	576	3	BP262642	BP262642 BP262642
	35	38	100.0	577	7	CN298919	CN298919 170006001
	36	38	100.0	580	3	BP261787	BP261787 BP261787
	37	38	100.0	580	3	BP274368	BP274368 BP274368
	38	38	100.0	581	1	AV717253	AV717253 AV717253
	39	38	100.0	581	3	BP199948	BP199948 BP199948
	40	38	100.0	581	3	BP207751	BP207751 BP207751
	41	38	100.0	581	3	BP225088	BP225088 BP225088
	42	38	100.0	581	3	BP262638	BP262638 BP262638
	43	38	100.0	581	3	BP264149	BP264149 BP264149
	44	38	100.0	581	3	BP270744	BP270744 BP270744
	45	38	100.0	581	3	BP349936	BP349936 BP349936
	46	38	100.0	581	3	BP379659	BP379659 BP379659
	47	38	100.0	582	3	BM507091	BM507091 ih24h01.y
	48	38	100.0	582	3	BP219740	BP219740 BP219740
	49	38	100.0	582	3	BP233592	BP233592 BP233592
	50	38	100.0	582	3	BP234218	BP234218 BP234218
	51	38	100.0	582	3	BP261141	BP261141 BP261141
	52	38	100.0	582	3	BP263516	BP263516 BP263516
	53	38	100.0	582	3	BP270415	BP270415 BP270415
	54	38	100.0	582	3	BP275810	BP275810 BP275810
	55	38	100.0	583	3	BP195721	BP195721 BP195721
	56	38	100.0	583	3	BP262103	BP262103 BP262103
	57	38	100.0	583	3	BP262741	BP262741 BP262741
	58	38	100.0	583	3	BP319548	BP319548 BP319548
	59	38	100.0	583	5	BU783229	BU783229 in01a08.y
	60	38	100.0	584	3	BP263059	BP263059 BP263059
	61	38	100.0	584	3	BP274942	BP274942 BP274942
	62	38	100.0	584	3	BP348623	BP348623 BP348623
	63	38	100.0	588	3	BP236074	BP236074 BP236074
	64	38	100.0	588	3	BP348796	BP348796 BP348796
	65	38	100.0	600	2	BI064472	BI064472 pgf1n.pk0
	66	38	100.0	602	1	AL703449	AL703449 DKFZp686G
	67	38	100.0	604	1	AV730762	AV730762 AV730762
c	68	38	100.0	618	11	CR163431	CR163431 Forward s
	69	38	100.0	621	6	CB069952	CB069952 is31e10.y
	70	38	100.0	625	5	BU471395	BU471395 603363114
	71	38	100.0	634	6	CB157906	CB157906 K-EST0217
	72	38	100.0	634	7	CV023369	CV023369 288 Full
	73	38	100.0	644	3	BI333830	BI333830 602999339
	74	38	100.0	644	9	AZ360347	AZ360347 1M0103M01
	75	38	100.0	645	3	BM426270	BM426270 pgf2n.pk0

ALIGNMENTS

RESULT 1

```
CR196611/c
LOCUS       CR196611                216 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Forward strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN168n23, genomic survey sequence.
ACCESSION   CR196611
VERSION     CR196611.1  GI:49975460
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 216)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .216
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN168n23"
                     /clone_lib="MHPN"
ORIGIN

Alignment Scores:
Pred. No.:                 1.29e+03        Length:          216
Score:                     38.00           Matches:         7
Percent Similarity:        100.0%          Conservative:    0
Best Local Similarity:     100.0%          Mismatches:      0
Query Match:               100.0%          Indels:          0
DB:                        11              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x CR196611 (1-216)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        191 TCATTAATAACCTGGAATATC 171

RESULT 2
CR404572
LOCUS       CR404572                221 bp    DNA     linear   GSS 02-MAY-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-872B08-026119,
            genomic survey sequence.
ACCESSION   CR404572
VERSION     CR404572.1  GI:46945300
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
  TITLE     GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
            the identification of T-DNA insertion mutants in Arabidopsis
            thaliana
  JOURNAL   Bioinformatics 19 (11), 1441-1442 (2003)
  PUBMED    12874060
REFERENCE   2
  AUTHORS   Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
            Weisshaar,B.
  TITLE     An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
            flanking sequence tag-based reverse genetics
  JOURNAL   Plant Mol. Biol. 53 (1-2), 247-259 (2003)
  PUBMED    14756321
REFERENCE   3
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and
            Weisshaar,B.
  TITLE     High-throughput generation of sequence indexes from T-DNA
            mutagenized Arabidopsis thaliana lines
  JOURNAL   BioTechniques 35 (6), 1164-1168 (2003)
  PUBMED    14682050
REFERENCE   4  (bases 1 to 221)
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2004) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence has been recovered from the left border of the T-DNA.
            It indicates an insertion within the locus defined by BAC clone
            MYF24. Details on the protocols used for generation of the sequence
            are described in References 1-3. The sequences are generated at the
            MPI for Plant Breeding Research in the context of the GABI-Kat
            project. GABI-Kat is part of the German Plant Genomics program
            designated 'GABI'. Information on line availability can be found
            at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .221
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="GK-872B08-026119"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161 (GenBank accession number: AJ537514). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN

Alignment Scores:
Pred. No.:                 1.32e+03        Length:          221
Score:                     38.00           Matches:         7
Percent Similarity:        100.0%          Conservative:    0
Best Local Similarity:     100.0%          Mismatches:      0
Query Match:               100.0%          Indels:          0
DB:                        11              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x CR404572 (1-221)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        161 AGTTTGATTACTTGGAACATA 181

RESULT 3
CR094068/c
LOCUS       CR094068                227 bp    DNA     linear   GSS 05-JUL-2004
DEFINITION  Forward strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN414109, genomic survey sequence.
ACCESSION   CR094068
VERSION     CR094068.1  GI:49827897
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 227)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .227
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN414109"
```

/clone_lib="MHPN"
ORIGIN

Alignment Scores:

Pred. No.:	1.36e+03	Length:	227
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x CR094068 (1-227)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        202 TCATTAATAACCTGGAATATC 182
```

RESULT 4
CNS00VF5/c
LOCUS CNS00VF5 246 bp DNA linear GSS 28-JUN-1999
DEFINITION Arabidopsis thaliana genome survey sequence SP6 end of BAC T9H2 of TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.
ACCESSION AL092207
VERSION AL092207.1 GI:5293361
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 246)
AUTHORS Salanoubat,M., Choisne,N., Artiguenave,F., Brottier,P., Wincker,P., Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 246)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
FEATURES Location/Qualifiers
source 1. .246
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="T9H2"
/clone_lib="TAMU"
/ecotype="Columbia"
/note="end : SP6"
ORIGIN

Alignment Scores:

Pred. No.:	1.5e+03	Length:	246
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x CNS00VF5 (1-246)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        111 AGTTTGATTACTTGGAACATA 91
```

RESULT 5
BB125347/c
LOCUS BB125347 289 bp mRNA linear EST 28-JUN-2000
DEFINITION BB125347 RIKEN full-length enriched, 16 days neonate cerebellum Mus musculus cDNA clone 9630004K16 3' similar to AF046001 Homo sapiens zinc finger transcription factor (ZNF207) mRNA, mRNA sequence.
ACCESSION BB125347
VERSION BB125347.1 GI:8779673
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 289)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .289
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9630004K16"
/tissue_type="cerebellum"
/dev_stage="16 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 16 days neonate cerebellum"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda

FLC I."
ORIGIN

Alignment Scores:
Pred. No.: 1.82e+03 Length: 289
Score: 38.00 Matches: 7
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x BB125347 (1-289)

Qy 1 SerLeuIleThrTrpAsnIle 7
||||||||||||||||||||
Db 83 TCTCTAATTACATGGAATATC 63

RESULT 6
T26956
LOCUS T26956 310 bp mRNA linear EST 12-JUN-1996
DEFINITION 11d155proj Clontech adult human fat cell library HL1108A Homo sapiens cDNA clone 11d155, mRNA sequence.
ACCESSION T26956 T27254
VERSION T26956.1 GI:575899
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 310)
AUTHORS Bouillaud,F.
TITLE Study of expressed sequences tags in adipose tissue 1994
JOURNAL Unpublished (1994)
COMMENT On Nov 17, 1994 this sequence version replaced gi:624970.
Contact: Frederic Bouillaud
Centre de Recherche sur l'Endocrinologie moleculaire et le Developpement
CNRS
9, Rue Jules Hetzel, Meudon Bellevue, 92190 France
Tel: 33 1 45 07 52 87
Fax: 33 1 45 07 58 90
Email: bouillau@infobiogen.fr
Southern blot human DNA EcoRI : 2 major bands 3.1, 2.3 Kb and 2 weak : 4.2 and 1.9 Kb (wash 0.5xSSC).
FEATURES Location/Qualifiers
source 1. .310
/organism="Homo sapiens"
/mol_type="mRNA"
/strain="Caucasian"
/db_xref="taxon:9606"
/clone="11d155"
/clone_lib="Clontech adult human fat cell library HL1108A"
/note="Vector: lambda gt10; Site_1: EcoRI; dev-stage=adult; tissue-type=adipose tissue; lab_host=Bacteriophage lambda; first strand priming with random and poly-d(T) oligonucleotides."
ORIGIN

Alignment Scores:
Pred. No.: 1.98e+03 Length: 310
Score: 38.00 Matches: 7
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x T26956 (1-310)

Qy 1 SerLeuIleThrTrpAsnIle 7
||||||||||||||||||||
Db 239 TCTCTCATYACCTGGAATATT 259

RESULT 7
BU070378/c
LOCUS BU070378 347 bp mRNA linear EST 27-AUG-2002
DEFINITION im09c09.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:6034553 3' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION BU070378
VERSION BU070378.1 GI:22511567
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 347)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Other_ESTs: im09c09.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. J. Ferrer In vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40UP from Gibco.
FEATURES Location/Qualifiers
source 1. .347
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6034553"
/tissue_type="insulinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Human insulinoma"
/note="Organ: pancreas; Vector: pBluescript SK-; Site_1: XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.: 2.26e+03 Length: 347
Score: 38.00 Matches: 7
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x BU070378 (1-347)

Qy 1 SerLeuIleThrTrpAsnIle 7
||||||||||||||||||||
Db 51 TCTCTCATTACCTGGAATATT 31

RESULT 8

```
BQ787387
LOCUS       BQ787387                352 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  im09c09.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6034553
            5' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION   BQ787387
VERSION     BQ787387.1  GI:21995859
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 352)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .352
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6034553"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:                2.3e+03        Length:          352
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       5              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BQ787387 (1-352)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        298 TCTCTCATTACCTGGAATATT 318

RESULT 9
BE694706/c
LOCUS       BE694706                368 bp    mRNA    linear   EST 11-SEP-2000
DEFINITION  PM0-BT0757-140800-006-d10 BT0757 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE694706
VERSION     BE694706.1  GI:10081866
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 368)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  PUBMED    10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=PM0-BT0757-140
            800-006-d10&t3=2000-08-14&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 37
            High quality sequence stop: 368.
FEATURES             Location/Qualifiers
     source          1. .368
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="BT0757"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:                2.43e+03       Length:          368
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       2              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BE694706 (1-368)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        266 TCTCTCATTACCTGGAATATT 246

RESULT 10
AV661333
LOCUS       AV661333                394 bp    mRNA    linear   EST 16-JAN-2002
DEFINITION  AV661333 GLC Homo sapiens cDNA clone GLCGRD01 3', mRNA sequence.
ACCESSION   AV661333
VERSION     AV661333.1  GI:9882347
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 394)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
   PUBMED   11752456
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .394
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="GLCGRD01"
                     /tissue_type="corresponding non cancerous liver tissue"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /clone_lib="GLC"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

Alignment Scores:
Pred. No.:                2.63e+03        Length:         394
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       1               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AV661333 (1-394)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        131 TCTCTCATTACCTGGAATATT 151

RESULT 11
BQ582059
LOCUS       BQ582059                 429 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  il12c11.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6029589
            5' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION   BQ582059
VERSION     BQ582059.1  GI:21494955
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: il12c11.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6029589"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact Hiroshi Inoue, MD/PhD for further
                     information on this library (Metabolism Division, Permutt
                     Laboratory, Washington University School of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:                2.91e+03        Length:         429
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BQ582059 (1-429)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        286 TCTCTCATTACCTGGAATATT 306

RESULT 12
CR404571
LOCUS       CR404571                 457 bp    DNA     linear   GSS 02-MAY-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-872B08-026084,
            genomic survey sequence.
ACCESSION   CR404571
VERSION     CR404571.1  GI:46945299
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
  TITLE     GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
            the identification of T-DNA insertion mutants in Arabidopsis
            thaliana
  JOURNAL   Bioinformatics 19 (11), 1441-1442 (2003)
   PUBMED   12874060
REFERENCE   2
  AUTHORS   Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
            Weisshaar,B.
  TITLE     An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
```

flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
PUBMED 14756321
REFERENCE 3
AUTHORS Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and Weisshaar,B.
TITLE High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
JOURNAL BioTechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
REFERENCE 4 (bases 1 to 457)
AUTHORS Li,Y., Strizhov,N., Rosso,M.G. and Weisshaar,B.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2004) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence has been recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by BAC clone MYF24. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES Location/Qualifiers
source 1. .457
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GK-872B08-026084"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161 (GenBank accession number: AJ537514). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."
ORIGIN

Alignment Scores:

Pred. No.:	3.14e+03	Length:	457
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x CR404571 (1-457)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        103 AGTTTGATTACTTGGAACATA 123
```

RESULT 13
DN851694
LOCUS DN851694 463 bp mRNA linear EST 20-APR-2005
DEFINITION 4147331 BARC_3GAL chicken mixed tissue Gallus gallus cDNA clone 3GAL_45I09 5', mRNA sequence.
ACCESSION DN851694
VERSION DN851694.1 GI:62810854
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 463)
AUTHORS Evock-Clover,C.M., Ashwell,C.M., McMurtry,J.P., Lillehoj,H.S., Matukumalli,L.K. and Van Tassell,C.P.
TITLE Characterization of expressed sequence tags generated from multiple chicken tissues
JOURNAL Unpublished (2004)
COMMENT Contact: Christina M. Clover
Growth Biology Laboratory
Animal and Natural Resources Institute
Bldg. 200 Rm215 BARC-East, Beltsville, MD 20705, USA
Tel: 3015048224
Fax: 3015048623
Email: chrisc@anri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim_alt '' -trim_fasta. Vector identified by cross_match using options -minmatch 12 -minscore 18
Plate: 45 row: I column: 09
Seq primer: CCCAGTCACGACGTTGTAAAACG
High quality sequence stop: 463.
FEATURES Location/Qualifiers
source 1. .463
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Leghorn and broiler"
/db_xref="taxon:9031"
/clone="3GAL_45I09"
/lab_host="DH5alpha"
/clone_lib="BARC_3GAL chicken mixed tissue"
/note="Vector: pBluescript, SK+, Stratagene; Site_1: NotI; Site_2: EcoRI; Normalized library of pooled RNA isolated from whole brain, ultimobranchial gland, parathyroid gland, cecal tonsil and primordial germ cells Multiple"
ORIGIN

Alignment Scores:

Pred. No.:	3.19e+03	Length:	463
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x DN851694 (1-463)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        155 TCGCTGATAACTTGGAACATT 175
```

RESULT 14
CR543841
LOCUS CR543841 471 bp mRNA linear EST 07-JUL-2004
DEFINITION DKFZp459O0742_r1 459 (synonym: pcor1) Pongo pygmaeus cDNA clone DKFZp459O0742 5', mRNA sequence.
ACCESSION CR543841
VERSION CR543841.1 GI:49895253
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE 1 (bases 1 to 471)
AUTHORS Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
TITLE Pongo pygmaeus mRNA (Bloecker,H., Boecher,M., Brandt,P., et al.)
JOURNAL Unpublished (2004)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp459O0742) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

```
FEATURES             Location/Qualifiers
     source          1. .471
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp459O0742"
                     /tissue_type="cortex"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="459 (synonym: pcor1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN
```

Alignment Scores:

Pred. No.:	3.26e+03	Length:	471
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x CR543841 (1-471)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        344 TCTCTCATTACCTGGAATATT 364
```

```
RESULT 15
CB120234
LOCUS       CB120234                473 bp    mRNA    linear   EST 28-JAN-2003
DEFINITION  K-EST0167337 L8SCK0 Homo sapiens cDNA clone L8SCK0-29-F06 5', mRNA
            sequence.
ACCESSION   CB120234
VERSION     CB120234.1  GI:27946036
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 473)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 29  row: F  column: 06
            High quality sequence stop: 473.
FEATURES             Location/Qualifiers
     source          1. .473
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L8SCK0-29-F06"
                     /sex="M"
                     /cell_line="SCK"
                     /lab_host="Top10F'"
                     /clone_lib="L8SCK0"
                     /note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI;
                     Site_2: NotI; The library was contributed by the Soares
                     laboratory and it was constructed as described by Bonaldo,
                     M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
                     6(9): 791-806. RNA was prepared from harvested cell
                     culture."
ORIGIN
```

Alignment Scores:

Pred. No.:	3.27e+03	Length:	473
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x CB120234 (1-473)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        292 TCTCTCATTACCTGGAATATT 312
```

```
RESULT 16
AI750554
LOCUS       AI750554                480 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  cn04a01.y1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
            clone NHTBC_cn04a01 random, mRNA sequence.
ACCESSION   AI750554
VERSION     AI750554.1  GI:5128818
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
            Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
            Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
  TITLE     SGAP: The Skeletal Genome Anatomy Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Libin Jia
            Medical Genetics Branch
            National Human Genome Research Institute
            10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
            Tel: 301-402-4877
            Fax: 301-496-7157
            Email: libin@helix.nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 04  row: a  column: 01
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NHTBC_cn04a01"
                     /sex="Female"
                     /tissue_type="Bone"
                     /cell_type="Trabecular Bone Cells"
                     /lab_host="SURE"
                     /clone_lib="Normal Human Trabecular Bone Cells"
                     /note="Organ: Hip; Vector: pBluescript; Site_1: EcoRI;
                     Library constructed by Dr. Marian Young and Dr. Pamela
                     Gehron Robey (NIDCR)"
ORIGIN
```

Alignment Scores:

Pred. No.:	3.33e+03	Length:	480
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	1	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x AI750554 (1-480)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        337 TCTCTCATTACCTGGAATATT 357
```

```
RESULT 17
AA486032
LOCUS       AA486032                485 bp    mRNA    linear   EST 06-MAR-1998
DEFINITION  ab40b10.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone
            IMAGE:843259 5', mRNA sequence.
ACCESSION   AA486032
VERSION     AA486032.1  GI:2216248
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 485)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1913   Std Error: 0.00
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .485
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:843259"
                     /sex="female"
                     /dev_stage="HeLa S3 cell line"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene HeLa cell s3 937216"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. HeLa S3
                     epithelioid carcinoma cells grown to semi-confluency
                     without induction. Average insert size: 1.5 kb; Uni-ZAP XR
                     Vector. ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

Alignment Scores:
Pred. No.:              3.37e+03        Length:         485
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AA486032 (1-485)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        333 TCTCTCATTACCTGGAATATT 353

RESULT 18
BU659432/c
LOCUS       BU659432                511 bp    mRNA    linear   EST 30-SEP-2002
DEFINITION  cl47f06.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo
            sapiens cDNA clone cl47f06 5', mRNA sequence.
ACCESSION   BU659432
VERSION     BU659432.1  GI:23371614
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 511)
  AUTHORS   Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
  TITLE     Gene Expression in Human Erythroid Precursor Cells
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Jeffery L. Miller
            Laboratory of Chemical Biology
            National Institute of Diabetes and Digestive and Kidney Diseases
            Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
            20892, USA
            Tel: 301 402 2373
            Fax: 301 435 5148
            Email: jm7f@nih.gov
            The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
            Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
            analyses by National Institutes of Health Intramural Sequencing
            Center (NISC). More information available at:
            http://hembase.niddk.nih.gov
            Plate: 47  row: f  column: 06
            Seq primer: 5' lambda-TriplEx2 Sequencing Primer.
FEATURES             Location/Qualifiers
     source          1. .511
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="cl47f06"
                     /sex="unknown"
                     /tissue_type="blood"
                     /cell_type="Erythroid Precursor Cells"
                     /cell_line="Primary Culture of Peripheral Blood
                     Mononuclear Cells"
                     /dev_stage="Precursor erythroblasts; GPA++"
                     /lab_host="DH5alpha"
                     /clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl
                     library)"
                     /note="Organ: blood; Vector: pTriplEx2; Site_1: SfiI;
                     Site_2: SfiI; A complementary DNA (cDNA) library from
                     human erythroid precursor cells was constructed using
                     SMART PCR (polymerase chain reaction) cDNA Library
                     Construction Kit (Clontech, Palo Alto, CA) according to
                     the manufacturer's directions, but with slight
                     modifications. Briefly, reverse transcription was
                     performed in the presence of 1 umol/L peptide nucleic acid
                     (PNA) oligos
                     (N-terminal)-biotin-GTC-CAC-CCG-AAG-CTT-G-(C-terminal) and
                     (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)
                     . Synthesized cDNA was digested with SfiI and
                     size-selected on a 1% agarose gel (>800bp). Large-scale
                     sequencing of the library was performed by the NIH
                     Intramural Sequencing Center (NISC;
                     Http://www.nisc.nih.gov/)."
ORIGIN

Alignment Scores:
Pred. No.:              3.59e+03        Length:         511
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BU659432 (1-511)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db         30 TCATTGATAACTTGGAATATA 10

RESULT 19
CC153114
LOCUS       CC153114                518 bp    DNA     linear   GSS 25-APR-2003
```

```
DEFINITION  CSU-K34.128K7.SP6 CSU-K34 Aedes aegypti genomic clone
            CSU-K34-128K7, genomic survey sequence.
ACCESSION   CC153114
VERSION     CC153114.1  GI:30106410
KEYWORDS    GSS.
SOURCE      Aedes aegypti (yellow fever mosquito)
  ORGANISM  Aedes aegypti
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Culicidae; Culicinae; Culicini; Aedes; Stegomyia.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Loftus,B., Shetty,J., Knudson,D. and Severson,D.
  TITLE     BAC end sequencing of Aedes aegypti
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: CSU-K34.128K7.T7
            Contact: Brendan Loftus
            Department of Eukaryotic Genomics
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-3543
            Fax: 301-838-0208
            Email: enta@tigr.org
            Library was provided by Susan Brown and Dennis Knudson at Colorado
            State University.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="Aedes aegypti"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7159"
                     /clone="CSU-K34-128K7"
                     /clone_lib="CSU-K34"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Source DNA: Aedes
                     aegypti; strain unknown [derived from freshly hatched
                     larvae at the Virus Research Centre, Poona, India.
                     Reference: SINGH, K. R. P., 1967 Cell cultures derived
                     from larvae of Aedes albopictus (Skuse) and Aedes aegypti
                     (L.). Current Science 36: 506-508]; ATC-10 cell line ATCC
                     CCL-125"
ORIGIN
```

Alignment Scores:

Pred. No.:	3.65e+03	Length:	518
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x CC153114 (1-518)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        167 TCGCTCATAACCTGGAACATT 187
```

```
RESULT 20
BM846221
LOCUS       BM846221                530 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0125204 S13KMS5 Homo sapiens cDNA clone S13KMS5-59-C04 5',
            mRNA sequence.
ACCESSION   BM846221
VERSION     BM846221.1  GI:19202620
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 59  row: C  column: 04
            High quality sequence stop: 530.
FEATURES             Location/Qualifiers
     source          1. .530
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S13KMS5-59-C04"
                     /tissue_type="myeloma"
                     /cell_line="KMS-5"
                     /lab_host="Top10F'"
                     /clone_lib="S13KMS5"
                     /note="Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly
                     (A)+ RNA was dephosphorylated with bacterial alkaline
                     phosphatase (BAP) and then decapped with tabacco acid
                     pyrophosphatase (TAP). The decapped intact mRNA was
                     ligated with DNA-RNA linker including EcoR I site by
                     treatment of T4 RNA ligase and the first strand cDNA was
                     synthesized from oligo dT-selected mRNA by priming with
                     dT-tailed vector. The dT-tailed vector was adjusted to
                     have about 60nt. The cDNA vector was circularized with E.
                     coli DNA ligase after digestion of EcoRI which site is
                     also included in vector. An RNA strand converted to a DNA
                     strand by Okayama-Berg method. The obtained cDNA vectors
                     were used for transformation of competent cells E. coli
                     Top10F' by electroporation method. The cDNA libraries
                     constructed by this method are full-length enriched cDNA
                     library."
ORIGIN
```

Alignment Scores:

Pred. No.:	3.75e+03	Length:	530
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BM846221 (1-530)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        347 TCTCTCATTACCTGGAATATT 367
```

```
RESULT 21
CN233559
LOCUS       CN233559                531 bp    mRNA    linear   EST 09-APR-2004
DEFINITION  RJA113F02.ab1 RJbrain Gallus gallus cDNA 5', mRNA sequence.
ACCESSION   CN233559
VERSION     CN233559.1  GI:46337303
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Savolainen,P., Fitzsimmons,C.J., Arvestad,L., Andersson,L. and
            Lundeberg,J.
  TITLE     EST analysis of brain and testis cDNA libraries from White leghorn
            and Red Jungle Fowl
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Peter Savolainen
            Department of Biotechnology
```

```
            Royal Institute of Technology, KTH
            SE-106 91 Stockholm, SWEDEN
            Tel: +46 (0)8 5537 8481
            Pax: +46 (0)8 5537 8335
            Email: Peter.Savolainen@biotech.kth.se
            Seq primer: M13 reverse primer.
FEATURES              Location/Qualifiers
     source           1. .531
                      /organism="Gallus gallus"
                      /mol_type="mRNA"
                      /strain="Red junglefowl"
                      /db_xref="taxon:9031"
                      /sex="female"
                      /lab_host="ElectroMAX DH10B (Invitrogen)"
                      /clone_lib="RJbrain"
                      /note="Organ: brain; Vector: pSPORT-1; Site_1: Hind III;
                      Site_2: EcoRI; The cDNA libraries were created with the
                      Superscript Plasmid System  (Invitrogen)."
ORIGIN

Alignment Scores:
Pred. No.:              3.76e+03        Length:         531
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x CN233559 (1-531)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        303 TCGCTGATAACTTGGAACATT 323

RESULT 22
BP220985
LOCUS       BP220985                554 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP220985 Sugano cDNA library, colon Homo sapiens cDNA clone
            COL06450, mRNA sequence.
ACCESSION   BP220985
VERSION     BP220985.1  GI:52093890
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 554)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES              Location/Qualifiers
     source           1. .554
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="COL06450"
                      /tissue_type="colon"
                      /clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:
Pred. No.:              3.95e+03        Length:         554
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP220985 (1-554)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        252 TCTCTCATTACCTGGAATATT 272

RESULT 23
BX997515/c
LOCUS       BX997515                558 bp    DNA     linear   GSS 05-JUL-2004
DEFINITION  Forward strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHPP341i12, genomic survey sequence.
ACCESSION   BX997515
VERSION     BX997515.1  GI:49728973
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES              Location/Qualifiers
     source           1. .558
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:10090"
                      /clone="MHPP341i12"
                      /clone_lib="MHPP"
ORIGIN

Alignment Scores:
Pred. No.:              3.99e+03        Length:         558
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     10              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BX997515 (1-558)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        472 TCGTTGATTACTTGGAACATT 452

RESULT 24
BP226620
LOCUS       BP226620                565 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP226620 Sugano cDNA library, dermoid cancer Homo sapiens cDNA
            clone DMC03190, mRNA sequence.
ACCESSION   BP226620
VERSION     BP226620.1  GI:52099525
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 565)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
```

```
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .565
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DMC03190"
                     /clone_lib="Sugano cDNA library, dermoid cancer"
                     /note="dermoid cancer"
ORIGIN

Alignment Scores:
Pred. No.:              4.05e+03        Length:         565
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP226620 (1-565)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        344 TCTCTCATTACCTGGAATATT 364

RESULT 25
BP220508
LOCUS       BP220508                  566 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP220508 Sugano cDNA library, colon Homo sapiens cDNA clone
            COL03160, mRNA sequence.
ACCESSION   BP220508
VERSION     BP220508.1  GI:52093413
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .566
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL03160"
                     /tissue_type="colon"
                     /clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:
Pred. No.:              4.06e+03        Length:         566
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP220508 (1-566)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        262 TCTCTCATTACCTGGAATATT 282

RESULT 26
BP257211
LOCUS       BP257211                  568 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP257211 Sugano cDNA library, heart Homo sapiens cDNA clone
            HRT00430, mRNA sequence.
ACCESSION   BP257211
VERSION     BP257211.1  GI:52172441
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 568)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .568
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRT00430"
                     /tissue_type="heart"
                     /clone_lib="Sugano cDNA library, heart"
ORIGIN

Alignment Scores:
Pred. No.:              4.07e+03        Length:         568
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP257211 (1-568)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        313 TCTCTCATTACCTGGAATATT 333

RESULT 27
BM172060
LOCUS       BM172060                  570 bp    mRNA    linear   EST 04-DEC-2001
DEFINITION  imageqc_3_2001/sml293bdff41.x1 NIH_MGC_76 Homo sapiens cDNA clone
            IMAGE:4700059 5', mRNA sequence.
ACCESSION   BM172060
VERSION     BM172060.1  GI:17311623
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Kale,P.I., Harsch,T.J., Folta,P.A., Nelson,D.O., Sanders,C.G. and
            Prange,C.K.
  TITLE     The I.M.A.G.E. Consortium quality control effort: clone
            resequencing for verification
  JOURNAL   Unpublished (2001)
```

```
COMMENT     Other_ESTs: BG533717
            Contact: Prange CK
            The I.M.A.G.E. Consortium
            Lawrence Livermore National Laboratory
            Livermore, CA, USA
            Email: help@image.llnl.gov
            This read has been verified (found to hit its original self in the
            correct orientation), as part of the I.M.A.G.E. Consortium quality
            control effort. High quality sequence is defined as having 100 or
            more base pairs with a phred quality value of 20 or greater, where
            a sliding window of 4 base pairs with a phred quality value of 15
            or greater marks the beginning and end of the sequence. For
            information on obtaining this clone, please contact
            info@image.llnl.gov.
            Plate: LLCM1533  row: a  column: 20
            Seq primer: -21m13
            High quality sequence stop: 570.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4700059"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_76"
                     /note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.85
                     kb (range 1.0-4.0 kb). 15/15 colonies contained  inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                4.09e+03      Length:         570
Score:                    38.00         Matches:        7
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BM172060 (1-570)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        195 TCTCTCATTACCTGGAATATT 215

RESULT 28
BP221518
LOCUS       BP221518                570 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP221518 Sugano cDNA library, colon Homo sapiens cDNA clone
            COL09203, mRNA sequence.
ACCESSION   BP221518
VERSION     BP221518.1  GI:52094423
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL09203"
                     /tissue_type="colon"
                     /clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:
Pred. No.:                4.09e+03      Length:         570
Score:                    38.00         Matches:        7
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP221518 (1-570)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        251 TCTCTCATTACCTGGAATATT 271

RESULT 29
BU397611
LOCUS       BU397611                570 bp    mRNA    linear   EST 27-NOV-2002
DEFINITION  603534313F1 CSEQCHN58 Gallus gallus cDNA clone ChEST493j4 5', mRNA
            sequence.
ACCESSION   BU397611
VERSION     BU397611.1  GI:25766654
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Compton Line 15I"
                     /db_xref="taxon:9031"
                     /clone="ChEST493j4"
                     /sex="Female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN58"
                     /note="Organ: small intestine; Vector: pBluescript II
                     KS(+); Site_1: EcoRI; Site_2: Not1; This normalized
                     library was constructed from 1 million independent clones.
                     cDNA synthesis was initiated using an oligo(dT) primer,
                     using methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
```

compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."
ORIGIN

Alignment Scores:
Pred. No.: 4.09e+03 Length: 570
Score: 38.00 Matches: 7
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x BU397611 (1-570)

Qy 1 SerLeuIleThrTrpAsnIle 7
||||||||||||||||||||
Db 80 TCGCTGATAACTTGGAACATT 100

RESULT 30
BP262485
LOCUS BP262485 572 bp mRNA linear EST 16-SEP-2004
DEFINITION BP262485 Sugano cDNA library, small intestine Homo sapiens cDNA clone HSI07111, mRNA sequence.
ACCESSION BP262485
VERSION BP262485.1 GI:52177716
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 572)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .572
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HSI07111"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.: 4.11e+03 Length: 572
Score: 38.00 Matches: 7
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x BP262485 (1-572)

Qy 1 SerLeuIleThrTrpAsnIle 7
||||||||||||||||||||
Db 299 TCTCTCATTACCTGGAATATT 319

RESULT 31
AU279894
LOCUS AU279894 573 bp mRNA linear EST 31-JUL-2003
DEFINITION AU279894 CHONS2 Homo sapiens cDNA clone CHONS2002038 5', mRNA sequence.
ACCESSION AU279894
VERSION AU279894.1 GI:28299121
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Imabayashi,H., Mori,T., Gojo,S., Kiyono,T., Sugiyama,T., Irie,R., Isogai,T., Hata,J., Tomoya,Y. and Umezawa,A.
TITLE Redifferentiation of dedifferentiated chondrocytes and chondrogenesis of human bone marrow stromal cells via chondrosphere formation with expression profiling by large-scale cDNA analysis
JOURNAL Exp. Cell Res. 288 (1), 35-50 (2003)
PUBMED 12878157
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA Project, Sugiyama,T.; Wakamatsu,A.; Irie,R.; Umezawa,A.; Fukuma,M.; Kusakari,S.; Hata,J.; Ishii,S.; Yamamoto,J.; Isono,Y.; Saito,K.; Nakamura,Y.; Masuho,Y.; Nagai,K.; Isogai,T.
HRI human cDNA project; cDNA library construction & 5'-end one pass sequencing: Helix Research Institute.
FEATURES Location/Qualifiers
source 1. .573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CHONS2002038"
/cell_type="chondrocytes"
/clone_lib="CHONS2"
/note="Vector: pME18SFL3"
ORIGIN

Alignment Scores:
Pred. No.: 4.12e+03 Length: 573
Score: 38.00 Matches: 7
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x AU279894 (1-573)

Qy 1 SerLeuIleThrTrpAsnIle 7
||||||||||||||||||||
Db 380 TCTCTCATTACCTGGAATATT 400

RESULT 32
BP221260
LOCUS BP221260 573 bp mRNA linear EST 15-SEP-2004
DEFINITION BP221260 Sugano cDNA library, colon Homo sapiens cDNA clone COL08083, mRNA sequence.
ACCESSION BP221260
VERSION BP221260.1 GI:52094165
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL08083"
/tissue_type="colon"
/clone_lib="Sugano cDNA library, colon"
ORIGIN

Alignment Scores:

Pred. No.:	4.12e+03	Length:	573
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BP221260 (1-573)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        265 TCTCTCATTACCTGGAATATT 285
```

RESULT 33
BP261500
LOCUS BP261500 573 bp mRNA linear EST 16-SEP-2004
DEFINITION BP261500 Sugano cDNA library, small intestine Homo sapiens cDNA clone HSI03816, mRNA sequence.
ACCESSION BP261500
VERSION BP261500.1 GI:52176731
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HSI03816"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:

Pred. No.:	4.12e+03	Length:	573
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BP261500 (1-573)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        249 TCTCTCATTACCTGGAATATT 269
```

RESULT 34
BP262642
LOCUS BP262642 576 bp mRNA linear EST 16-SEP-2004
DEFINITION BP262642 Sugano cDNA library, small intestine Homo sapiens cDNA clone HSI07584, mRNA sequence.
ACCESSION BP262642
VERSION BP262642.1 GI:52177873
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 576)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HSI07584"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:

Pred. No.:	4.14e+03	Length:	576
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BP262642 (1-576)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        267 TCTCTCATTACCTGGAATATT 287
```

RESULT 35
CN298919
LOCUS CN298919 577 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000600175228 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN298919
VERSION CN298919.1 GI:47315333
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 577)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.

```
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 577   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

Alignment Scores:
Pred. No.:                4.15e+03        Length:          577
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x CN298919 (1-577)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363

RESULT 36
BP261787
LOCUS       BP261787                 580 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP261787 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI05008, mRNA sequence.
ACCESSION   BP261787
VERSION     BP261787.1  GI:52177018
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI05008"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                4.18e+03        Length:          580
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BP261787 (1-580)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        278 TCTCTCATTACCTGGAATATT 298

RESULT 37
BP274368
LOCUS       BP274368                 580 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP274368 Sugano cDNA library, kidney Homo sapiens cDNA clone
            KDN00894, mRNA sequence.
ACCESSION   BP274368
VERSION     BP274368.1  GI:52188100
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KDN00894"
                     /tissue_type="kidney"
                     /clone_lib="Sugano cDNA library, kidney"
ORIGIN

Alignment Scores:
Pred. No.:                4.18e+03        Length:          580
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BP274368 (1-580)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        363 TCTCTCATTACCTGGAATATT 383

RESULT 38
AV717253
LOCUS       AV717253                 581 bp    mRNA    linear   EST 16-OCT-2000
DEFINITION  AV717253 DCB Homo sapiens cDNA clone DCBBMA04 5', mRNA sequence.
ACCESSION   AV717253
VERSION     AV717253.1  GI:10814405
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
            Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
            Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
            Lu,G., Cheng,Z. and Han,Z.
  TITLE     Homo sapiens cDNA DCB clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DCBBMA04"
                     /cell_type="dendritic cells"
                     /dev_stage="mature"
                     /lab_host="BM25.8"
                     /clone_lib="DCB"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
ORIGIN

Alignment Scores:
Pred. No.:              4.18e+03        Length:         581
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AV717253 (1-581)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        186 TCTCTCATTACCTGGAATATT 206

RESULT 39
BP199948
LOCUS       BP199948                581 bp    mRNA    linear   EST 14-SEP-2004
DEFINITION  BP199948 Sugano cDNA library, amygdala Homo sapiens cDNA clone
            AMR00446, mRNA sequence.
ACCESSION   BP199948
VERSION     BP199948.1  GI:52048428
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="AMR00446"
                     /tissue_type="amygdala"
                     /clone_lib="Sugano cDNA library, amygdala"
ORIGIN

Alignment Scores:
Pred. No.:              4.18e+03        Length:         581
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP199948 (1-581)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        379 TCTCTCATTACCTGGAATATT 399

RESULT 40
BP207751
LOCUS       BP207751                581 bp    mRNA    linear   EST 14-SEP-2004
DEFINITION  BP207751 Sugano cDNA library, coronary artery smooth muscle cell
            Homo sapiens cDNA clone CAS11735, mRNA sequence.
ACCESSION   BP207751
VERSION     BP207751.1  GI:52064157
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CAS11735"
                     /tissue_type="coronary artery"
                     /cell_type="smooth muscle cell"
                     /clone_lib="Sugano cDNA library, coronary artery smooth
                     muscle cell"
ORIGIN

Alignment Scores:
Pred. No.:              4.18e+03        Length:         581
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP207751 (1-581)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        251 TCTCTCATTACCTGGAATATT 271

RESULT 41
```

BP225088
LOCUS BP225088 581 bp mRNA linear EST 15-SEP-2004
DEFINITION BP225088 Sugano cDNA library, lymphocyte Daudi Homo sapiens cDNA clone DAT05184, mRNA sequence.
ACCESSION BP225088
VERSION BP225088.1 GI:52097993
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DAT05184"
/cell_type="lymphocyte"
/cell_line="Daudi"
/clone_lib="Sugano cDNA library, lymphocyte Daudi"
/note="Burkitt's lymphoma"
ORIGIN

Alignment Scores:

Pred. No.:	4.18e+03	Length:	581
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BP225088 (1-581)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        364 TCTCTCATTACCTGGAATATT 384
```

RESULT 42
BP262638
LOCUS BP262638 581 bp mRNA linear EST 16-SEP-2004
DEFINITION BP262638 Sugano cDNA library, small intestine Homo sapiens cDNA clone HSI07576, mRNA sequence.
ACCESSION BP262638
VERSION BP262638.1 GI:52177869
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HSI07576"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:

Pred. No.:	4.18e+03	Length:	581
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BP262638 (1-581)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        260 TCTCTCATTACCTGGAATATT 280
```

RESULT 43
BP264149
LOCUS BP264149 581 bp mRNA linear EST 16-SEP-2004
DEFINITION BP264149 Sugano cDNA library, small intestine Homo sapiens cDNA clone HSI14850, mRNA sequence.
ACCESSION BP264149
VERSION BP264149.1 GI:52179380
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HSI14850"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:

Pred. No.:	4.18e+03	Length:	581
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BP264149 (1-581)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
```

```
Db        309 TCTCTCATTACCTGGAATATT 329

RESULT 44
BP270744
LOCUS       BP270744                  581 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP270744 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone KAR05146, mRNA sequence.
ACCESSION   BP270744
VERSION     BP270744.1  GI:52220093
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAR05146"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:              4.18e+03        Length:         581
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP270744 (1-581)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        251 TCTCTCATTACCTGGAATATT 271

RESULT 45
BP349936
LOCUS       BP349936                  581 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP349936 Sugano cDNA library, brain Homo sapiens cDNA clone
            SZR06923, mRNA sequence.
ACCESSION   BP349936
VERSION     BP349936.1  GI:52279921
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SZR06923"
                     /tissue_type="brain"
                     /clone_lib="Sugano cDNA library, brain"
ORIGIN

Alignment Scores:
Pred. No.:              4.18e+03        Length:         581
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP349936 (1-581)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        120 TCTCTCATTACCTGGAATATT 140

RESULT 46
BP379659
LOCUS       BP379659                  581 bp    mRNA    linear   EST 21-SEP-2004
DEFINITION  BP379659 Sugano cDNA library, uterus Homo sapiens cDNA clone
            WMD06642, mRNA sequence.
ACCESSION   BP379659
VERSION     BP379659.1  GI:52412762
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="WMD06642"
                     /tissue_type="uterus"
                     /clone_lib="Sugano cDNA library, uterus"
ORIGIN

Alignment Scores:
Pred. No.:              4.18e+03        Length:         581
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP379659 (1-581)

Qy          1 SerLeuIleThrTrpAsnIle 7
```

```
                    ||||||||||||||||||||
Db        252 TCTCTCATTACCTGGAATATT 272

RESULT 47
BM507091
LOCUS       BM507091                582 bp    mRNA    linear   EST 15-FEB-2002
DEFINITION  ih24h01.y1 Human insulinoma Homo sapiens cDNA 5' similar to
            TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION   BM507091
VERSION     BM507091.1  GI:18678234
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ih24h01.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 269.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact Hiroshi Inoue, MD/PhD for further
                     information on this library (Metabolism Division, Permutt
                     Laboratory, Washington University School of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:                4.19e+03        Length:          582
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BM507091 (1-582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||

Db        145 TCTCTCATTACCTGGAATATT 165

RESULT 48
BP219740
LOCUS       BP219740                582 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP219740 Sugano cDNA library, caudate nucleus Homo sapiens cDNA
            clone CNR08985, mRNA sequence.
ACCESSION   BP219740
VERSION     BP219740.1  GI:52092643
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CNR08985"
                     /tissue_type="caudate nucleus"
                     /clone_lib="Sugano cDNA library, caudate nucleus"
ORIGIN

Alignment Scores:
Pred. No.:                4.19e+03        Length:          582
Score:                    38.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BP219740 (1-582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||

Db        358 TCTCTCATTACCTGGAATATT 378

RESULT 49
BP233592
LOCUS       BP233592                582 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP233592 Sugano cDNA library, coronary artery endothelial cell Homo
            sapiens cDNA clone HCR00774, mRNA sequence.
ACCESSION   BP233592
VERSION     BP233592.1  GI:52106502
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
```

```
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HCR00774"
                     /tissue_type="coronary artery"
                     /cell_type="endothelial cell"
                     /clone_lib="Sugano cDNA library, coronary artery
                     endothelial cell"
ORIGIN

Alignment Scores:
Pred. No.:              4.19e+03        Length:         582
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP233592 (1-582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        358 TCTCTCATTACCTGGAATATT 378

RESULT 50
BP234218
LOCUS       BP234218                582 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP234218 Sugano cDNA library, coronary artery endothelial cell Homo
            sapiens cDNA clone HCR02378, mRNA sequence.
ACCESSION   BP234218
VERSION     BP234218.1  GI:52107128
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HCR02378"
                     /tissue_type="coronary artery"
                     /cell_type="endothelial cell"
                     /clone_lib="Sugano cDNA library, coronary artery
                     endothelial cell"
ORIGIN

Alignment Scores:
Pred. No.:              4.19e+03        Length:         582
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP234218 (1-582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        370 TCTCTCATTACCTGGAATATT 390

RESULT 51
BP261141
LOCUS       BP261141                582 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP261141 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI02606, mRNA sequence.
ACCESSION   BP261141
VERSION     BP261141.1  GI:52176372
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI02606"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:              4.19e+03        Length:         582
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP261141 (1-582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        249 TCTCTCATTACCTGGAATATT 269

RESULT 52
BP263516
LOCUS       BP263516                582 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP263516 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI11929, mRNA sequence.
ACCESSION   BP263516
VERSION     BP263516.1  GI:52178747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
```

```
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI11929"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                4.19e+03        Length:         582
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP263516 (1-582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        309 TCTCTCATTACCTGGAATATT 329

RESULT 53
BP270415
LOCUS       BP270415                  582 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP270415 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone KAR02166, mRNA sequence.
ACCESSION   BP270415
VERSION     BP270415.1  GI:52219764
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAR02166"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                4.19e+03        Length:         582
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP270415 (1-582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        341 TCTCTCATTACCTGGAATATT 361

RESULT 54
BP275810
LOCUS       BP275810                  582 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP275810 Sugano cDNA library, kidney Homo sapiens cDNA clone
            KDN04583, mRNA sequence.
ACCESSION   BP275810
VERSION     BP275810.1  GI:52189542
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KDN04583"
                     /tissue_type="kidney"
                     /clone_lib="Sugano cDNA library, kidney"
ORIGIN

Alignment Scores:
Pred. No.:                4.19e+03        Length:         582
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP275810 (1-582)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        332 TCTCTCATTACCTGGAATATT 352

RESULT 55
BP195721
LOCUS       BP195721                  583 bp    mRNA    linear   EST 14-SEP-2004
DEFINITION  BP195721 Sugano cDNA library, brain Homo sapiens cDNA clone
            ADB08961, mRNA sequence.
ACCESSION   BP195721
VERSION     BP195721.1  GI:52039998
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
```

```
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="ADB08961"
                     /tissue_type="brain"
                     /clone_lib="Sugano cDNA library, brain"
ORIGIN

Alignment Scores:
Pred. No.:                4.2e+03        Length:          583
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       3              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BP195721 (1-583)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        400 TCTCTCATTACCTGGAATATT 420

RESULT 56
BP262103
LOCUS       BP262103                583 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262103 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI05902, mRNA sequence.
ACCESSION   BP262103
VERSION     BP262103.1  GI:52177334
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI05902"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                4.2e+03        Length:          583
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       3              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BP262103 (1-583)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        277 TCTCTCATTACCTGGAATATT 297

RESULT 57
BP262741
LOCUS       BP262741                583 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262741 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI07919, mRNA sequence.
ACCESSION   BP262741
VERSION     BP262741.1  GI:52177972
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI07919"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                4.2e+03        Length:          583
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       3              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BP262741 (1-583)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        213 TCTCTCATTACCTGGAATATT 233

RESULT 58
BP319548
LOCUS       BP319548                583 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP319548 Sugano cDNA library, pericardium Homo sapiens cDNA clone
            PCD11128, mRNA sequence.
ACCESSION   BP319548
VERSION     BP319548.1  GI:52248523
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
```

```
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
  PUBMED    15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PCD11128"
                     /tissue_type="pericardium"
                     /clone_lib="Sugano cDNA library, pericardium"
ORIGIN

Alignment Scores:
Pred. No.:                4.2e+03        Length:          583
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       3              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BP319548 (1-583)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        392 TCTCTCATTACCTGGAATATT 412

RESULT 59
BU783229
LOCUS       BU783229                583 bp    mRNA    linear   EST 11-OCT-2002
DEFINITION  in01a08.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123158
            5' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION   BU783229
VERSION     BU783229.1  GI:23827207
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 583)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: in01a08.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 430.
FEATURES             Location/Qualifiers
     source          1. .583
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6123158"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact Hiroshi Inoue, MD/PhD for further
                     information on this library (Metabolism Division, Permutt
                     Laboratory, Washington University School of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:                4.2e+03        Length:          583
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              100.0%         Indels:          0
DB:                       5              Gaps:            0

US-10-757-745-2_COPY_115_121 (1-7) x BU783229 (1-583)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        341 TCTCTCATTACCTGGAATATT 361

RESULT 60
BP263059
LOCUS       BP263059                584 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP263059 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI10591, mRNA sequence.
ACCESSION   BP263059
VERSION     BP263059.1  GI:52178290
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
  PUBMED    15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .584
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI10591"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:                4.21e+03       Length:          584
Score:                    38.00          Matches:         7
Percent Similarity:       100.0%         Conservative:    0
```

```
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP263059 (1-584)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        266 TCTCTCATTACCTGGAATATT 286

RESULT 61
BP274942
LOCUS       BP274942                584 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP274942 Sugano cDNA library, kidney Homo sapiens cDNA clone
            KDN02414, mRNA sequence.
ACCESSION   BP274942
VERSION     BP274942.1  GI:52188674
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .584
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KDN02414"
                     /tissue_type="kidney"
                     /clone_lib="Sugano cDNA library, kidney"
ORIGIN

Alignment Scores:
Pred. No.:              4.21e+03        Length:         584
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP274942 (1-584)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        310 TCTCTCATTACCTGGAATATT 330

RESULT 62
BP348623
LOCUS       BP348623                584 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP348623 Sugano cDNA library, brain Homo sapiens cDNA clone
            SZR01143, mRNA sequence.
ACCESSION   BP348623
VERSION     BP348623.1  GI:52278608
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .584
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SZR01143"
                     /tissue_type="brain"
                     /clone_lib="Sugano cDNA library, brain"
ORIGIN

Alignment Scores:
Pred. No.:              4.21e+03        Length:         584
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP348623 (1-584)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        373 TCTCTCATTACCTGGAATATT 393

RESULT 63
BP236074
LOCUS       BP236074                588 bp    mRNA    linear   EST 15-SEP-2004
DEFINITION  BP236074 Sugano cDNA library, coronary artery endothelial cell Homo
            sapiens cDNA clone HCR06784, mRNA sequence.
ACCESSION   BP236074
VERSION     BP236074.1  GI:52108984
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 588)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .588
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HCR06784"
                     /tissue_type="coronary artery"
                     /cell_type="endothelial cell"
                     /clone_lib="Sugano cDNA library, coronary artery
                     endothelial cell"
ORIGIN

Alignment Scores:
```

```
Pred. No.:                4.24e+03        Length:         588
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP236074 (1-588)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        379 TCTCTCATTACCTGGAATATT 399

RESULT 64
BP348796
LOCUS       BP348796                 588 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP348796 Sugano cDNA library, brain Homo sapiens cDNA clone
            SZR01880, mRNA sequence.
ACCESSION   BP348796
VERSION     BP348796.1  GI:52278781
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 588)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
  PUBMED    15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .588
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SZR01880"
                     /tissue_type="brain"
                     /clone_lib="Sugano cDNA library, brain"
ORIGIN

Alignment Scores:
Pred. No.:                4.24e+03        Length:         588
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BP348796 (1-588)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        370 TCTCTCATTACCTGGAATATT 390

RESULT 65
BI064472
LOCUS       BI064472                 600 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  pgf1n.pk001.h17 normalized chicken fat cDNA library Gallus gallus
            cDNA clone pgf1n.pk001.h17 5' similar to gi|7705262
            ref|NP_057698.1| TRAF and TNF receptor-associated protein [Homo
            sapiens] gb|AAF64144.1|AF223469_1 (AF223469) AD022 protein [Homo
            sapiens]G, mRNA sequence.
ACCESSION   BI064472
VERSION     BI064472.1  GI:14471994
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Cogburn,L.A., Morgan,R.W. and Burnside,J.
  TITLE     Chicken ESTs from fat
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9031"
                     /clone="pgf1n.pk001.h17"
                     /sex="Male and Female"
                     /tissue_type="fat"
                     /lab_host="E.coli EMDH10B"
                     /clone_lib="normalized chicken fat cDNA library"
                     /note="Vector: pSPORT1"
ORIGIN

Alignment Scores:
Pred. No.:                4.35e+03        Length:         600
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              100.0%          Indels:         0
DB:                       2               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BI064472 (1-600)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        239 TCGCTGATAACTTGGAACATT 259

RESULT 66
AL703449
LOCUS       AL703449                 602 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686G0621_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686G0621 5', mRNA sequence.
ACCESSION   AL703449
VERSION     AL703449.1  GI:19686804
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   Ansorge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Ansorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686G0621) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
```

```
     source          1. .602
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686G0621"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

Alignment Scores:
Pred. No.:              4.37e+03        Length:         602
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AL703449 (1-602)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        344 TCTCTCATTACCTGGAATATT 364

RESULT 67
AV730762
LOCUS       AV730762                604 bp    mRNA    linear   EST 17-OCT-2000
DEFINITION  AV730762 HTF Homo sapiens cDNA clone HTFAQB10 5', mRNA sequence.
ACCESSION   AV730762
VERSION     AV730762.1  GI:10840183
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 604)
  AUTHORS   Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,
            Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L.,
            Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
            Chen,J., Chen,Z. and Han,Z.
  TITLE     Homo sapiens cDNA HTF clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES              Location/Qualifiers
     source           1. .604
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="HTFAQB10"
                      /tissue_type="Hypothalamus"
                      /dev_stage="Adult"
                      /lab_host="SOLR"
                      /clone_lib="HTF"
                      /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                      XhoI"
ORIGIN

Alignment Scores:
Pred. No.:              4.38e+03        Length:         604
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AV730762 (1-604)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        385 TCACTTATAACCTGGAATATT 405

RESULT 68
CR163431/c
LOCUS       CR163431                618 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Forward strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN168d24, genomic survey sequence.
ACCESSION   CR163431
VERSION     CR163431.1  GI:49942280
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 618)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES              Location/Qualifiers
     source           1. .618
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:10090"
                      /clone="MHPN168d24"
                      /clone_lib="MHPN"
ORIGIN

Alignment Scores:
Pred. No.:              4.5e+03         Length:         618
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x CR163431 (1-618)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        194 TCATTAATAACCTGGAATATC 174

RESULT 69
CB069952
LOCUS       CB069952                621 bp    mRNA    linear   EST 21-JAN-2003
DEFINITION  is31e10.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6554035 5'
            similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION   CB069952
VERSION     CB069952.1  GI:27814472
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
```

```
COMMENT     Other_ESTs: is31e10.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@im.wustl.edu)
            Seq primer: -40UP from Gibco
            High quality sequence stop: 476.
FEATURES              Location/Qualifiers
     source           1. .621
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:6554035"
                      /tissue_type="Purified pancreatic islet"
                      /lab_host="DH10B"
                      /clone_lib="HR85 islet"
                      /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
                      NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
                      Size-selected on agarose gel. Average insert size ~1kb. 5'
                      XhoI site was destroyed after directional cloning.
                      Amplified once. Contact information: Hiroshi Inoue, MD,
                      Metabolism Div. (Alan Permutt Lab), Washington University
                      School of Medicine, Box 8127, 660 South Euclid Ave., St.
                      Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
                      314-362-1916, Fax: 314-747-2692."
ORIGIN
```

Alignment Scores:

Pred. No.:	4.53e+03	Length:	621
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x CB069952 (1-621)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        349 TCTCTCATTACCTGGAATATT 369
```

```
RESULT 70
BU471395
LOCUS       BU471395                625 bp    mRNA    linear   EST 30-NOV-2002
DEFINITION  603363114F1 CSEQRBN21 Gallus gallus cDNA clone ChEST258k3 5', mRNA
            sequence.
ACCESSION   BU471395
VERSION     BU471395.1  GI:25964972
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 625)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES              Location/Qualifiers
     source           1. .625
                      /organism="Gallus gallus"
                      /mol_type="mRNA"
                      /strain="Layer"
                      /db_xref="taxon:9031"
                      /clone="ChEST258k3"
                      /sex="Female"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="CSEQRBN21"
                      /note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
                      EcoRI; Site_2: NotI; This normalized library was
                      constructed from 1 million independent clones. cDNA
                      synthesis was initiated using an oligo(dT) primer, using
                      methylated C in the first strand synthesis reaction.
                      Following this first strand reaction, double-stranded cDNA
                      was blunted, ligated to NotI adapters, digested with
                      EcoRI, size-selected, and cloned into the NotI and EcoRI
                      compatible sites of a custom modified MCS of the
                      pBluescript (KS+) vector. The library was normalized in 2
                      rounds using conditions adapted from Soares et al., PNAS
                      (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                      (1996): 791, except that a significantly longer
                      reannealing hybridization was used."
ORIGIN
```

Alignment Scores:

Pred. No.:	4.56e+03	Length:	625
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_115_121 (1-7) x BU471395 (1-625)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        286 TCGCTGATAACTTGGAACATT 306
```

```
RESULT 71
CB157906
LOCUS       CB157906                634 bp    mRNA    linear   EST 29-JAN-2003
DEFINITION  K-EST0217150 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-5-F06
            5', mRNA sequence.
ACCESSION   CB157906
VERSION     CB157906.1  GI:28143040
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 5  row: F  column: 06
            High quality sequence stop: 634.
FEATURES              Location/Qualifiers
```

```
     source          1. .634
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L18POOL1n1-5-F06"
                     /cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
                     /lab_host="Top10F'"
                     /clone_lib="L18POOL1n1"
                     /note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI;
                     Site_2: NotI; The library was contributed by the Soares
                     laboratory and it was constructed as described by Bonaldo,
                     M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
                     6(9): 791-806. RNA was prepared from harvested cell
                     culture."
ORIGIN

Alignment Scores:
Pred. No.:              4.64e+03        Length:         634
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x CB157906 (1-634)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        295 TCTCTCATTACCTGGAATATT 315

RESULT 72
CV023369
LOCUS       CV023369                634 bp    mRNA    linear   EST 20-AUG-2004
DEFINITION  288 Full Length cDNA from the Mammalian Gene Collection Homo
            sapiens cDNA 5' similar to BC017553, mRNA sequence.
ACCESSION   CV023369
VERSION     CV023369.1  GI:51481130
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
            Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
            Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
            Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
            Vandenhaute,J., Cusick,M.E., Albala,J.S., Hill,D.E. and Vidal,M.
  TITLE     Human ORFeome Version 1.1: a Platform for Reverse Proteomics
  JOURNAL   Genome Res. (2004) In press
COMMENT     Contact: Vidal M
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 5739
            Email: Marc_Vidal@dfci.harvard.edu
            ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
            results from a PCR reaction using an MGC full-length cDNA as
            template DNA and ORF specific primers
            PCR PRimers
            FORWARD: ATGGAGTTGGGGAGTTGCCTGGA
            BACKWARD: TACAATATTATATCTAAGTTGCACAGAAG
            Insert Length: 634    Std Error: 32.00
            Plate: 11008  row: 05  column: F
            Seq primer: ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
            High quality sequence start: 97
            High quality sequence stop: 633
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .634
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /tissue_type="mixed"
                      /clone_lib="Full Length cDNA from the Mammalian Gene
                      Collection"
                      /note="Vector: mixed; The ORFs were PCR amplified from the
                      MGC (Mammalian Gene Collection) as of April 2004 and
                      cloned by recombinational Gateway cloning into pDONR223
                      Donor vector. Reference : MGC (Mammalian Gene Collection)
                      Program Team, Generation and Initial Analysis of more than
                      15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
                      2002, 99(26), 16899-16903"
ORIGIN

Alignment Scores:
Pred. No.:              4.64e+03        Length:         634
Score:                  38.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x CV023369 (1-634)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        343 TCTCTCATTACCTGGAATATT 363

RESULT 73
BI333830
LOCUS       BI333830                644 bp    mRNA    linear   EST 30-JUL-2001
DEFINITION  602999339F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141577 5',
            mRNA sequence.
ACCESSION   BI333830
VERSION     BI333830.1  GI:15018487
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 644)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11348  row: n  column: 10
            High quality sequence stop: 531.
FEATURES              Location/Qualifiers
     source           1. .644
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5141577"
                      /tissue_type="cervical carcinoma cell line"
                      /lab_host="DH10B"
                      /clone_lib="NIH_MGC_12"
                      /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 1.4 kb. Library prepared by Life
                      Technologies."
ORIGIN

Alignment Scores:
Pred. No.:              4.73e+03        Length:         644
```

```
Score:                 38.00          Matches:        7
Percent Similarity:    100.0%         Conservative:   0
Best Local Similarity: 100.0%         Mismatches:     0
Query Match:           100.0%         Indels:         0
DB:                    3              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x BI333830 (1-644)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db        330 TCTCTCATTACCTGGAATATT 350

RESULT 74
AZ360347
LOCUS       AZ360347                644 bp    DNA     linear   GSS 02-OCT-2000
DEFINITION  1M0103M01R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0103M01 R, genomic survey sequence.
ACCESSION   AZ360347
VERSION     AZ360347.1  GI:10474047
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 644)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0103  row: M  column: 01
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 644.
FEATURES             Location/Qualifiers
     source          1. .644
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0103M01"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

Alignment Scores:
Pred. No.:             4.73e+03       Length:         644
Score:                 38.00          Matches:        7
Percent Similarity:    100.0%         Conservative:   0
Best Local Similarity: 100.0%         Mismatches:     0
Query Match:           100.0%         Indels:         0
DB:                    9              Gaps:           0

US-10-757-745-2_COPY_115_121 (1-7) x AZ360347 (1-644)

Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||||
Db         62 TCATTAATAACCTGGAATATC 82

RESULT 75
BM426270
LOCUS       BM426270                645 bp    mRNA    linear   EST 30-JAN-2002
DEFINITION  pgf2n.pk002.c6 Normalized Chicken Abdominal Fat Library (pgf2n)
            Gallus gallus cDNA clone pgf2n.pk002.c6 5' similar to gi|11418470
            ref|XP_004263.1| TRAF and TNF receptor-associated protein [Homo
            sapiens] gi|14747264 ref|XP_041298.1| TRAF and TNF
            receptor-associated protein [Homo sapiens] emb|CAA21141.1|
            (AL031775) dJ30M3.3 (novel protein similar to C. elegans, mRNA
            sequence.
ACCESSION   BM426270
VERSION     BM426270.1  GI:18430455
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Cogburn,L.A., Morgan,R. and Burnside,J.
  TITLE     ESTs from Normalized Chicken fat cDNA library-USDA/IFAFS Animal
            Genome Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Commercial broiler, Ottawa Research Centre,
                     leghorn"
                     /db_xref="taxon:9031"
                     /clone="pgf2n.pk002.c6"
                     /sex="Male and Female"
                     /tissue_type="Abdominal Fat"
                     /dev_stage="Embryonic (d18,d19); post-hatch
                     (d1,w3,w7,w9,w16,1yr)"
                     /lab_host="E. coli EMDH10B"
                     /clone_lib="Normalized Chicken Abdominal Fat Library
                     (pgf2n)"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
                     pools of total RNA isolated  from each developmental age
                     (across strains); Single pass sequencing from 5'-end"
ORIGIN

Alignment Scores:
Pred. No.:             4.74e+03       Length:         645
Score:                 38.00          Matches:        7
Percent Similarity:    100.0%         Conservative:   0
Best Local Similarity: 100.0%         Mismatches:     0
Query Match:           100.0%         Indels:         0
```

DB: 3 Gaps: 0

US-10-757-745-2_COPY_115_121 (1-7) x BM426270 (1-645)

```
Qy          1 SerLeuIleThrTrpAsnIle 7
              |||||||||||||||||||
Db        373 TCGCTGATAACTTGGAACATT 393
```

Search completed: March 11, 2006, 03:39:35
Job time : 201.478 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:40:54 ; Search time 296.731 Seconds
(without alignments)
1724.091 Million cell updates/sec

Title: US-10-757-745-2_COPY_145_153
Perfect score: 45
Sequence: 1 PDVIFLQEV 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170327_4011/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_5954_@runat_09032006_170327_4011 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	45	100.0	176	6	AX351374	AX351374 Sequence
c	2	45	100.0	179	6	AX339803	AX339803 Sequence
	3	45	100.0	444	6	CQ069444	CQ069444 Sequence
	4	45	100.0	483	6	BD076937	BD076937 5' EST of
	5	45	100.0	691	6	CQ078617	CQ078617 Sequence
	6	45	100.0	752	6	BD150065	BD150065 Primer fo
	7	45	100.0	752	6	AX870003	AX870003 Sequence
c	8	45	100.0	1079	6	BD139846	BD139846 Compounds
c	9	45	100.0	1079	6	AR202950	AR202950 Sequence
c	10	45	100.0	1079	6	AR208051	AR208051 Sequence
c	11	45	100.0	1079	6	AX429924	AX429924 Sequence
	12	45	100.0	1261	8	AF201687	AF201687 Homo sapi
	13	45	100.0	1898	6	BD157138	BD157138 Primer fo
	14	45	100.0	1898	6	AX878304	AX878304 Sequence
	15	45	100.0	1898	8	AK002168	AK002168 Homo sapi
	16	45	100.0	1920	6	BD205490	BD205490 CD40-Inte
	17	45	100.0	1920	6	AR594294	AR594294 Sequence
	18	45	100.0	1920	6	AX011599	AX011599 Sequence
	19	45	100.0	1921	8	HSA269473	AJ269473 Homo sapi
	20	45	100.0	1936	8	AF223469	AF223469 Homo sapi
	21	45	100.0	1948	6	AR339398	AR339398 Sequence
	22	45	100.0	1952	8	BC017553	BC017553 Homo sapi
	23	45	100.0	2151	8	HSA420495	AJ420495 Homo sapi
	24	45	100.0	2499	6	E23195	E23195 Topoisomera
	25	45	100.0	20938	8	AY613922	AY613922 Homo sapi
	26	45	100.0	102200	8	HS30M3	AL031775 Human DNA
c	27	45	100.0	150344	14	AC152027	AC152027 Dasypus n
	28	45	100.0	167847	8	CR925830	CR925830 Human DNA
	29	45	100.0	176819	8	CR942205	CR942205 Human DNA
	30	44	97.8	1236	5	CR761829	CR761829 Xenopus t
	31	44	97.8	1312	6	BD205491	BD205491 CD40-Inte
	32	44	97.8	1312	6	AR594295	AR594295 Sequence
	33	44	97.8	1312	6	AX011601	AX011601 Sequence
	34	44	97.8	1943	9	MMU251328	AJ251328 Mus muscu
	35	44	97.8	110000	1	BA000019_50	Continuation (51 o
c	36	44	97.8	245394	14	AC125756	AC125756 Rattus no
	37	44	97.8	256608	9	AL589699	AL589699 Mouse DNA
c	38	43	95.6	110000	1	BA000045_42	Continuation (43 o
	39	43	95.6	110000	1	CP000089_40	Continuation (41 o
c	40	42	93.3	33899	2	CEY63D3A	AL032652 Caenorhab
	41	41	91.1	1451	5	BC083404	BC083404 Danio rer
	42	41	91.1	1455	5	BC097117	BC097117 Danio rer
	43	41	91.1	198752	5	BX957279	BX957279 Zebrafish
	44	41	91.1	200076	5	BX511258	BX511258 Zebrafish
c	45	41	91.1	223042	14	AC102963	AC102963 Rattus no
c	46	41	91.1	261222	14	AC155120	AC155120 Bos tauru
c	47	41	91.1	347071	1	BX640415	BX640415 Bordetell
	48	41	91.1	347356	1	BX640437	BX640437 Bordetell
	49	41	91.1	348251	1	BX640423	BX640423 Bordetell
	50	41	91.1	348706	1	BX640445	BX640445 Bordetell
	51	41	91.1	349028	1	BX640413	BX640413 Bordetell
c	52	40	88.9	18284	15	AP000599	AP000599 Arabidops
c	53	40	88.9	61555	14	AC090334	AC090334 Homo sapi
	54	40	88.9	96128	9	AL663077	AL663077 Mouse DNA
c	55	40	88.9	137836	5	BX545850	BX545850 Zebrafish
	56	40	88.9	158050	8	AC023394	AC023394 Homo sapi
c	57	40	88.9	167610	14	CR974453	CR974453 Sus scrof
c	58	40	88.9	174290	5	BX649417	BX649417 Zebrafish
c	59	40	88.9	176333	9	AC135187	AC135187 Mus muscu
c	60	40	88.9	195087	5	BX000438	BX000438 Zebrafish
c	61	40	88.9	211493	14	CR956635	CR956635 Sus scrof
	62	40	88.9	215046	5	CR318663	CR318663 Zebrafish
c	63	40	88.9	220420	5	BX000452	BX000452 Zebrafish
c	64	40	88.9	250603	14	CR848052	CR848052 Danio rer
	65	40	88.9	302178	1	AE016918	AE016918 Chromobac
	66	39	86.7	587	10	BV547170	BV547170 rrj99b02.
	67	39	86.7	725	10	BV553589	BV553589 S217P6232
	68	39	86.7	739	10	BV542059	BV542059 G591P5309
c	69	39	86.7	748	10	BV518605	BV518605 ssw41e11.
	70	39	86.7	752	10	BV553590	BV553590 S217P6232
c	71	39	86.7	801	10	BV505641	BV505641 qrf49e12.
c	72	39	86.7	834	10	BV547859	BV547859 rnx59b01.
	73	39	86.7	850	10	BV543970	BV543970 qpc96h04.
	74	39	86.7	5395	2	DME302712	AJ302712 Drosophil
c	75	39	86.7	7568	2	AY183922	AY183922 Drosophil

ALIGNMENTS

```
RESULT 1
AX351374
LOCUS       AX351374                  176 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 121 from Patent WO0196390.
ACCESSION   AX351374
VERSION     AX351374.1  GI:18616721
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196390-A 121 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .176
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                0.395          Length:         176
Score:                    45.00          Matches:        9
Percent Similarity:       100.0%         Conservative:   0
Best Local Similarity:    100.0%         Mismatches:     0
Query Match:              100.0%         Indels:         0
DB:                       6              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AX351374 (1-176)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          6 CCAGATGTGATATTTCTACAGGAAGTT 32

RESULT 2
AX339803/c
LOCUS       AX339803                  179 bp    DNA     linear   PAT 10-JAN-2002
DEFINITION  Sequence 50 from Patent WO0196388.
ACCESSION   AX339803
VERSION     AX339803.1  GI:18135784
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Harlocker,S.L. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196388-A 50 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .179
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                0.402          Length:         179
Score:                    45.00          Matches:        9
Percent Similarity:       100.0%         Conservative:   0
Best Local Similarity:    100.0%         Mismatches:     0
Query Match:              100.0%         Indels:         0
DB:                       6              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AX339803 (1-179)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145

RESULT 3
CQ069444
LOCUS       CQ069444                  444 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 5244 from Patent WO0157278.
ACCESSION   CQ069444
VERSION     CQ069444.1  GI:41039313
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 5244 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL031775.1
                     EXPRESSED IN HELA, SIGNAL = 1.9"
ORIGIN

Alignment Scores:
Pred. No.:                1.04           Length:         444
Score:                    45.00          Matches:        9
Percent Similarity:       100.0%         Conservative:   0
Best Local Similarity:    100.0%         Mismatches:     0
Query Match:              100.0%         Indels:         0
DB:                       6              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CQ069444 (1-444)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        260 CCAGATGTGATATTTCTACAGGAAGTT 286

RESULT 4
BD076937
LOCUS       BD076937                  483 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  5' EST of secretory protein expressed in prostate.
ACCESSION   BD076937
VERSION     BD076937.1  GI:22622540
KEYWORDS    JP 2001512013-A/184.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
  TITLE     5' EST of secretory protein expressed in prostate
  JOURNAL   Patent: JP 2001512013-A 184 21-AUG-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001512013-A/184
            PD   21-AUG-2001
            PF   31-JUL-1998 JP 2000505291
            PR   01-AUG-1997 US    08/905144
            PI   JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO  PI
```

```
            LACROIX
            PC   C12N15/09,C07K14/47,C12P21/02,C12Q1/02,C12Q1/68,C12N15/00 CC
            blastn
            CC   identity 99
            CC   region 1. .310
            CC   id T26956
            CC   est
            CC   blastn
            CC   identity 98
            CC   region 1. .315
            CC   id T31666
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 137. .267
            CC   id R14990
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 63. .137
            CC   id R14990
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 1. .50
            CC   id R14990
            CC   est
            CC   Von Heijne matrix
            CC   score 5.2
            CC   seq LCVEFASVASCDA/AV
            FH   Key             Location/Qualifiers
            FT   misc_feature    105. .414
            FT   misc_feature    45. .359
            FT   misc_feature    202. .332
            FT   misc_feature    127. .201
            FT   misc_feature    65. .114
            FT   sig_peptide     1. .120.
FEATURES                Location/Qualifiers
     source             1. .483
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                1.14          Length:         483
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       6             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BD076937 (1-483)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 5
CQ078617
LOCUS       CQ078617                691 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 14417 from Patent WO0157278.
ACCESSION   CQ078617
VERSION     CQ078617.1  GI:41048486
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 14417 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES                Location/Qualifiers
     source             1. .691
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"
                        /note="MAP TO AL031775.1
                        EXPRESSED IN HELA, SIGNAL = 1.9
                        SWISSPROT HIT: P18004, EVALUE 7.00e-01
                        NT HIT: AF201687.1, EVALUE 3.00e-45
                        EST_HUMAN HIT: AV717253.1, EVALUE 5.00e-45"
ORIGIN

Alignment Scores:
Pred. No.:                1.66          Length:         691
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       6             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CQ078617 (1-691)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        122 CCAGATGTGATATTTCTACAGGAAGTT 148

RESULT 6
BD150065
LOCUS       BD150065                752 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD150065
VERSION     BD150065.1  GI:27855823
KEYWORDS    JP 2002191363-A/4908.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 4908 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/4908
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                      Location/Qualifiers
            FT   source          1. .752
            FT                   /organism='Homo sapiens (human)'.
FEATURES                Location/Qualifiers
     source             1. .752
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                1.82          Length:         752
```

```
Score:                  45.00          Matches:        9
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BD150065 (1-752)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 7
AX870003
LOCUS       AX870003                752 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 4908 from Patent EP1074617.
ACCESSION   AX870003
VERSION     AX870003.1  GI:40024866
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 4908 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.82           Length:         752
Score:                  45.00          Matches:        9
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AX870003 (1-752)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 8
BD139846/c
LOCUS       BD139846               1079 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Compounds for immunotherapy and diagnosis of breast cancer and
            methods for their use.
ACCESSION   BD139846
VERSION     BD139846.1  GI:23234791
KEYWORDS    JP 2002507387-A/63.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for immunotherapy and diagnosis of breast cancer and
            methods for their use
  JOURNAL   Patent: JP 2002507387-A 63 12-MAR-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002507387-A/63
            PD   12-MAR-2002
            PF   22-DEC-1998 JP 2000526543
            PR   24-DEC-1997 US    08/998253,24-DEC-1997 US    08/998255 PR
            17-JUL-1998 US    09/118627,17-JUL-1998 US    09/118554 PI   STEVEN
            G REED,JIANGCHUN XU
            PC   C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,
            PC   C07K14/47,
            PC
            C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/ PC
             53,
            PC   G01N33/577,C12N15/00,A61K37/02,C12N5/00
            CC   Compounds for immunotherapy and diagnosis
            of breast cancer and
            CC           methods for
            CC   their use
            FH   Key             Location/Qualifiers
            FT   source          1. .1079
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.65           Length:         1079
Score:                  45.00          Matches:        9
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BD139846 (1-1079)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024

RESULT 9
AR202950/c
LOCUS       AR202950               1079 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 63 from patent US 6365348.
ACCESSION   AR202950
VERSION     AR202950.1  GI:21499210
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for diagnosis of Breast cancer and methods for their use
  JOURNAL   Patent: US 6365348-A 63 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              2.65           Length:         1079
Score:                  45.00          Matches:        9
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AR202950 (1-1079)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024

RESULT 10
```

```
AR208051/c
LOCUS       AR208051                1079 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 63 from patent US 6379951.
ACCESSION   AR208051
VERSION     AR208051.1  GI:21507966
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for immunotherapy of breast cancer and methods for their
            use
  JOURNAL   Patent: US 6379951-A 63 30-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                2.65          Length:          1079
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AR208051 (1-1079)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024

RESULT 11
AX429924/c
LOCUS       AX429924                1079 bp    DNA     linear   PAT 21-JUN-2002
DEFINITION  Sequence 63 from Patent WO0198339.
ACCESSION   AX429924
VERSION     AX429924.1  GI:21541088
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.
  TITLE     Compositions and methods for the therapy and diagnosis of breast
            cancer
  JOURNAL   Patent: WO 0198339-A 63 27-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                2.65          Length:          1079
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AX429924 (1-1079)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024

RESULT 12
AF201687
LOCUS       AF201687                1261 bp    mRNA    linear   PRI 16-MAY-2003
DEFINITION  Homo sapiens ETS1-associated protein 2 (EAP2) mRNA, complete cds.
ACCESSION   AF201687
VERSION     AF201687.1  GI:11493668
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1261)
  AUTHORS   Pei,H., Yordy,J.S., Leng,Q., Zhao,Q., Watson,D.K. and Li,R.
  TITLE     EAPII interacts with ETS1 and modulates its transcriptional
            function
  JOURNAL   Oncogene 22 (18), 2699-2709 (2003)
   PUBMED   12743594
REFERENCE   2  (bases 1 to 1261)
  AUTHORS   Li,R., Pei,H. and Papas,T.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1999) Center for Molecular and Structural
            Biology, Department of Medicine, and Hollings Cancer Center,
            Medical University of South Carolina, 86 Jonathan Lucas St.,
            Charleston, SC 29425, USA
FEATURES             Location/Qualifiers
     source          1. .1261
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22.1-p22.3"
                     /cell_line="Jurkat"
     gene            1. .1261
                     /gene="EAP2"
     CDS             26. .1114
                     /gene="EAP2"
                     /codon_start=1
                     /product="ETS1-associated protein 2"
                     /protein_id="AAG35600.1"
                     /db_xref="GI:11493669"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                3.13          Length:          1261
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       8             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AF201687 (1-1261)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        458 CCAGATGTGATATTTCTACAGGAAGTT 484

RESULT 13
BD157138
LOCUS       BD157138                1898 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD157138
VERSION     BD157138.1  GI:27862896
KEYWORDS    JP 2002191363-A/11981.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE 1  (bases 1 to 1898)
  AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
          Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE   Primer for synthesizing full-length cDNA and use thereof
  JOURNAL Patent: JP 2002191363-A 11981 09-JUL-2002;
          HELIX RESEARCH INSTITUTE
COMMENT   OS   Homo sapiens (human)
          PN   JP 2002191363-A/11981
          PD   09-JUL-2002
          PF   28-JUL-2000 JP 2000280990
          PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
          PI   SAITO,
          PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
          PI   KEIICHI NAGAI,TETSUJI OTSUKI
          PC
          C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
           10,
          PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
           Primer for synthesizing full-length cDNA and use thereof FH   Key
                    Location/Qualifiers
          FT   CDS             (1). .(1086).
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.8          Length:          1898
Score:                    45.00        Matches:         9
Percent Similarity:       100.0%       Conservative:    0
Best Local Similarity:    100.0%       Mismatches:      0
Query Match:              100.0%       Indels:          0
DB:                       6            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BD157138 (1-1898)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 14
AX878304
LOCUS       AX878304                1898 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 13209 from Patent EP1074617.
ACCESSION   AX878304
VERSION     AX878304.1  GI:40033040
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE 1
  AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
          Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE   Primers for synthesising full-length cDNA and their use
  JOURNAL Patent: EP 1074617-A 13209 07-FEB-2001;
          Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .1089
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE90095.1"
                     /db_xref="GI:40033041"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                4.8          Length:          1898
Score:                    45.00        Matches:         9
Percent Similarity:       100.0%       Conservative:    0
Best Local Similarity:    100.0%       Mismatches:      0
Query Match:              100.0%       Indels:          0
DB:                       6            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AX878304 (1-1898)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 15
AK002168
LOCUS       AK002168                1898 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ11306 fis, clone PLACE1010031.
ACCESSION   AK002168
VERSION     AK002168.1  GI:7023882
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE 1
  AUTHORS Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
          Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
          Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
          Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
          Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
          Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
          Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
          Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
          Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
          Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
          Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
          Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
          Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
          Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
          Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
          Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
          Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
          Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
          Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
          Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
          Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
          Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
          Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
          Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
          Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
          Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
          Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
          Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
          Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
  TITLE   Complete sequencing and characterization of 21,243 full-length
          human cDNAs
  JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED 14702039
REFERENCE 2
  AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
          Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
          Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
```

```
            Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
            Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
            Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
            Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1898)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1010031"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="cloning vector: pME18SFL3"
     CDS             1. .1089
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAA92119.1"
                     /db_xref="GI:7023883"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                4.8          Length:        1898
Score:                    45.00        Matches:       9
Percent Similarity:       100.0%       Conservative:  0
Best Local Similarity:    100.0%       Mismatches:    0
Query Match:              100.0%       Indels:        0
DB:                       8            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x AK002168 (1-1898)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 16
BD205490
LOCUS       BD205490               1920 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  CD40-Interacting and TRAF-interacting protein.
ACCESSION   BD205490
VERSION     BD205490.1  GI:33015260
KEYWORDS    JP 2002512796-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1920)
  AUTHORS   Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting protein
  JOURNAL   Patent: JP 2002512796-A 1 08-MAY-2002;
            VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002512796-A/1
            PD   08-MAY-2002
            PF   28-APR-1999 JP 2000546003
            PR   29-APR-1998 EP   98201392.2
            PI   STEFAN MARIA CHRISTIAAN PYPE,
            PI   JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI   DANNY
            FRANCOIS EVELINE HUYLEBROECK
            PC   C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
              A61P37/02,
            PC   A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
            PC   G01N33/566,
            PC   C12N15/00,A61K37/02
            CC   CD40-Interacting and TRAF-interacting protein. FH   Key
                 Location/Qualifiers
            FT   CDS             (20). .(1108).
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.86         Length:        1920
Score:                    45.00        Matches:       9
Percent Similarity:       100.0%       Conservative:  0
Best Local Similarity:    100.0%       Mismatches:    0
Query Match:              100.0%       Indels:        0
DB:                       6            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x BD205490 (1-1920)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        452 CCAGATGTGATATTTCTACAGGAAGTT 478

RESULT 17
AR594294
LOCUS       AR594294               1920 bp    DNA     linear   PAT 15-DEC-2004
DEFINITION  Sequence 1 from patent US 6812203.
ACCESSION   AR594294
VERSION     AR594294.1  GI:56643900
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1920)
  AUTHORS   Pype,S.M.C., Remacle,J.E.E. and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting proteins
  JOURNAL   Patent: US 6812203-A 1 02-NOV-2004;
            Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
            Zwijnaarde;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                4.86         Length:        1920
Score:                    45.00        Matches:       9
Percent Similarity:       100.0%       Conservative:  0
Best Local Similarity:    100.0%       Mismatches:    0
Query Match:              100.0%       Indels:        0
DB:                       6            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x AR594294 (1-1920)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        452 CCAGATGTGATATTTCTACAGGAAGTT 478
```

```
RESULT 18
AX011599
LOCUS       AX011599               1920 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 1 from Patent WO9955859.
ACCESSION   AX011599
VERSION     AX011599.1  GI:9998123
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 1 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             20. .1108
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07555.1"
                     /db_xref="GI:9998124"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                4.86          Length:          1920
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AX011599 (1-1920)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        452 CCAGATGTGATATTTCTACAGGAAGTT 478

RESULT 19
HSA269473
LOCUS       HSA269473              1921 bp    mRNA    linear   PRI 27-MAY-2000
DEFINITION  Homo sapiens mRNA for TRAF and TNF receptor associated protein
            (ttrap gene).
ACCESSION   AJ269473
VERSION     AJ269473.1  GI:8247253
KEYWORDS    TRAF and TNF receptor associated protein; ttrap gene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van
            Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,
            Huylebroeck,D. and Remacle,J.E.
  TITLE     TTRAP, a novel protein that associates with CD40, tumor necrosis
            factor (TNF) receptor-75 and TNF receptor-associated factors
            (TRAFs), and that inhibits nuclear factor-kappa B activation
  JOURNAL   J. Biol. Chem. 275 (24), 18586-18593 (2000)
   PUBMED   10764746
REFERENCE   2  (bases 1 to 1921)
  AUTHORS   Pype,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1999) Pype S., VIB07, Dept. Cell Growth,
            Differentiation and Development, Flanders Interuniversity Institute
            for Biotechnology, CELGEN, K.U.Leuven, Gasthuisberg Campus,
            Herestraat 49, B-3000 Leuven, BELGIUM
COMMENT     Related sequence: AL031775.
FEATURES             Location/Qualifiers
     source          1. .1921
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .1921
                     /gene="ttrap"
     5'UTR           <1. .21
                     /gene="ttrap"
     CDS             22. .1110
                     /gene="ttrap"
                     /function="putative role in signal transduction"
                     /codon_start=1
                     /product="TRAF and TNF receptor associated protein"
                     /protein_id="CAB92966.1"
                     /db_xref="GI:8247254"
                     /db_xref="GOA:O95551"
                     /db_xref="UniProt/TrEMBL:O95551"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
     3'UTR           1111. .1921
                     /gene="ttrap"
     polyA_signal    1900. .1905
                     /gene="ttrap"
ORIGIN

Alignment Scores:
Pred. No.:                4.87          Length:          1921
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       8             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x HSA269473 (1-1921)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        454 CCAGATGTGATATTTCTACAGGAAGTT 480

RESULT 20
AF223469
LOCUS       AF223469               1936 bp    mRNA    linear   PRI 17-APR-2000
DEFINITION  Homo sapiens AD022 protein (AD022) mRNA, complete cds.
ACCESSION   AF223469
VERSION     AF223469.1  GI:7578788
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1936)
  AUTHORS   Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
  TITLE     A novel gene expressed in human adrenal gland
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1936)
  AUTHORS   Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-2000) Chinese National Human Genome Center at
            Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
            Shanghai 201203, China
FEATURES             Location/Qualifiers
     source          1. .1936
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="adrenal gland"
     gene            1. .1936
                     /gene="AD022"
     CDS             17. .1105
                     /gene="AD022"
                     /codon_start=1
                     /product="AD022 protein"
                     /protein_id="AAF64144.1"
                     /db_xref="GI:7578789"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVRFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                4.91          Length:          1936
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       8             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AF223469 (1-1936)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        449 CCAGATGTGATATTTCTACAGGAAGTT 475

RESULT 21
AR339398
LOCUS       AR339398               1948 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 889 from patent US 6569662.
ACCESSION   AR339398
VERSION     AR339398.1  GI:33726255
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Tang,Y.T., Zhou,P. and Drmanac,R.T.
  TITLE     Nucleic acids and polypeptides
  JOURNAL   Patent: US 6569662-A 889 27-MAY-2003;
            Hyseq, Inc.; Sunnyvale, CA
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                4.94          Length:          1948
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AR339398 (1-1948)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        471 CCAGATGTGATATTTCTACAGGAAGTT 497

RESULT 22
BC017553
LOCUS       BC017553               1952 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein, mRNA (cDNA
            clone MGC:9099 IMAGE:3920790), complete cds.
ACCESSION   BC017553
VERSION     BC017553.2  GI:34782842
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1952)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED    12477932
REFERENCE   2  (bases 1 to 1952)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Sep 16, 2003 this sequence version replaced gi:17028464.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 15 Row: m Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 23510347.
FEATURES             Location/Qualifiers
     source          1. .1952
                     /organism="Homo sapiens"
```

```
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:9099 IMAGE:3920790"
                     /tissue_type="Skin, melanotic melanoma."
                     /clone_lib="NIH_MGC_72"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1952
                     /gene="TTRAP"
                     /note="synonyms: AD022, EAP2, MGC9099, dJ30M3.3"
                     /db_xref="GeneID:51567"
                     /db_xref="MIM:605764"
     CDS             38. .1126
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TRAF and TNF receptor-associated protein"
                     /protein_id="AAH17553.1"
                     /db_xref="GI:17028465"
                     /db_xref="GeneID:51567"
                     /db_xref="MIM:605764"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Alignment Scores:
Pred. No.:                4.95          Length:         1952
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BC017553 (1-1952)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        470 CCAGATGTGATATTTCTACAGGAAGTT 496

RESULT 23
HSA420495
LOCUS       HSA420495               2151 bp    mRNA    linear   PRI 23-NOV-2001
DEFINITION  Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1761756.
ACCESSION   AJ420495
VERSION     AJ420495.1  GI:17066359
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Auffray,C., Ansorge,W., Ballabio,A., Estivill,X., Gibson,K.,
            Lehrach,H., Poustka,A. and Lundeberg,J.
  TITLE     The European IMAGE consortium for integrated Molecular analysis of
            human gene transcripts
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2151)
  AUTHORS   Persson,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2001) Persson A., Center for Molecular
            Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagsvagen
            30B, 106 91 Stockholm, SWEDEN
COMMENT     This clone is available royalty-free through IMAGE Consortium
            Distributors. IMPORTANT: This sequence represents the full insert
            of this IMAGE cDNA clone. No attempt has been made to verify
            whether this corresponds to the full-length of the original mRNA
            from which it was derived.
FEATURES             Location/Qualifiers
     source          1. .2151
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="EUROIMAGE 1761756"
                     /clone_lib="NCI_CGAP_Kid3"
     polyA_signal    2105. .2110
ORIGIN

Alignment Scores:
Pred. No.:                5.48          Length:         2151
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x HSA420495 (1-2151)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        659 CCAGATGTGATATTTCTACAGGAAGTT 685

RESULT 24
E23195
LOCUS       E23195                  2499 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Topoisomerase binding protein.
ACCESSION   E23195
VERSION     E23195.1  GI:13024277
KEYWORDS    JP 1999075856-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2499)
  AUTHORS   Takashi,T. and Kazuhiko,Y.
  TITLE     Topoisomerase binding protein
  JOURNAL   Patent: JP 1999075856-A 1 23-MAR-1999;
            TAKASHI TSURUO,CHUGAI PHARMACEUT CO LTD
COMMENT     OS   Unidentified
            PN   JP 1999075856-A/1
            PD   23-MAR-1999
            PF   17-SEP-1997 JP 1997251544
            PR
            PI   TAKASHI TSURUO,KAZUHIKO YAMANE
            PC   C12N15/09,C07K14/47,C12N9/90,C12N15/00
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .2499
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2499
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                6.41          Length:         2499
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       6             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x E23195 (1-2499)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        914 CCAGATGTGATATTTCTACAGGAAGTT 940

RESULT 25
```

```
AY613922
LOCUS       AY613922               20938 bp    DNA     linear   PRI 11-MAY-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein (TTRAP) gene,
            complete cds.
ACCESSION   AY613922
VERSION     AY613922.1  GI:47060296
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 20938)
  AUTHORS   Rieder,M.J., Daniels,R.L., da Ponte,S.H., Hastings,N.C.,
            Ahearn,M.O., Rajkumar,N., Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2004) Genome Sciences, University of Washington,
            1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  SeattleSNPs. NHLBI HL66682 Program
            for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.gs.washington.edu).
FEATURES             Location/Qualifiers
     source          1. .20938
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     repeat_region   1. .32
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     variation       81
                     /frequency="0.67"
                     /replace=""
     repeat_region   87. .368
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     repeat_region   369. .673
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       449
                     /frequency="0.96"
                     /replace=""
     variation       540
                     /frequency="0.01"
                     /replace="t"
     variation       586
                     /frequency="0.01"
                     /replace="a"
     variation       650
                     /frequency="0.03"
                     /replace="g"
     variation       658. .660
                     /frequency="0.01"
                     /replace=""
     repeat_region   683. .989
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       695
                     /frequency="0.18"
                     /replace="t"
     variation       922
                     /frequency="0.02"
                     /replace="t"
     variation       1319
                     /frequency="0.15"
                     /replace="t"
     variation       1456
                     /frequency="0.19"
                     /replace="a"
     variation       1532
                     /frequency="0.01"
                     /replace="c"
     variation       1711. .1713
                     /frequency="0.04"
                     /replace=""
     variation       1795
                     /frequency="0.10"
                     /replace="a"
     variation       1821
                     /frequency="0.02"
                     /replace="t"
     variation       1830. .1862
                     /frequency="0.04"
                     /replace=""
     variation       1913. .1914
                     /frequency="0.13"
                     /replace=""
     gene            <2005. .18938
                     /gene="TTRAP"
     mRNA            join(<2005. .2194,2281. .2366,10172. .10345,11003. .11094,
                     14376. .14494,15753. .15923,17846. .18938)
                     /gene="TTRAP"
                     /product="TRAF and TNF receptor associated protein"
     CDS             join(2030. .2194,2281. .2366,10172. .10345,11003. .11094,
                     14376. .14494,15753. .15923,17846. .18127)
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TRAF and TNF receptor associated protein"
                     /protein_id="AAT09764.1"
                     /db_xref="GI:47060297"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
     variation       2381
                     /gene="TTRAP"
                     /frequency="0.08"
                     /replace="a"
     variation       2767
                     /gene="TTRAP"
                     /frequency="0.08"
                     /replace="g"
     variation       2800
                     /gene="TTRAP"
                     /frequency="0.79"
                     /replace=""
     repeat_region   2919. .3323
                     /rpt_family="L2"
                     /rpt_type=dispersed
     variation       2927
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="g"
     variation       3004
                     /gene="TTRAP"
                     /frequency="0.04"
                     /replace="t"
     variation       3061
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="t"
     variation       3062
                     /gene="TTRAP"
                     /frequency="0.13"
                     /replace="t"
     variation       3114
                     /gene="TTRAP"
                     /frequency="0.21"
                     /replace="c"
     variation       3418
                     /gene="TTRAP"
                     /frequency="0.03"
                     /replace="c"
     variation       3781
```

```
                     /gene="TTRAP"
                     /frequency="0.15"
                     /replace="g"
     variation       3988
                     /gene="TTRAP"
                     /frequency="0.01"
                     /replace="a"
     variation       4191
                     /gene="TTRAP"
                     /frequency="0.01"
                     /replace="c"
     variation       4393
                     /gene="TTRAP"
                     /frequency="0.97"
                     /replace=""
     variation       4414
                     /gene="TTRAP"
                     /frequency="0.03"
                     /replace="c"
     variation       4437
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="g"
     repeat_region   4446. .4530
                     /rpt_family="MER1_type"
                     /rpt_type=dispersed
     repeat_region   4559. .4609
                     /rpt_family="MER1_type"
                     /rpt_type=dispersed
     variation       4565
                     /gene="TTRAP"
                     /frequency="0.52"
                     /replace=""
     repeat_region   4625. .4760
                     /rpt_family="MER1_type"
                     /rpt_type=dispersed
     variation       4658
                     /gene="TTRAP"
                     /frequency="0.01"
                     /replace="a"
     variation       4861
                     /gene="TTRAP"
                     /frequency="0.02"
                     /replace="a"
     variation       4941
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="g"
     variation       4983
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="t"
     variation       5024
                     /gene="TTRAP"
                     /frequency="0.07"

Alignment Scores:
Pred. No.:              59.8          Length:         20938
Score:                  45.00         Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     8             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AY613922 (1-20938)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db      11010 CCAGATGTGATATTTCTACAGGAAGTT 11036

RESULT 26
HS30M3
LOCUS       HS30M3                102200 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP1-30M3 on chromosome 6p22.1-22.3
            Contains the 5' end of gene KIAA0319, the gene for TRAF and TNF
            receptor-associated protein (TTRAP)(EAP2, AD022, MGC9099), gene
            FLJ20501, the 3' end of gene FLJ12619 (DKFZP564G182) and two CpG
            islands, complete sequence.
ACCESSION   AL031775
VERSION     AL031775.1  GI:4071041
KEYWORDS    HTG; AD022; CpG island; DKFZP564G182; FLJ12619; FLJ20501; KIAA0319;
            TNF; TRAF; TTRAP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 102200)
  AUTHORS   Phillips,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 29, 1998 this sequence version replaced gi:4007165.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP1-30M3 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .102200
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704M0330"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="p22.1-22.3"
                     /clone="RP1-30M3"
                     /clone_lib="RPCI-1"
     gene            join(complement(AL133264.10:4611. .6907),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-001"
     mRNA            join(complement(AL133264.10:4611. .6907),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-001"
                     /product="chromosome 6 open reading frame 62"
                     /note="match: ESTs: AA084220 AA093151 AA168907 AA197971
                     AA200765 AA237701 AA292313 AA295166 AA402388 AA478637
                     AA591228 AA612170 AA638148 AA647423 AA647435 AA647742
                     AA672357 AA798123 AA798467 AA815944 AA833812 AA863774
                     AI053561 AI054330 AI122327 AI156738 AI191279 AI194277
```

```
                AI284633 AI316639 AI390324 AI458673 AI688647 AI694146
                AI694199 AI788240 AI802319 AI807413 AI831412 AI932363
                AI935446 AI982078 AJ396992 AJ398211 AL529211 AL563388.1
                AL601897.1 AU051650 AU125496 AU135776 AU137791 AU138464
                AV656408 AV686324 AV698092 AV703346 AV708342 AV708873
                AV709145 AV709526 AV739602 AV742411 AV743612 AV744587
                AV745018 AW050065 AW063931 AW270998 AW300107 AW300121
                AW301984 AW302318 AW302718 AW302841 AW302861 AW303029
                AW355281 AW822881 AW859606 AW959243 BB067326 BB123906
                BB609972 BB624305 BB628344 BB654156 BB823708 BE007675
                BE049029 BE049182 BE120619 BE178692 BE178693 BE244058
                BE246922 BE287246 BE289710 BE311050 BE369398 BE377181
                BE386264 BE465032 BE484573 BE554360 BE571717 BE625094
                BE766876 BE766886 BE770856 BE785982.1 BE813002 BE817780
                BE829341 BE829344 BE829355 BE829359 BE893564 BE895461
                BE896141 BE908143 BE913030 BE945891 BE991271 BE992722
                BF015474 BF021751 BF022025 BF029248 BF037603 BF060909
                BF064161 BF109576 BF111215 BF122334 BF137583 BF139381
                BF140373 BF145038 BF148549 BF150174 BF165619 BF207643
                BF326037 BF326041 BF326042 BF326045 BF334390 BF341149
                BF376614 BF378829 BF380539 BF380550 BF465841 BF535383
                BF578793 BF582307 BF748822 BF748938 BF748963 BF765346
                BF766403 BF784199 BF789335 BF789696 BF885392 BF896088
                BF940327 BF984391 BG030354 BG035327 BG088552 BG116557
                BG148023 BG167018 BG168325 BG247693 BG252842 BG295272
                BG295445 BG342500 BG610013 BG655426 BG667600 BG677384.1
                BG678705 BG692520 BG717389 BG801183 BG802707 BG803992
                BG863302 BG914583 BG917993 BG940215 BG943615 BG975946
                BI080618 BI092275 BI106992 BI107793 BI108344 BI222784
                BI258149 BI326830 BI651746 BI659175 BI666111 BI688353
                BI689042 BI694119 BI696541 BI696575 BI696638 BI733610
                BI753721 BI754669 BI851772 BI856689.1 BI872006 BI872428
                BI903238 BI903250 BI960823 BI986171 BM016286 BM069622
                BM069876 BM147734 BM237580 BM238719 BM238848 BM423047
                BM423107 F11344 F20959 H17162 H98057 N56057 R25494 T11219
                T51727 W47233
                match: cDNAs: AB046018.1 AJ420475.1 AK022681.1 AL136632.1
                BC002217.1 BC005537.1"
     gene       join(complement(AL133264.10:101. .211),5119. .5253,
                7768. .8194)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-002"
     mRNA       join(complement(AL133264.10:101. .211),5119. .5253,
                7768. .8194)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-002"
                /product="chromosome 6 open reading frame 62"
                /note="match: ESTs: BG772628 BG772628.1"
     gene       join(complement(AL133264.10:5931. .6069),
                complement(AL133264.10:2219. .2395),
                complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-003"
     mRNA       join(complement(AL133264.10:5931. .6069),
                complement(AL133264.10:2219. .2395),
                complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-003"
                /product="chromosome 6 open reading frame 62"
                /note="match: ESTs: BG342500.1 BG692520.1 BG803992.1"
     CDS        join(complement(AL133264.10:4611. .4739),
                complement(AL133264.10:2219. .2395),
                complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-003"
                /standard_name="OTTHUMP00000016093"
                /note="match: proteins: O95550 Q961D0 Q99K03 Q99LU8 Q9GZU0
                Q9H1Z2 Q9N0B2 Q9QEJ3 Q9W384 Q9XWI7"
                /codon_start=1
                /product="chromosome 6 open reading frame 62"
                /protein_id="CAI20135.1"
                /db_xref="GI:56202469"
                /translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
                KKSALFEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
                DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEFHHGDYEKQFLHV
                LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
                IEDHLRPYMPE"
     CDS        join(complement(AL133264.10:5931. .5972),
                complement(AL133264.10:2219. .2395),
                complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-003"
                /standard_name="OTTHUMP00000016094"
                /codon_start=1
                /product="chromosome 6 open reading frame 62"
                /protein_id="CAI20136.1"
                /db_xref="GI:56202470"
                /db_xref="UniProt/TrEMBL:Q5JVZ2"
                /translation="MSENLSDPVSPVVRKKKSALFEVSEVIPVMTNNYEENILKGVRD
                SSYSLESSLELLQKDVVQLHAPRYQSMRRDVIGCTQEMDFILWPRNDIEKIVCLLFSR
                WKESDEPFRPVQAKFEFHHGDYEKQFLHVLSRKDKTGIVVNNPNQSVFLFIDRQHLQT
                PKNKATIFKLCSICLYLPQEQLTHWAVGTIEDHLRPYMPE"
     polyA_signal   9124. .9129
                /gene="C6orf62"
                /locus_tag="RP1-30M3.4-001"
     gene       complement(join(12316. .12571,15963. .16147,46686. .46959))
                /gene="THEM2"
                /locus_tag="RP1-30M3.2-001"
     mRNA       complement(join(12316. .12571,15963. .16147,46686. .46959))
                /gene="THEM2"
                /locus_tag="RP1-30M3.2-001"
                /product="thioesterase superfamily member 2"
                /note="match: ESTs: AA614605 AA928745 AI017897 AI052396
                BF407527 BF475709 BF594161
                match: cDNAs: AF155649.1 AF220186.1 AF274952.1 AK000508.1
                AK002261.1 AK008624.1 BC000894.1"
     gene       complement(join(12316. .12571,15963. .16147,26337. .26628))
                /gene="THEM2"
                /locus_tag="RP1-30M3.2-002"
     mRNA       complement(join(12316. .12571,15963. .16147,26337. .26628))
                /gene="THEM2"
                /locus_tag="RP1-30M3.2-002"
                /product="thioesterase superfamily member 2"
                /note="match: ESTs: BG219890 BG256069"
     polyA_site 12318
     polyA_signal 12336. .12341
     CDS        complement(join(12415. .12571,15963. .16147,46686. .46766))
                /gene="THEM2"
                /locus_tag="RP1-30M3.2-001"
                /standard_name="OTTHUMP00000016090"
                /note="match: proteins: AAA28024 AAH18165 AAL48890 O95549
                P34419 P87304 P93828 P95914 Q18187 Q9ASQ1 Q9NPJ3 Q9VZZ5
                Q9VZZ6"
                /codon_start=1
                /product="thioesterase superfamily member 2"
                /protein_id="CAD92509.1"

Alignment Scores:
Pred. No.:              316             Length:         102200
Score:                  45.00           Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x HS30M3 (1-102200)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db      56134 CCAGATGTGATATTTCTACAGGAAGTT 56160

RESULT 27
AC152027/c
LOCUS       AC152027              150344 bp    DNA     linear   HTG 11-NOV-2004
DEFINITION  Dasypus novemcinctus clone VMRC5-12D21, WORKING DRAFT SEQUENCE, 5
            ordered pieces.
```

```
ACCESSION   AC152027
VERSION     AC152027.2  GI:55700150
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Dasypus novemcinctus (nine-banded armadillo)
  ORGANISM  Dasypus novemcinctus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Xenarthra; Dasypodidae; Dasypus.
REFERENCE   1  (bases 1 to 150344)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
            Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
            Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Laric,P.,
            Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L.,
            Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A.,
            Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,
            Shah,K., Sison,C., Stantripop,S., Stephen,E., Thomas,J.W.,
            Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
            Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625
            Fishers Lane, Rockville, MD 20852, USA
COMMENT     On Nov 11, 2004 this sequence version replaced gi:54312159.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: ifi
                 Center clone name: 012D21

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            -------------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 149480 bases at least Q40
                 Consensus quality: 149754 bases at least Q30
                 Consensus quality: 149909 bases at least Q20
                 Insert size: 155000; agarose-fp
                 Insert size: 149944; sum-of-contigs
                 Quality coverage: 9.52x in Q20 bases; agarose-fp
                 Quality coverage: 9.84x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    36273: contig of 36273 bp in length
            *    36274    36373: gap of unknown length
            *    36374    55049: contig of 18676 bp in length
            *    55050    55149: gap of unknown length
            *    55150   101939: contig of 46790 bp in length
            *   101940   102039: gap of unknown length
            *   102040   114806: contig of 12767 bp in length
            *   114807   114906: gap of unknown length
            *   114907   150344: contig of 35438 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150344
                     /organism="Dasypus novemcinctus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9361"
                     /clone="VMRC5-12D21"
                     /clone_lib="VMRC5"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .36273
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    1. .23436
                     /note="clone overlaps with GenBank Accession Number
                     AC152029 clone VMRC5-168B11 (center project name ifj)"
     gap             36274. .36373
                     /estimated_length=unknown
     misc_feature    36374. .55049
                     /note="assembly_fragment"
     gap             55050. .55149
                     /estimated_length=unknown
     misc_feature    55150. .101939
                     /note="assembly_fragment"
     gap             101940. .102039
                     /estimated_length=unknown
     misc_feature    102040. .114806
                     /note="assembly_fragment"
     gap             114807. .114906
                     /estimated_length=unknown
     misc_feature    114907. .150344
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:              474          Length:         150344
Score:                  45.00        Matches:        9
Percent Similarity:     100.0%       Conservative:   0
Best Local Similarity:  100.0%       Mismatches:     0
Query Match:            100.0%       Indels:         0
DB:                     14           Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AC152027 (1-150344)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db      53165 CCAGATGTGATATTTCTACAGGAAGTT 53139

RESULT 28
CR925830
LOCUS       CR925830              167847 bp    DNA     linear   PRI 08-JAN-2005
DEFINITION  Human DNA sequence from clone RP11-195J19 on chromosome 6, complete
            sequence.
ACCESSION   CR925830
VERSION     CR925830.3  GI:56798018
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 167847)
```

```
  AUTHORS   Palmer,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 24, 2004 this sequence version replaced gi:56542150.
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-195J19 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .167847
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-195J19"
                     /clone_lib="RPCI-11.1"
ORIGIN

Alignment Scores:
Pred. No.:                533           Length:         167847
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CR925830 (1-167847)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db      36377 CCAGATGTGATATTTCTACAGGAAGTT 36403

RESULT 29
CR942205
LOCUS       CR942205               176819 bp    DNA     linear   PRI 29-APR-2005
DEFINITION  Human DNA sequence from clone RP11-948M1 on chromosome 6, complete
            sequence.
ACCESSION   CR942205
VERSION     CR942205.4  GI:62989869
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 176819)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:62551221.
            ------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-948M1 is from the library RPCI-11.4 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .176819
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-948M1"
                     /clone_lib="RPCI-11.4"
ORIGIN

Alignment Scores:
Pred. No.:                563           Length:         176819
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CR942205 (1-176819)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db      30356 CCAGATGTGATATTTCTACAGGAAGTT 30382

RESULT 30
CR761829
LOCUS       CR761829               1236 bp    mRNA    linear   VRT 15-SEP-2004
DEFINITION  Xenopus tropicalis finished cDNA, clone TGas102g02.
ACCESSION   CR761829
VERSION     CR761829.1  GI:51967738
```

```
KEYWORDS    .
SOURCE      Xenopus tropicalis (Silurana tropicalis)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1236)
  AUTHORS   Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
            Francis,M.D., Garrett,N., Gilchrist,M.J., Grafham,D.V.,
            McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
            Voigt,J. and Zorn,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
COMMENT     Sanger Xenopus tropicalis EST/cDNA project.
            This sequence is from a Xenopus Gene Collection (XGC) library, from
            a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
            extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
            ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'
            end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1-blue.
FEATURES             Location/Qualifiers
     source          1. .1236
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TGas102g02"
                     /clone_lib="XGC-gastrula"
                     /dev_stage="gastrula (stage 10.5-13 mixed)"
ORIGIN

Alignment Scores:
Pred. No.:              5.18          Length:        1236
Score:                  44.00         Matches:       8
Percent Similarity:     100.0%        Conservative:  1
Best Local Similarity:  88.9%         Mismatches:    0
Query Match:            97.8%         Indels:        0
DB:                     5             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x CR761829 (1-1236)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||||
Db        514 CCAGATGTAGTATTCTTACAGGAGGTT 540

RESULT 31
BD205491
LOCUS       BD205491               1312 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  CD40-Interacting and TRAF-interacting protein.
ACCESSION   BD205491
VERSION     BD205491.1  GI:33015261
KEYWORDS    JP 2002512796-A/2.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting protein
  JOURNAL   Patent: JP 2002512796-A 2 08-MAY-2002;
            VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT     OS   Mus musculus (mouse)
            PN   JP 2002512796-A/2
            PD   08-MAY-2002
            PF   28-APR-1999 JP 2000546003
            PR   29-APR-1998 EP   98201392.2
            PI   STEFAN MARIA CHRISTIAAN PYPE,
            PI   JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI   DANNY
            FRANCOIS EVELINE HUYLEBROECK
            PC   C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
              A61P37/02,
            PC   A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
            PC   G01N33/566,
            PC   C12N15/00,A61K37/02
            CC   CD40-Interacting and TRAF-interacting protein. FH   Key
                 Location/Qualifiers
            FT   CDS             (122). .(1234).
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
ORIGIN

Alignment Scores:
Pred. No.:              5.52          Length:        1312
Score:                  44.00         Matches:       8
Percent Similarity:     100.0%        Conservative:  1
Best Local Similarity:  88.9%         Mismatches:    0
Query Match:            97.8%         Indels:        0
DB:                     6             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x BD205491 (1-1312)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||||
Db        584 CCAGATGTGGTATTTCTACAGGAAGTT 610

RESULT 32
AR594295
LOCUS       AR594295               1312 bp    DNA     linear   PAT 15-DEC-2004
DEFINITION  Sequence 3 from patent US 6812203.
ACCESSION   AR594295
VERSION     AR594295.1  GI:56643901
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting proteins
  JOURNAL   Patent: US 6812203-A 3 02-NOV-2004;
            Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
            Zwijnaarde;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              5.52          Length:        1312
Score:                  44.00         Matches:       8
Percent Similarity:     100.0%        Conservative:  1
Best Local Similarity:  88.9%         Mismatches:    0
Query Match:            97.8%         Indels:        0
DB:                     6             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x AR594295 (1-1312)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||||
Db        584 CCAGATGTGGTATTTCTACAGGAAGTT 610

RESULT 33
AX011601
LOCUS       AX011601               1312 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 3 from Patent WO9955859.
ACCESSION   AX011601
VERSION     AX011601.1  GI:9998125
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 3 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
     CDS             122. .1234
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07556.1"
                     /db_xref="GI:9998126"
                     /translation="MASGSSSDAAEPAGPAGRAASAPEAAQAEEDRVKRRRLQCLGFA
                     LVGGCDPTMVPSVLRENDWQTQKALSAYFELPENDQGWPRQPPTSFKSEAYVDLTNED
                     ANDTTILEASPSGTPLEDSSTISFITWNIDGLDGCNLPERARGVCSCLALYSPDVVFL
                     QEVIPPYCAYLKKRAASYTIITGNEEGYFTAILLKKGRVKFKSQEIIPFPNTKMMRNL
                     LCVNVSLGGNEFCLMTSHLESTREHSAERIRQLKTVLGKMQEAPDSTTVIFAGDTNLR
                     DQEVIKCGGLPDNVFDAWEFLGKPKHCQYTWDTKANNNLRIPAAYKHRFDRIFFRAEE
                     GHLIPQSLDLVGLEKLDCGRFPSDHWGLLCTLNVVL"
ORIGIN

Alignment Scores:
Pred. No.:                5.52          Length:          1312
Score:                    44.00         Matches:         8
Percent Similarity:       100.0%        Conservative:    1
Best Local Similarity:    88.9%         Mismatches:      0
Query Match:              97.8%         Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AX011601 (1-1312)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||
Db        584 CCAGATGTGGTATTTCTACAGGAAGTT 610

RESULT 34
MMU251328
LOCUS       MMU251328               1943 bp    mRNA    linear   ROD 27-MAY-2000
DEFINITION  Mus musculus mRNA for putative TRAF and TNF receptor associated
            protein (TTRAP gene).
ACCESSION   AJ251328
VERSION     AJ251328.1  GI:8247282
KEYWORDS    ttrap gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van
            Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,
            Huylebroeck,D. and Remacle,J.E.
  TITLE     TTRAP, a novel protein that associates with CD40, tumor necrosis
            factor (TNF) receptor-75 and TNF receptor-associated factors
            (TRAFs), and that inhibits nuclear factor-kappa B activation
  JOURNAL   J. Biol. Chem. 275 (24), 18586-18593 (2000)
   PUBMED   10764746
REFERENCE   2  (bases 1 to 1943)
  AUTHORS   Pype,S.M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-NOV-1999) Pype S.M.C., VIB07, Department Cell Growth,
            Differentiation and Developme, Flanders Interuniversity Institute
            for Biotechnology, K.U.Leuven, Campus Gasthuisberg, Herestraat 49,
            3000 Leuven, BELGIUM
COMMENT     Related sequence: AJ269473.
```

```
FEATURES             Location/Qualifiers
     source          1. .1943
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /note="IMAGE cDNA EST clone 1262914"
     gene            1. .1943
                     /gene="TTRAP"
     CDS             48. .1160
                     /gene="TTRAP"
                     /codon_start=1
                     /product="putative TRAF and TNF receptor associated
                     protein"
                     /protein_id="CAB92971.1"
                     /db_xref="GI:8247283"
                     /db_xref="GOA:Q9JJX7"
                     /db_xref="UniProt/TrEMBL:Q9JJX7"
                     /translation="MASGSSSDAAEPAGPAGRAASAPEAAQAEEDRVKRRRLQCLGFA
                     LVGGCDPTMVPSVLRENDWQTQKALSAYFELPENDQGWPRQPPTSFKSEAYVDLTNED
                     ANDTTILEASPSGTPLEDSSTISFITWNIDGLDGCNLPERARGVCSCLALYSPDVVFL
                     QEVIPPYCAYLKKRAASYTIITGNEEGYFTAILLKKGRVKFKSQEIIPFPNTKMMRNL
                     LCVNVSLGGNEFCLMTSHLESTREHSAERIRQLKTVLGKMQEAPDSTTVIFAGDTNLR
                     DQEVIKCGGLPDNVFDAWEFLGKPKHCQYTWDTKANNNLRIPAAYKHRFDRIFFRAEE
                     GHLIPQSLDLVGLEKLDCGRFPSDHWGLLCTLNVVL"
ORIGIN

Alignment Scores:
Pred. No.:                8.33          Length:          1943
Score:                    44.00         Matches:         8
Percent Similarity:       100.0%        Conservative:    1
Best Local Similarity:    88.9%         Mismatches:      0
Query Match:              97.8%         Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x MMU251328 (1-1943)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||
Db        510 CCAGATGTGGTATTTCTACAGGAAGTT 536

RESULT 35
BA000019_50
WPCOMMENT
  Sequence split into 65 fragments  LOCUS BA000019 Accession BA000019
     Fragment Name      Begin      End
     BA000019_00            1    110000
     BA000019_01       100001    210000
     BA000019_02       200001    310000
     BA000019_03       300001    410000
     BA000019_04       400001    510000
     BA000019_05       500001    610000
     BA000019_06       600001    710000
     BA000019_07       700001    810000
     BA000019_08       800001    910000
     BA000019_09       900001   1010000
     BA000019_10      1000001   1110000
     BA000019_11      1100001   1210000
     BA000019_12      1200001   1310000
     BA000019_13      1300001   1410000
     BA000019_14      1400001   1510000
     BA000019_15      1500001   1610000
     BA000019_16      1600001   1710000
     BA000019_17      1700001   1810000
     BA000019_18      1800001   1910000
     BA000019_19      1900001   2010000
     BA000019_20      2000001   2110000
     BA000019_21      2100001   2210000
     BA000019_22      2200001   2310000
     BA000019_23      2300001   2410000
     BA000019_24      2400001   2510000
     BA000019_25      2500001   2610000
     BA000019_26      2600001   2710000
     BA000019_27      2700001   2810000
```

```
    BA000019_28       2800001   2910000
    BA000019_29       2900001   3010000
    BA000019_30       3000001   3110000
    BA000019_31       3100001   3210000
    BA000019_32       3200001   3310000
    BA000019_33       3300001   3410000
    BA000019_34       3400001   3510000
    BA000019_35       3500001   3610000
    BA000019_36       3600001   3710000
    BA000019_37       3700001   3810000
    BA000019_38       3800001   3910000
    BA000019_39       3900001   4010000
    BA000019_40       4000001   4110000
    BA000019_41       4100001   4210000
    BA000019_42       4200001   4310000
    BA000019_43       4300001   4410000
    BA000019_44       4400001   4510000
    BA000019_45       4500001   4610000
    BA000019_46       4600001   4710000
    BA000019_47       4700001   4810000
    BA000019_48       4800001   4910000
    BA000019_49       4900001   5010000
    BA000019_50       5000001   5110000
    BA000019_51       5100001   5210000
    BA000019_52       5200001   5310000
    BA000019_53       5300001   5410000
    BA000019_54       5400001   5510000
    BA000019_55       5500001   5610000
    BA000019_56       5600001   5710000
    BA000019_57       5700001   5810000
    BA000019_58       5800001   5910000
    BA000019_59       5900001   6010000
    BA000019_60       6000001   6110000
    BA000019_61       6100001   6210000
    BA000019_62       6200001   6310000
    BA000019_63       6300001   6410000
    BA000019_64       6400001   6413771
Continuation (51 of 65) of BA000019 from base 5000001 (BA000019 Nostoc sp. PCC 7120 DNA,

Alignment Scores:
Pred. No.:                578             Length:         110000
Score:                    44.00           Matches:        8
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    88.9%           Mismatches:     0
Query Match:              97.8%           Indels:         0
DB:                       1               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BA000019_50 (1-110000)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||:::|||||||||||||||||||
Db      58702 CCAGATATCATATTTCTACAAGAAGTC 58728

RESULT 36
AC125756/c
LOCUS       AC125756               245394 bp    DNA     linear   HTG 13-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-42H3, WORKING DRAFT SEQUENCE, 2
            unordered pieces.
ACCESSION   AC125756
VERSION     AC125756.5  GI:30578987
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 245394)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 245394)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 245394)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 13, 2003 this sequence version replaced gi:23265926.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
```

```
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
          ------------- Project Information
            Center project name: GJVU
            Center clone name: CH230-42H3
          ------------- Summary Statistics
            Assembly program: Atlas 3.0;
            Consensus quality: 238380 bases at least Q40
            Consensus quality: 239660 bases at least Q30
            Consensus quality: 240760 bases at least Q20
            Estimated insert size: 255581; sum-of-contigs estimation
            Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
          -------------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
          * NOTE: This sequence may represent more than one clone.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 2 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *       1   242009: contig of 242009 bp in length
          *  242010   242109: gap of unknown length
          *  242110   245394: contig of 3285 bp in length.
FEATURES             Location/Qualifiers
     source          1. .245394
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-42H3"
     misc_feature    1. .1025
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    1076. .2080
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    3723. .4536
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH348771"
     misc_feature    128656. .128715
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH260088"
     misc_feature    236171. .238412
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    239605. .242009
                     /note="wgs_end_extension
                     clone_end:Sp6"
     gap             242010. .242109
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:              1.34e+03        Length:         245394
Score:                  44.00           Matches:        8
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  88.9%           Mismatches:     0
Query Match:            97.8%           Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AC125756 (1-245394)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||
Db     144885 CCGGACGTGGTATTTCTACAGGAAGTG 144859
```

```
RESULT 37
AL589699
LOCUS       AL589699              256608 bp    DNA     linear   ROD 17-DEC-2004
DEFINITION  Mouse DNA sequence from clone RP23-92G13 on chromosome 13 Contains
            the 3' end of the gene for a novel protein similar to KIAA0386,
            five novel genes, the Gmnn gene for geminin, gene PNAS-27, the
            Ttrap gene for Traf and Tnf receptor associated protein, the gene
            for a novel protein similar to KIAA0319, the gene for the ortholog
            of human and rat aldehyde dehydrogenase 5A1 (succinate-semialdehyde
            dehydrogenase) (ALDH5A1) (SSADH), the 5' end of the Gpld1 gene for
            glycosylphosphatidylinositol specific phospholipase D1 and four CpG
            islands, complete sequence.
ACCESSION   AL589699
VERSION     AL589699.4  GI:13810107
KEYWORDS    HTG; Aldh5a1; CpG island; geminin; Gmnn; Gpld1; KIAA0319; KIAA0386;
            phospholipase; SSADH; Ttrap.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 256608)
  AUTHORS   Pelan,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 26, 2001 this sequence version replaced gi:13751568.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep ----------
            Genome Center
            Center: UK Medical Research Council
            Center code: UK-MRC
            Web site: http://mrcseq.har.mrc.ac.uk
            Contact: mouseq@har.mrc.ac.uk
            ----------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence.  Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
            RP23-92G13 is from the RPCI-23 Mouse BAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .256608
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="13"
                     /clone="RP23-92G13"
                     /clone_lib="RPCI-23"
     gene            join(AL513014.9:11808. .79376,4770. .28457)
                     /locus_tag="RP23-461O24.1-003"
     mRNA            join(AL513014.9:11808. .12037,AL513014.9:43061. .43187,
                     AL513014.9:44856. .45011,AL513014.9:46133. .46211,
                     AL513014.9:47748. .47771,AL513014.9:48876. .48929,
                     AL513014.9:50825. .50974,AL513014.9:53729. .53792,
                     AL513014.9:62567. .62610,AL513014.9:66867. .66992,
                     AL513014.9:67218. .67366,AL513014.9:68674. .68803,
                     AL513014.9:74163. .74831,AL513014.9:79192. .79376,
                     4770. .4932,8263. .8398,9931. .10092,12188. .12346,
                     16272. .16474,18350. .18433,19587. .19677,26226. .28457)
```

```
                     /locus_tag="RP23-461O24.1-003"
                     /note="match: cDNAs: AK007145.1 AB002384.1
                     match: ESTs: AW822704.1 BB674120.1 AA041998.1 BM144846.1
                     Z42557.1 R20596.1"
     CDS             join(AL513014.9:11989. .12037,AL513014.9:43061. .43187,
                     AL513014.9:44856. .45011,AL513014.9:46133. .46211,
                     AL513014.9:47748. .47771,AL513014.9:48876. .48929,
                     AL513014.9:50825. .50974,AL513014.9:53729. .53792,
                     AL513014.9:62567. .62610,AL513014.9:66867. .66992,
                     AL513014.9:67218. .67366,AL513014.9:68674. .68803,
                     AL513014.9:74163. .74831,AL513014.9:79192. .79376,
                     4770. .4932,8263. .8398,9931. .10092,12188. .12346,
                     16272. .16474,18350. .18433,19587. .19677,26226. .26326)
                     /locus_tag="RP23-461O24.1-003"
                     /standard_name="OTTMUSP00000000566"
                     /note="match: proteins: Q9D9C2 Q9Y4F9"
                     /codon_start=1
                     /protein_id="CAI26080.1"
                     /db_xref="GI:56237839"
                     /translation="MPPGTKRLRALGAFSAGLPTRLPEIMLVGSQSFSPGGPNGIIRS
                     QSFAGFSGLQERRSRCNSFIENASALKKPQAKLKKMHNLGHKNNNTPKEPQPKRVEEV
                     YRALKNGLDEYLEFHQTELDKLTAQLKDMKRNSRLGVLYDLDKQIKTIERYMRRLEFH
                     ISKVDELYEAYCIQRRLQDGASKMKQAFATSPASKAARESLSEINRSYKEYTENMCAI
                     EAELESLLGEFSIKMKGLAGFARLCPGDQYEIFMKYGRQRWKLKGKIEVNGKQSWDGA
                     ETVFLPLIVGFISIKVTELKGLATHILVGSVTCETKELFAARPQVVAVDINDLGTIKL
                     NLEITWYPFDVEDTTPSSGPGNKTAALQRRMSMYSQGTPETPTFKDQSFFSNLPDDIF
                     ESGKAAEEKRPLSLSFSDLQDGDCVFTSSSATSPSSSHSAHPEITITPAELTHSSLSS
                     QNEGTEDSSSASSRNSLGEDHEPKSHPKSDTVEPGKPGVATRSGTESLFLESSVAEAL
                     LQESDEASELKPVELDTFEGNITKQLVKRLTSAEGPVTTDKLFFEGSVGSESEAGRSF
                     LDGSLEDAFNGLFLALDPHKKQYKEFQDLNQEVTHLDDVLKCKPAGSRSRSSSLSLTV
                     ESALESFDFLNTSDFDEEEEDGDDVCHVGGGADSVFSDTETEKSGSVHPEARGHLSEA
                     LTEDTGVGTSVAGSPLPLTTGNESLDITIVKHLQYCTQLIQQIVFSSKTPFVARSLLE
                     KLSRQVLVMQKLAAVSDENLGNITSVVEAIPEFHKKLSLLAFWTKCCSPSGVYHSSAA
                     RLIKQLEASFARSINKDYPGLAEPVFRTLVSQILDRAEPLLSSSLSSEVITVFQYYSF
                     FTSHGVSDLETYLGQLTRQVAMVQTLQSLRDEKLLQTMSDLAPSNLPAQQEVLRTLAL
                     LLTRDDNEVSEAVTLYLAAASKNEHFREKALLYYCEALTKANLQLQKAACLALKSLEA
                     TESIKMLVTLCQSDTEEIRTVASETLLSLGEDGRLAYEQLDKFPRDCVKVGGRHGTEV
                     ATAF"
     polyA_signal    26424. .26429
                     /locus_tag="RP23-461O24.1-003"
     polyA_site      26446
                     /locus_tag="RP23-461O24.1-003"
     polyA_site      26447
                     /locus_tag="RP23-461O24.1-003"
     polyA_site      26457
                     /locus_tag="RP23-461O24.1-003"
     gene            join(35873. .36201,40700. .41233)
                     /locus_tag="RP23-92G13.3-001"
     mRNA            join(35873. .36201,40700. .41233)
                     /locus_tag="RP23-92G13.3-001"
                     /note="match: cDNAs: AK006029.1
                     match: ESTs: AV045211 BF018417 BB012811 BB014319 BB020000
                     AV263251 BF319523 AV206925 AV041162 AV259155 AV041846"
     polyA_signal    39211. .39216
                     /locus_tag="RP23-92G13.3-001"
     polyA_site      39233
                     /locus_tag="RP23-92G13.3-001"
     polyA_site      complement(44486)
     polyA_site      complement(44487)
     polyA_site      complement(44488)
     polyA_signal    complement(44511. .44516)
     polyA_signal    complement(44515. .44520)
     gene            complement(join(46486. .46871,47989. .48099,48299. .48381,
                     51241. .51376,52237. .52314,54801. .54858,56454. .56564))
                     /gene="Gmnn"
                     /locus_tag="RP23-92G13.2-001"
     mRNA            complement(join(46486. .46871,47989. .48099,48299. .48381,
                     51241. .51376,52237. .52314,54801. .54858,56454. .56564))
                     /gene="Gmnn"
                     /locus_tag="RP23-92G13.2-001"
                     /note="match: ESTs: BI789536.1 BQ178448.1 BF722692.1
                     AA510015.1 BG797776.1 BB868504.1 AA014584.1 AA032471.1
                     BI454155.1 BB448331.1 W13399.1 BF556131.1 BG085437.1
                     AI528457.1 AA277821.1 AA681520.1 W99871.1 AA004051.1"
     gene            complement(join(46486. .46871,47989. .48099,48299. .48381,
                     51241. .51376,52237. .52314,54801. .54858,56228. .56502))
                     /gene="Gmnn"
                     /locus_tag="RP23-92G13.2-002"
     mRNA            complement(join(46486. .46871,47989. .48099,48299. .48381,
                     51241. .51376,52237. .52314,54801. .54858,56228. .56502))
                     /gene="Gmnn"
                     /locus_tag="RP23-92G13.2-002"
                     /note="match: cDNAs: AF068780.1 BC020061.1
                     match: ESTs: BF160648.1 AA437903.1"
     CDS             complement(join(46710. .46871,47989. .48099,48299. .48381,
                     51241. .51376,52237. .52314,54801. .54851))
                     /gene="Gmnn"
                     /locus_tag="RP23-92G13.2-001"
                     /standard_name="OTTMUSP00000000569"
                     /note="match: proteins: O88513 O75496"
                     /codon_start=1
                     /protein_id="CAI26081.1"
                     /db_xref="GI:56237840"
                     /translation="MNLSMKQKQEGAQENVKNSPVPRRTLKMIQPSADGSLVGRENEL
                     PKGLFKRKLWDDQLASQTSSCGPEANENKDVGDLTQEAFDLISKENPSSQYWKEVAEQ
                     RRKALYEALKENEKLHKEIEQKDSEIARLRKENKDLAEVAEHVQYMAEVIERLSNEPL
                     DNFESPDSQEFDSEEEAVEYSELEDSGAGTCAEETVSSSTDARPCT"
     gene            96664. .108654
                     /gene="RP23-92G13.4"
                     /locus_tag="RP23-92G13.4-001"
     mRNA            join(96664. .98150,99423. .99599,100503. .100625,
                     104527. .104661,107278. .108654)
                     /gene="RP23-92G13.4"
                     /locus_tag="RP23-92G13.4-001"
                     /note="match: cDNAs: BC002217.1 BC005537.1 AL136632.1
                     AK022681.1 AJ420475.1
                     match: ESTs: BI078093.1 BG295445.1 BI696541.1 BG717389.1
                     BI108344.1 BG975946.1 BI452750.1 BG168325.1 BG295139.1
                     BG914583.1 BI753721.1"
     CDS             join(98022. .98150,99423. .99599,100503. .100625,
                     104527. .104661,107278. .107403)
                     /gene="RP23-92G13.4"
                     /locus_tag="RP23-92G13.4-001"
                     /standard_name="OTTMUSP00000000563"
                     /note="match: proteins: Q9H1Z2 Q99LU8 Q9GZU0 O95550"
                     /codon_start=1
                     /protein_id="CAI26082.1"
                     /db_xref="GI:56237841"
                     /translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
                     KKSALPEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
                     DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEFHHGDYEKQFLHV
                     LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
                     IEDHLRPYMPE"
     polyA_signal    106607. .106612
                     /gene="RP23-92G13.4"
                     /locus_tag="RP23-92G13.4-001"
     polyA_site      106654
                     /gene="RP23-92G13.4"
                     /locus_tag="RP23-92G13.4-001"

Alignment Scores:
Pred. No.:              1.41e+03        Length:         256608
Score:                  44.00           Matches:        8
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  88.9%           Mismatches:     0
Query Match:            97.8%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AL589699 (1-256608)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||
Db     131503 CCAGATGTGGTATTTCTACAGGAAGTT 131529

RESULT 38
BA000045_42/c
```

```
WPCOMMENT
  Sequence split into 47 fragments  LOCUS BA000045 Accession BA000045
     Fragment Name      Begin     End
     BA000045_00            1    110000
     BA000045_01       100001    210000
     BA000045_02       200001    310000
     BA000045_03       300001    410000
     BA000045_04       400001    510000
     BA000045_05       500001    610000
     BA000045_06       600001    710000
     BA000045_07       700001    810000
     BA000045_08       800001    910000
     BA000045_09       900001   1010000
     BA000045_10      1000001   1110000
     BA000045_11      1100001   1210000
     BA000045_12      1200001   1310000
     BA000045_13      1300001   1410000
     BA000045_14      1400001   1510000
     BA000045_15      1500001   1610000
     BA000045_16      1600001   1710000
     BA000045_17      1700001   1810000
     BA000045_18      1800001   1910000
     BA000045_19      1900001   2010000
     BA000045_20      2000001   2110000
     BA000045_21      2100001   2210000
     BA000045_22      2200001   2310000
     BA000045_23      2300001   2410000
     BA000045_24      2400001   2510000
     BA000045_25      2500001   2610000
     BA000045_26      2600001   2710000
     BA000045_27      2700001   2810000
     BA000045_28      2800001   2910000
     BA000045_29      2900001   3010000
     BA000045_30      3000001   3110000
     BA000045_31      3100001   3210000
     BA000045_32      3200001   3310000
     BA000045_33      3300001   3410000
     BA000045_34      3400001   3510000
     BA000045_35      3500001   3610000
     BA000045_36      3600001   3710000
     BA000045_37      3700001   3810000
     BA000045_38      3800001   3910000
     BA000045_39      3900001   4010000
     BA000045_40      4000001   4110000
     BA000045_41      4100001   4210000
     BA000045_42      4200001   4310000
     BA000045_43      4300001   4410000
     BA000045_44      4400001   4510000
     BA000045_45      4500001   4610000
     BA000045_46      4600001   4659019
Continuation (43 of 47) of BA000045 from base 4200001 (BA000045 Gloeobacter violaceus PC

Alignment Scores:
Pred. No.:              978           Length:         110000
Score:                  43.00         Matches:        7
Percent Similarity:     100.0%        Conservative:   2
Best Local Similarity:  77.8%         Mismatches:     0
Query Match:            95.6%         Indels:         0
DB:                     1             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BA000045_42 (1-110000)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||::::::|||||||||||||||||
Db      76087 CCGGATATTGTGTTTTTGCAGGAGGTG 76061

RESULT 39
CP000089_40
WPCOMMENT
  Sequence split into 45 fragments  LOCUS CP000089 Accession CP000089
     Fragment Name      Begin     End
     CP000089_00            1    110000
     CP000089_01       100001    210000
     CP000089_02       200001    310000
     CP000089_03       300001    410000
     CP000089_04       400001    510000
     CP000089_05       500001    610000
     CP000089_06       600001    710000
     CP000089_07       700001    810000
     CP000089_08       800001    910000
     CP000089_09       900001   1010000
     CP000089_10      1000001   1110000
     CP000089_11      1100001   1210000
     CP000089_12      1200001   1310000
     CP000089_13      1300001   1410000
     CP000089_14      1400001   1510000
     CP000089_15      1500001   1610000
     CP000089_16      1600001   1710000
     CP000089_17      1700001   1810000
     CP000089_18      1800001   1910000
     CP000089_19      1900001   2010000
     CP000089_20      2000001   2110000
     CP000089_21      2100001   2210000
     CP000089_22      2200001   2310000
     CP000089_23      2300001   2410000
     CP000089_24      2400001   2510000
     CP000089_25      2500001   2610000
     CP000089_26      2600001   2710000
     CP000089_27      2700001   2810000
     CP000089_28      2800001   2910000
     CP000089_29      2900001   3010000
     CP000089_30      3000001   3110000
     CP000089_31      3100001   3210000
     CP000089_32      3200001   3310000
     CP000089_33      3300001   3410000
     CP000089_34      3400001   3510000
     CP000089_35      3500001   3610000
     CP000089_36      3600001   3710000
     CP000089_37      3700001   3810000
     CP000089_38      3800001   3910000
     CP000089_39      3900001   4010000
     CP000089_40      4000001   4110000
     CP000089_41      4100001   4210000
     CP000089_42      4200001   4310000
     CP000089_43      4300001   4410000
     CP000089_44      4400001   4501104
Continuation (41 of 45) of CP000089 from base 4000001 (CP000089 Dechloromonas aromatica

Alignment Scores:
Pred. No.:              978           Length:         110000
Score:                  43.00         Matches:        7
Percent Similarity:     100.0%        Conservative:   2
Best Local Similarity:  77.8%         Mismatches:     0
Query Match:            95.6%         Indels:         0
DB:                     1             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CP000089_40 (1-110000)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||::::::|||||||||||||||||
Db      52636 CCGGACATCGTCTTTCTGCAGGAGGTG 52662

RESULT 40
CEY63D3A/c
LOCUS       CEY63D3A               33899 bp    DNA     linear   INV 09-AUG-2005
DEFINITION  Caenorhabditis elegans YAC Y63D3A, complete sequence.
ACCESSION   AL032652
VERSION     AL032652.1  GI:3810723
KEYWORDS    HTG.
SOURCE      Caenorhabditis elegans
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 33899)
  AUTHORS   .
  CONSRTM   C. elegans Sequencing Consortium
```

```
  TITLE     Genome sequence of the nematode C. elegans: a platform for
            investigating biology
  JOURNAL   Science 282 (5396), 2012-2018 (1998)
   PUBMED   9851916
REFERENCE   2  (bases 1 to 33899)
  AUTHORS   White,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-OCT-1998) Nematode Sequencing Project, Sanger
            Institute, Hinxton, Cambridge CB10 1SA, England and Department of
            Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
            worm@sanger.ac.uk
COMMENT     On May 14, 2001 this sequence version replaced gi:3646980.
            Coding sequences below are predicted from computer analysis, using
            predictions from Genefinder (P. Green, U. Washington), and other
            available information.
            Current sequence finishing criteria for the C. elegans genome
            sequencing consortium are that all bases are either sequenced
            unambiguously on both strands, or on a single strand with both a
            dye primer and dye terminator reaction, from distinct subclones.
            Exceptions are indicated by an explicit note.
            IMPORTANT:  This sequence is NOT necessarily the entire insert of
            the specified clone.  It may be shorter because we only sequence
            overlapping sections once, or longer because we arrange for a small
            overlap between neighbouring submissions.
            For a graphical representation of this sequence and its analysis
            see:- http://www.wormbase.org/perl/ace/elegans/seq/sequence?
            name=Y63D3A;class=Sequence
            IMPORTANT: This sequence is not the entire insert of clone Y63D3A.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we arrange for a small overlap between
            neighbouring submissions.
            The true left end of clone K03D10 is at 33796 in this sequence. The
            true right end of clone F47G4 is at 103 in this sequence. The true
            right end of clone Y71A12 is at 113 in this sequence. The start of
            this sequence (1. .103) overlaps with the end of sequence Z99171.
            The end of this sequence (33796. .33899) overlaps with the start of
            sequence Z81561.
FEATURES             Location/Qualifiers
     source          1. .33899
                     /organism="Caenorhabditis elegans"
                     /mol_type="genomic DNA"
                     /strain="Bristol N2"
                     /db_xref="taxon:6239"
                     /chromosome="I"
                     /clone="Y63D3A"
     gene            complement(join(49. .345,612. .1065,1117. .1199))
                     /locus_tag="Y63D3A.9"
     CDS             complement(join(49. .345,612. .1065,1117. .1199))
                     /locus_tag="Y63D3A.9"
                     /standard_name="Y63D3A.9"
                     /note="contains similarity to Pfam domain PF00646 (F-box
                     domain)"
                     /codon_start=1
                     /product="Hypothetical protein Y63D3A.9"
                     /protein_id="CAA21711.1"
                     /db_xref="GI:3881237"
                     /db_xref="InterPro:IPR001810"
                     /db_xref="UniProt/TrEMBL:Q9XWG0"
                     /translation="MSSPSFPILRLPSKALRTTLQNFDFIDLLSFSFVSPTPTKSLNM
                     KVTDFIIIKNYGIYIVARCCGIALRPITSQIFVDGQPQERQWEIQSITVKKCFENFLA
                     AFNHFEVEYICVRGEDDISYNLEEIKVVKNLEIFSADCGIFGQLYRVANKVTLSQRSL
                     LHEKPRDFFTENMDSLQLEKAVIYRPFTLSDLNLFLQHWRKGSNPRLEEIYIGLKEDN
                     LVDFKKTLLKGIEYWKYGIFRVYNNSPSIEGLEIERTDGIRAIIFVNPMHRSIGMGVL
                     Q"
     gene            join(1524. .1606,1651. .2086,4038. .4352)
                     /locus_tag="Y63D3A.10"
     CDS             join(1524. .1606,1651. .2086,4038. .4352)
                     /locus_tag="Y63D3A.10"
                     /standard_name="Y63D3A.10"
                     /note="contains similarity to Pfam domain PF00646 (F-box
                     domain)"
                     /codon_start=1
                     /product="Hypothetical protein Y63D3A.10"
                     /protein_id="CAA21706.1"
                     /db_xref="GI:3881232"
                     /db_xref="InterPro:IPR001810"
                     /db_xref="UniProt/TrEMBL:Q9XWG4"
                     /translation="MAHVAFNLLHLPAKTLRNVFQLMSHSELFIFSFASICAKRLVKS
                     MNFEMLELDLVYSRPCKLKISFGEHEEFIIEFDEFGTLSKQEPSDAFVFLIAYKLDIR
                     DCINHVLEVYKISKLHTLKVIGAINNPQNLLKIIGKLDLKVLWIFFNFSENSTTISVY
                     REILNHLGARCDVRSLRRGFYDKEPQKSFFLLRLSLSYALSQEPHSRPTATLAARRLK
                     RRQDTQPLDATTSVLMLRSCRTVGRGTSLPTHPPFSFVIAPHYYSCTCSDIQGEPDKL
                     Q"
     gene            complement(join(2689. .3289,3365. .3484,3604. .3686))
                     /locus_tag="Y63D3A.11"
     CDS             complement(join(2689. .3289,3365. .3484,3604. .3686))
                     /locus_tag="Y63D3A.11"
                     /standard_name="Y63D3A.11"
                     /note="contains similarity to Pfam domain PF00646 (F-box
                     domain)"
                     /codon_start=1
                     /product="Hypothetical protein Y63D3A.11"
                     /protein_id="CAB76748.1"
                     /db_xref="GI:7160769"
                     /db_xref="InterPro:IPR001810"
                     /db_xref="UniProt/TrEMBL:Q9NA68"
                     /translation="MSSPSLPILRLPSKALRTTLQNFEYLDLLSFSLVSPTPVKSLNM
                     KVTDFILISNDEIVIFGRATMFFRKSFENFLIAFNHFEVEHVYVNIENDIPYNLEEIK
                     VVKDLFVRSADRGIFSQFYRTANKISLSNRSLHNDKSRNFFTENIDSLWLDSSFHLEF
                     NDLLAINAKKAVIYRPFTLSDLNLFLQHWRKGSNPRLEEIYIGLEEDNLVDFNETLLK
                     GIVHWKSCLSKVYNNSPSIEGLEIERTDGTRAIIFVNPMKRSFEMGVLQ"
     gene            join(5095. .5177,5974. .6859)
                     /locus_tag="Y63D3A.3"
     CDS             join(5095. .5177,5974. .6859)
                     /locus_tag="Y63D3A.3"
                     /standard_name="Y63D3A.3"
                     /note="contains similarity to Pfam domain PF00646 (F-box
                     domain)"
                     /codon_start=1
                     /product="Hypothetical protein Y63D3A.3"
                     /protein_id="CAB63397.2"
                     /db_xref="GI:7160770"
                     /db_xref="InterPro:IPR001810"
                     /db_xref="UniProt/TrEMBL:Q9NA67"
                     /translation="MPNYPFPILWLPPKPLGHLLRNMDFRNLIIFSLTSKPAKWLTEL
                     LNMDAPMISVHVDHQINIIVSSVKSVTYHASNGSLRESRGKIDRQINFPGSLKQFVEH
                     LLHIFKEHKIDDLEIMCNSVYQYKPDIQALKGIKMRTLKLRYQESILKMFGTSADRIE
                     FSTIKSAELENSKDILIQNFDYLKFDFEGSLTADNLIGINGKVVNIWAADIEMKDVNI
                     FLKHWLKGSNPRLEKLILKTYRLDTDGNSVEDSVADLFRDLDCESIAESQFEILRDDE
                     SIAKISILEHGYWYNKIHCYFSLNEAMNLYTTYFWDIVHLQEKILH"
     gene            complement(join(7747. .8617,13260. .13342))
                     /locus_tag="Y63D3A.2"
     CDS             complement(join(7747. .8617,13260. .13342))
                     /locus_tag="Y63D3A.2"
                     /standard_name="Y63D3A.2"
                     /note="contains similarity to Pfam domain PF00646 (F-box
                     domain)"
                     /codon_start=1
                     /product="Hypothetical protein Y63D3A.2"
                     /protein_id="CAA21709.1"
                     /db_xref="GI:3881235"
                     /db_xref="InterPro:IPR001810"
                     /db_xref="UniProt/TrEMBL:Q9XWG1"
                     /translation="MPALPFPILSLPPKSRRIALKCMDTLDLINISFTHPLAKRLAIE
                     LSVGIEVGSLRLLVRNGFKVFFSFYAPPHNIIADISYLSADGETFDLIKDEDVLWINS
                     KIKLRDLLVDLLDVLHLRKLGEVSVHTSWDLILDELRGIEIHQLIVGESITRQFKLLK
                     LQAEELELNVPLTQADVICHNLHRIALKSPGNIQCEDLIGMNMMHIEVFMMFDHSSLT
                     HREINLFLLHWVKGSNPRMRYYRSPLQEPLQFEEQFVAQILKGIPHRAITPGELEKLR
                     VQFRGKTFQFRGYHIERNEGDAKAMISVDPWSPIVEVFVWF"
     gene            complement(join(14704. .14833,15019. .15302,16493. .16724,
                     16858. .16952,17625. .17845,17895. .17969,18033. .18084))
                     /locus_tag="Y63D3A.4"
     CDS             complement(join(14704. .14833,15019. .15302,16493. .16724,
                     16858. .16952,17625. .17845,17895. .17969,18033. .18084))
                     /locus_tag="Y63D3A.4"
                     /standard_name="Y63D3A.4"
```

```
                     /note="contains similarity to Pfam domain PF03372
                     (Endonuclease/Exonuclease/phosphatase family)"
                     /codon_start=1
                     /product="Hypothetical protein Y63D3A.4"
                     /protein_id="CAA21707.1"
                     /db_xref="GI:3881233"
                     /db_xref="InterPro:IPR005135"
                     /db_xref="UniProt/TrEMBL:Q9XWG3"
                     /translation="MSNSDDEIQEIEAKRQKMSQEDSEVEIEILDEPEQGKLKNSSMS
                     DEQKLHEFAIITATDEAFAQSILQDVDWDLKKALDVFYGSEAFAEARSAAVMGASSSM
                     ASSGAAVMTAEDLKGFEVSVMSWNIDGLDGRSLLTRMKAVAHIVKNVNPDILFLQEVV
                     DRDLAPIDKLQSLYKIYYSNKGCQYYTAILVSKMFDVEKHDVIHFQNSGMYRTLQILE
                     GSIGGLKVFLLNTHLESTREHRPQRCAQFGFCMDKVREIIAQNPGALVFFGGDLNLRD
                     EEVSRVPDGVKDAWEAAGSDNKTKFTWDTFKNDNKQGFHGAKMRFDRLYWSGPLDKVK
                     FTLEGRQRIRSCLCFPSDHWAINATFFA"
     gene            join(19048. .19228,19281. .19426,20077. .21210)
                     /gene="tfg-1"
                     /locus_tag="Y63D3A.5"
     CDS             join(19048. .19228,19281. .19426,20077. .21210)
                     /gene="tfg-1"
                     /locus_tag="Y63D3A.5"
                     /standard_name="Y63D3A.5"
                     /note="C. elegans TFG-1 protein; contains similarity to
                     Pfam domain PF00564 (PB1 domain)"
                     /codon_start=1
                     /product="Hypothetical protein Y63D3A.5"
                     /protein_id="CAB63398.1"
                     /db_xref="GI:6580329"
                     /db_xref="InterPro:IPR000270"
                     /db_xref="UniProt/TrEMBL:Q9U1W1"
                     /translation="MVHSNGAITSTILKARHADVVRKTSLHHANDLTLIDLVLNVQRL
                     LALPSDANFVLKYKDEEGDLVTLAEDSDLLLALHTSGATLDVTVVVDSRAREAVHDVQ
                     KQVEQIKLDVGKLLGALSALDVAQIAEQSNTSVANLSAPKQSHHDNIVFQKSFEAAPP
                     SPVPSEKAELPATIQPSVHEQFNHRPAHVEEEIPLENHYAPPPHQQIPDDLNTSFSSQ
                     PPPPIEQFGAIPPPNATIPSFPTSNAASPPVQEFAPPPPQQQQQQFQAPPPPMASHSS
                     ISSTPVQQQQGFAPPQQFGGPPPSGPPSEYGGYAPPQQQQQQFGAPPPQGAPQQGFGAP
                     PQGPPQGGPPQGSFGAPPPQQFHAPSPQSFGGPPPPVSSAPGNFAPPPQSGPPGAFAP
                     PPSAFGAPQGPGGPGGYGPPPPGGPGAPGSYGPPQGGPGGFGPPPPGGPGAYGPPPTG
                     FPPVGAPPPGAAGAPGGNPFARGPSATGYRQSPYQQ"
     gene            complement(join(21795. .21920,22061. .22165,23337. .23672,
                     24863. .25354,26233. .26446,27066. .27309,27804. .27912,
                     27973. .28209,29577. .29732,29793. .29908,29969. .30092))

Alignment Scores:
Pred. No.:                481          Length:         33899
Score:                    42.00        Matches:        7
Percent Similarity:       100.0%       Conservative:   2
Best Local Similarity:    77.8%        Mismatches:     0
Query Match:              93.3%        Indels:         0
DB:                       2            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CEY63D3A (1-33899)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||::::::|||||||||||||||||
Db      16717 CCGGACATCCTATTCCTCCAAGAAGTC 16691

RESULT 41
BC083404
LOCUS       BC083404               1451 bp    mRNA    linear   VRT 27-MAY-2005
DEFINITION  Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7238334), partial
            cds.
ACCESSION   BC083404
VERSION     BC083404.1  GI:53734140
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 1451)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED    12477932
REFERENCE   2  (bases 1 to 1451)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: John Postlethwait, University of Oregon
            cDNA Library Preparation: Dr. Sumio Sugano
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 192 Row: c Column: 18.
FEATURES              Location/Qualifiers
     source           1. .1451
                      /organism="Danio rerio"
                      /mol_type="mRNA"
                      /db_xref="taxon:7955"
                      /clone="IMAGE:7238334"
                      /tissue_type="Ovary"
                      /clone_lib="NIH_ZGC_5"
                      /lab_host="DH10B"
                      /note="Vector: pME18S-FL3"
     gene             <1. .1451
                      /gene="TTRAP"
                      /db_xref="GeneID:553516"
     CDS              <1. .1141
                      /gene="TTRAP"
                      /codon_start=2
                      /product="TTRAP protein"
                      /protein_id="AAH83404.1"
                      /db_xref="GI:53734141"
                      /db_xref="GeneID:553516"
                      /translation="FLSGNQTDQQMSALEESRTRLCDQFAFVSGSDSAVAQCYLAENE
                      WDMERALNSFFEAHMDSVFDEEAEKTEVTGNKRKDDTAEASGTKKKLKTDNADCIDLT
                      AEEPTCSITVNSKENQAENGTAKSEVEDSKLSIISWNVDGLDTLNLADRARGLCSYLA
                      LYTPDVVFLQELIPAYVQYLKKRAVSYLFFEGSDDGYFTGIMLRKSRVKFLESEIICF
                      PTTQMMRNLLIAQVTFSGQKLYLMTSHLESCKNQSQERTKQLRVVLQKIKEAPEDAIV
                      IFAGDTNLRDAEVANVGGLPAGVCDVWEQLGKQEHCRYTWDTKANSNKTVPYVSRCRF
                      DRIFLRSAKTGPPVTPDHMALIGMEKLDCGRYTSDHWGIYCTFNT"
ORIGIN

Alignment Scores:
```

```
Pred. No.:              29.7          Length:         1451
Score:                  41.00         Matches:        7
Percent Similarity:     100.0%        Conservative:   2
Best Local Similarity:  77.8%         Mismatches:     0
Query Match:            91.1%         Indels:         0
DB:                     5             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BC083404 (1-1451)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||:::
Db        491 CCAGATGTGGTTTTTCTTCAAGAGCTC 517

RESULT 42
BC097117
LOCUS       BC097117               1455 bp    mRNA    linear   VRT 23-JUN-2005
DEFINITION  Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7431816), partial
            cds.
ACCESSION   BC097117
VERSION     BC097117.1  GI:66910444
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 1455)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1455)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Will Talbot
            cDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 226 Row: g Column: 18
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 53734140.
FEATURES             Location/Qualifiers
     source          1. .1455
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:7431816"
                     /tissue_type="Embryo, whole, 2-8 hr, mixed wild type, pool
                     of 2500 embryos"
                     /clone_lib="NIH_ZGC_15"
                     /lab_host="DH10B"
                     /note="Vector: pME18S-FL3"
     gene            <1. .1455
                     /gene="TTRAP"
                     /db_xref="GeneID:553516"
     CDS             <1. .438
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TTRAP protein"
                     /protein_id="AAH97117.1"
                     /db_xref="GI:66910445"
                     /db_xref="GeneID:553516"
                     /translation="VFSGNQTDQQMSALGESRTRLCDQFAFVSGSDSAVAQCYLAENE
                     WDMERALNSFFEAHMDSVFDEEAEKTEVTGNKRKDDTAEASGTKKKLKTDNADCIDLT
                     AEEPTCSITVNSKKIRLKMAQPNLKWRTASSPSSAGMWTVWTH"
ORIGIN

Alignment Scores:
Pred. No.:              29.8          Length:         1455
Score:                  41.00         Matches:        7
Percent Similarity:     100.0%        Conservative:   2
Best Local Similarity:  77.8%         Mismatches:     0
Query Match:            91.1%         Indels:         0
DB:                     5             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BC097117 (1-1455)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||:::
Db        489 CCAGATGTGGTTTTTCTTCAAGAGCTC 515

RESULT 43
BX957279
LOCUS       BX957279               198752 bp    DNA     linear   VRT 29-APR-2005
DEFINITION  Zebrafish DNA sequence from clone DKEY-218N20 in linkage group 19,
            complete sequence.
ACCESSION   BX957279
VERSION     BX957279.13  GI:62989971
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 198752)
  AUTHORS   Clark,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:61673710.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
```

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-218N20 is from a Zebrafish BAC library
VECTOR: pIndigoBAC-5.

FEATURES Location/Qualifiers
source 1. .198752
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-218N20"
/clone_lib="DanioKey"

ORIGIN

Alignment Scores:

Pred. No.:	5.21e+03	Length:	198752
Score:	41.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	91.1%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BX957279 (1-198752)

Qy 1 ProAspValIlePheLeuGlnGluVal 9
||||||||||:::|||||||||||||:::
Db 144135 CCAGATGTGGTTTTTCTTCAAGAGCTC 144161

RESULT 44
BX511258
LOCUS BX511258 200076 bp DNA linear VRT 02-APR-2004
DEFINITION Zebrafish DNA sequence from clone CH211-81E5 in linkage group 19, complete sequence.
ACCESSION BX511258
VERSION BX511258.6 GI:42820869
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 200076)
AUTHORS Heath,P.
TITLE Direct Submission
JOURNAL Submitted (01-APR-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Feb 25, 2004 this sequence version replaced gi:42538803.
---------- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-81E5 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.

FEATURES Location/Qualifiers
source 1. .200076
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-81E5"
/clone_lib="CHORI-211"

ORIGIN

Alignment Scores:

Pred. No.:	5.25e+03	Length:	200076
Score:	41.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	91.1%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BX511258 (1-200076)

Qy 1 ProAspValIlePheLeuGlnGluVal 9
||||||||||:::|||||||||||||:::
Db 85197 CCAGATGTGGTTTTTCTTCAAGAGCTC 85223

RESULT 45
AC102963/c
LOCUS AC102963 223042 bp DNA linear HTG 10-MAY-2003
DEFINITION Rattus norvegicus clone CH230-184H18, WORKING DRAFT SEQUENCE, 3 unordered pieces.
ACCESSION AC102963
VERSION AC102963.5 GI:30522814
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 223042)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

```
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 223042)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 223042)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:23096458.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GGJK
                Center clone name: CH230-184H18
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 206116 bases at least Q40
                Consensus quality: 208918 bases at least Q30
                Consensus quality: 210843 bases at least Q20
                Estimated insert size: 213528; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   219905: contig of 219905 bp in length
            *   219906   220005: gap of unknown length
            *   220006   221450: contig of 1445 bp in length
            *   221451   221550: gap of unknown length
            *   221551   223042: contig of 1492 bp in length.
FEATURES             Location/Qualifiers
     source          1. .223042
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-184H18"
     misc_feature    1. .2275
                     /note="wgs_contig"
     misc_feature    216649. .218004
                     /note="wgs_contig"
     misc_feature    218598. .219905
                     /note="wgs_contig"
     gap             219906. .220005
                     /estimated_length=unknown
     gap             221451. .221550
                     /estimated_length=unknown
ORIGIN

Alignment Scores:
Pred. No.:              5.88e+03        Length:         223042
Score:                  41.00           Matches:        8
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            91.1%           Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AC102963 (1-223042)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||||||||||||
Db      79024 CCTGATGTCATATTTTACAGGAA 79001

RESULT 46
AC155120/c
LOCUS       AC155120              261222 bp    DNA     linear   HTG 01-JUL-2005
DEFINITION  Bos taurus clone CH240-44I24, *** SEQUENCING IN PROGRESS ***, 39
            unordered pieces.
ACCESSION   AC155120
VERSION     AC155120.2  GI:68265287
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
```

REFERENCE 1 (bases 1 to 261222)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 261222)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 261222)
AUTHORS Cow Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jun 28, 2005 this sequence version replaced gi:57337987.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```
-------------- Genome Center
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
-------------- Project Information
    Center project name: FCGL
    Center clone name: CH240-44I24
-------------- Summary Statistics
    Assembly program: Atlas 3.0;
    Consensus quality: 240483 bases at least Q40
    Consensus quality: 243544 bases at least Q30
    Consensus quality: 246587 bases at least Q20
    Estimated insert size: 244914; sum-of-contigs estimation
    Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
--------------
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*       1     4327: contig of 4327 bp in length
*    4328     4831: gap of 504 bp
*    4832     6238: contig of 1407 bp in length
*    6239     6338: gap of unknown length
*    6339     7998: contig of 1660 bp in length
*    7999     8338: gap of 340 bp
*    8339    13333: contig of 4995 bp in length
*   13334    14819: gap of 1486 bp
*   14820    20364: contig of 5545 bp in length
*   20365    20414: gap of 50 bp
*   20415    22807: contig of 2393 bp in length
*   22808    22857: gap of 50 bp
*   22858    29973: contig of 7116 bp in length
*   29974    30023: gap of 50 bp
*   30024    31125: contig of 1102 bp in length
*   31126    31225: gap of unknown length
*   31226    49964: contig of 18739 bp in length
*   49965    50014: gap of 50 bp
*   50015    60171: contig of 10157 bp in length
*   60172    60221: gap of 50 bp
*   60222    63982: contig of 3761 bp in length
*   63983    64032: gap of 50 bp
*   64033   128938: contig of 64906 bp in length
*  128939   128988: gap of 50 bp
*  128989   143211: contig of 14223 bp in length
*  143212   143412: gap of 201 bp
*  143413   157356: contig of 13944 bp in length
*  157357   159561: gap of 2205 bp
*  159562   161968: contig of 2407 bp in length
*  161969   162068: gap of unknown length
*  162069   165155: contig of 3087 bp in length
*  165156   165205: gap of 50 bp
*  165206   166749: contig of 1544 bp in length
*  166750   166849: gap of unknown length
*  166850   171056: contig of 4207 bp in length
*  171057   171303: gap of 247 bp
*  171304   179264: contig of 7961 bp in length
*  179265   179314: gap of 50 bp
*  179315   184923: contig of 5609 bp in length
*  184924   184973: gap of 50 bp
*  184974   192646: contig of 7673 bp in length
*  192647   192696: gap of 50 bp
*  192697   194342: contig of 1646 bp in length
```

```
            *   194343   197750: gap of 3408 bp
            *   197751   204681: contig of 6931 bp in length
            *   204682   204731: gap of 50 bp
            *   204732   208806: contig of 4075 bp in length
            *   208807   208856: gap of 50 bp
            *   208857   216297: contig of 7441 bp in length
            *   216298   216347: gap of 50 bp
            *   216348   222209: contig of 5862 bp in length
            *   222210   222259: gap of 50 bp
            *   222260   224701: contig of 2442 bp in length
            *   224702   224751: gap of 50 bp
            *   224752   235290: contig of 10539 bp in length
            *   235291   235340: gap of 50 bp
            *   235341   239385: contig of 4045 bp in length
            *   239386   239485: gap of unknown length
            *   239486   240486: contig of 1001 bp in length
            *   240487   240586: gap of unknown length
            *   240587   242297: contig of 1711 bp in length
            *   242298   242397: gap of unknown length
            *   242398   243833: contig of 1436 bp in length
            *   243834   243933: gap of unknown length
            *   243934   244986: contig of 1053 bp in length
            *   244987   245086: gap of unknown length
            *   245087   246696: contig of 1610 bp in length
            *   246697   246796: gap of unknown length
            *   246797   247820: contig of 1024 bp in length
            *   247821   247920: gap of unknown length
            *   247921   249325: contig of 1405 bp in length
            *   249326   249425: gap of unknown length
            *   249426   250929: contig of 1504 bp in length
            *   250930   251029: gap of unknown length
            *   251030   254479: contig of 3450 bp in length
            *   254480   254579: gap of unknown length
            *   254580   261222: contig of 6643 bp in length.
FEATURES             Location/Qualifiers
     source          1. .261222
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9913"
                     /clone="CH240-44I24"
     gap             4328. .4831
                     /estimated_length=504
     gap             6239. .6338

Alignment Scores:
Pred. No.:              6.95e+03        Length:         261222
Score:                  41.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  77.8%           Mismatches:     0
Query Match:            91.1%           Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AC155120 (1-261222)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||:::|||||||||||||||:::
Db     203419 CCTGATTTAATCTTCCTCCAAGAAATC 203393

RESULT 47
BX640415/c
LOCUS       BX640415               347071 bp    DNA     linear   BCT 17-APR-2005
DEFINITION  Bordetella pertussis strain Tohama I, complete genome; segment
            5/12.
ACCESSION   BX640415 BX470248
VERSION     BX640415.1  GI:33572084
KEYWORDS    complete genome.
SOURCE      Bordetella pertussis Tohama I
  ORGANISM  Bordetella pertussis Tohama I
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Bordetella.
REFERENCE   1
  AUTHORS   Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
            Harris,D.E., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,
            Mungall,K.L., Cerdeno-Tarraga,A.M., Temple,L., James,K., Harris,B.,
            Quail,M.A., Achtman,M., Atkin,R., Baker,S., Basham,D., Bason,N.,
            Cherevach,I., Chillingworth,T., Collins,M., Cronin,A., Davis,P.,
            Doggett,J., Feltwell,T., Goble,A., Hamlin,N., Hauser,H.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Norberczak,H.,
            O'Neil,S., Ormond,D., Price,C., Rabbinowitsch,E., Rutter,S.,
            Sanders,M., Saunders,D., Seeger,K., Sharp,S., Simmonds,M.,
            Skelton,J., Squares,R., Squares,S., Stevens,K., Unwin,L.,
            Whitehead,S., Barrell,B.G. and Maskell,D.J.
  TITLE     Comparative analysis of the genome sequences of Bordetella
            pertussis, Bordetella parapertussis and Bordetella bronchiseptica
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 347071)
  AUTHORS   Sebaihia,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-2003) Submitted on behalf of the Pathogen
            Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA, E-mail: ms5@sanger.ac.uk
FEATURES             Location/Qualifiers
     source          1. .347071
                     /organism="Bordetella pertussis Tohama I"
                     /mol_type="genomic DNA"
                     /strain="Tohama I"
                     /db_xref="taxon:257313"
     gene            100. .3071
                     /locus_tag="BP1322"
                     /pseudo
     CDS             join(100. .1953,1953. .3071)
                     /locus_tag="BP1322"
                     /note="Pseudogene. This CDS appears to have a frameshift
                     mutation following codon 618. The sequence has been
                     checked and believed to be correct. Similar, over its
                     C-terminal region, to Pseudomonas syringae histidine
                     protein kinase-response regulator hybrid protein CvgsY
                     TR:Q9WWH5 (EMBL:AF133263) (671 aa) fasta scores: E():
                     2.6e-60, 38.48% id in 569 aa, and to Rhizobium loti
                     two-component sensor histidine kinase mll8067 TR:Q984C1
                     (EMBL:AP003013) (1386 aa) fasta scores: E(): 7.7e-49,
                     41.6% id in 387 aa"
                     /pseudo
                     /codon_start=1
                     /transl_table=11
                     /product="two-component sensor histidine kinase
                     (Pseudogene)"
                     /db_xref="PSEUDO:CAE41616.1"
     misc_feature    226. .645
                     /locus_tag="BP1322"
                     /note="HMMPfam hit to PF01590, GAF domain"
                     /pseudo
     misc_feature    766. .849
                     /locus_tag="BP1322"
                     /note="HMMPfam hit to PF00989, PAS domain"
                     /pseudo
     misc_feature    1546. .1740
                     /locus_tag="BP1322"
                     /note="HMMPfam hit to PF00989, PAS domain"
                     /pseudo
     misc_feature    join(1942. .1953,1953. .2147)
                     /locus_tag="BP1322"
                     /note="HMMPfam hit to PF00512, His Kinase A
                     (phosphoacceptor) domain"
                     /pseudo
     variation       1954
                     /locus_tag="BP1322"
                     /note="C(-)G in pertussis; C(T)G in parapertussis and
                     bronchiseptica"
     misc_feature    2268. .2639
                     /locus_tag="BP1322"
                     /note="HMMPfam hit to PF02518, Histidine kinase-, DNA
                     gyrase B-, phytochrome-like ATPase"
                     /pseudo
     misc_feature    2703. .3041
                     /locus_tag="BP1322"
```

```
                /note="HMMPfam hit to PF00072, Response regulator receiver
                domain"
                /pseudo
     gene       complement(3062. .5605)
                /locus_tag="BP1323"
     CDS        complement(3062. .5605)
                /locus_tag="BP1323"
                /note="Similar to Streptomyces coelicolor putative alpha
                amylase Scbac1a6.05C TR:Q9ADI2 (EMBL:AL589708) (805 aa)
                fasta scores: E(): 8.8e-23, 38.42% id in 864 aa, and to
                Pseudomonas aeruginosa probable glycosyl hydrolase Pa2162
                TR:Q9I1V3 (EMBL:AE004643) (926 aa) fasta scores: E():
                4.1e-28, 45.05% id in 930 aa"
                /codon_start=1
                /transl_table=11
                /product="probable alpha amylase"
                /protein_id="CAE41617.1"
                /db_xref="GI:33572085"
                /db_xref="GOA:Q7VYK3"
                /db_xref="InterPro:IPR006047"
                /db_xref="UniProt/TrEMBL:Q7VYK3"
                /translation="MIRATVRLQLHAGFTLDAAGAQLDYYDALGISHLYLSPVARARC
                GSTHGYDVIDHAQVNPELGGLAALRRLAGAARARGMGLILDIVPNHMAAHPDNRWWRE
                VLRDGPASAYADWFDIDWRPPGRTLRGRVLLPVLTCGARAALASGEIGLRAGSDGTPA
                LHANGLDLPLAPASLPMAALLLHHGGTPAQWHALLRAQAYRLAWWRSAAQAINWRRFF
                EIGDLVGVRVERPEVFDAVHALPLALYRAGLIDGLRIDHVDGLAEPGAYCRRLRAELA
                QAGQARAARGLSAEPYVVVEKILGDKETLRTDWPVQGTTGYDFMNEASAWLHDARGAA
                PLAAFWRDNAGTWPHWREQLRRIRLRLLRRHFGAERRALARSLADLPELRASAAAIDP
                VLTQWLACFPVYRTYAEGGAGQAADQAWRQTAQTRANAHLDPAQQALLAALCLALDGR
                MADAPHRLALRRLQQLTPPLAAKALEDTWFYRDGTLLSRNEVGPAPAQFALGGGALRR
                SVAARARRHPQALLATASHDHKRGEDVRARLAVLSEMPQQWCDWARQWLADAGGSAAP
                PAWADRYMLLQTLVGAWPADLAPDDQAGMAALLRRVAAWQAKALREARQRSSWTCPDL
                AYETLCARYLRALADDGAALPAVAALARRLAPAGLVNSLAQTALRLGLPGVPDLYQGC
                ETWDLSLVDPDNRRAVDYAARRGALASLGNGALALPELWRDGRAKQALVARALQWRRE
                QPALFEHGACRWLRVSGARARHVLACLRTRGPQQALIVVPRASALALAGQTGPDRDAA
                AHWRDTCVHLPADAPPWRDILTHAGHQGRALALSELLARWPVALCRVQS"
misc_feature    complement(4154. .5593)
                /locus_tag="BP1323"
                /note="HMMPfam hit to PF00128, Alpha amylase, catalytic
                domain"
     gene       complement(join(5602. .6896,6898. .6975))
                /gene="malQ"
                /locus_tag="BP1324"
                /note="synonym: malA"
                /pseudo
     CDS        complement(join(5602. .6576,6576. .6896,6898. .6975))
                /gene="malQ"
                /locus_tag="BP1324"
                /EC_number="2.4.1.25"
                /note="Pseudogene. This CDS appears to contain a deletion
                that has lead to a shift in reading frame around codon
                133, and also appears to have a frameshift mutation
                following codon 26. The sequence has been checked and
                believed to be correct. Similar to Escherichia coli
                4-alpha-glucanotransferase MalQ or MalA or B3416
                SWALL:MALQ_ECOLI (SWALL:P15977) (694 aa) fasta scores:
                E(): 2.1e-25, 38.31% id in 327 aa"
                /pseudo
                /codon_start=1
                /transl_table=11
                /product="putative 4-alpha-glucanotransferase
                (pseudogene)"
                /db_xref="PSEUDO:CAE41618.1"
variation       complement(6576. .6578)
                /gene="malQ"
                /locus_tag="BP1324"
                /note="large deletion relative to parapertussis and
                bronchiseptica"
variation       complement(6898. .6900)
                /gene="malQ"
                /locus_tag="BP1324"
                /note="CGC(-) in pertussis; CGC(A) in parapertussis and
                bronchiseptica"
     gene       complement(6972. .8726)
                /locus_tag="BP1326"
     CDS        complement(6972. .8726)
                /locus_tag="BP1326"
                /note="Similar to Sulfolobus shibatae maltooligosyl
                trehalose trehalohydrolase TreZ TR:Q9UWN9 (EMBL:AF201335)
                (559 aa) fasta scores: E(): 3.4e-50, 34.79% id in 500 aa,
                and to Pseudomonas aeruginosa probable glycosyl hydrolase
                Pa2164 TR:Q9I1V1 (EMBL:AE004643) (583 aa) fasta scores:
                E(): 3.1e-101, 53.79% id in 554 aa"
                /codon_start=1
                /transl_table=11
                /product="probable alpha amylase"
                /protein_id="CAE41619.1"
                /db_xref="GI:33572086"
                /db_xref="GOA:Q7VYK2"
                /db_xref="InterPro:IPR004193"
                /db_xref="InterPro:IPR006047"
                /db_xref="UniProt/TrEMBL:Q7VYK2"
                /translation="MNEPAPLGATVCAPGKIRFRLWAPSAPPALALEIEGRAPIALRP
                GPDGYAQAIVDGAPGLRYRYRLAADCVVPDPASRLQDGDVHGHSVVVGPDRYPWRHTG
                WRGRPWQDSVIYEAHAGLCGGYAGLRRRLPEVARLGVTLLELMPIADFPGPRNWGYDG
                VLPYAPDTAYGTPDELKALIDAAHGLEMGVMLDVVYNHFGPDGNYLHRYAADFFDARA
                DTPWGAAIDYRQQAVQRYFIDNALYWLRKYRFDGLRLDAVHAMGRTDWLAGLAREARA
                ACPDRRVHLVVENDANQASLLEQGFDAQWNDDAHHVLHHLLTGERQGYYADYAQAPAA
                LLARALGEGYVYQGQAAPTRGGRARGEPSAHLPPWAFVLFLHNHDQVGNRADGLRLTS
                LLAPGSPALRAAIALQLLAPHIPLLFMGEEYGSTAPFFYFTSHGPELAAAVRAGRARE
                FAASMHDCDPPPDPNDPDTYRRSCPWPPPGAERQAWFEYYRELLRLRARLLGERLTGA
                RADGAAALGPAAVRAAWRLGDGARLTLYANLLPGSVIVPAPPAGQLAYATLAGAGQDL
                VAGVLRGGCALALWEPAP"
misc_feature    complement(7005. .8393)
                /locus_tag="BP1326"
                /note="HMMPfam hit to PF00128, Alpha amylase, catalytic
                domain"
     gene       complement(8723. .10843)
                /gene="glgX"
                /locus_tag="BP1327"
     CDS        complement(8723. .10843)
                /gene="glgX"
                /locus_tag="BP1327"
                /note="Similar to Escherichia coli glycogen operon protein
```

```
Alignment Scores:
Pred. No.:              9.36e+03        Length:         347071
Score:                  41.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  77.8%           Mismatches:     0
Query Match:            91.1%           Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BX640415 (1-347071)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||::::::|||||||||||||||||
Db      26269 CCGGATCTCGTGTTCCTGCAGGAAGTG 26243

RESULT 48
BX640437
LOCUS       BX640437               347356 bp    DNA     linear   BCT 17-APR-2005
DEFINITION  Bordetella bronchiseptica strain RB50, complete genome; segment
            1/16.
ACCESSION   BX640437 BX470250
VERSION     BX640437.1  GI:33575039
KEYWORDS    complete genome.
SOURCE      Bordetella bronchiseptica RB50
  ORGANISM  Bordetella bronchiseptica RB50
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Bordetella.
REFERENCE   1
  AUTHORS   Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
            Harris,D.E., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,
            Mungall,K.L., Cerdeno-Tarraga,A.M., Temple,L., James,K., Harris,B.,
            Quail,M.A., Achtman,M., Atkin,R., Baker,S., Basham,D., Bason,N.,
```

```
            Cherevach,I., Chillingworth,T., Collins,M., Cronin,A., Davis,P.,
            Doggett,J., Feltwell,T., Goble,A., Hamlin,N., Hauser,H.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Norberczak,H.,
            O'Neil,S., Ormond,D., Price,C., Rabbinowitsch,E., Rutter,S.,
            Sanders,M., Saunders,D., Seeger,K., Sharp,S., Simmonds,M.,
            Skelton,J., Squares,R., Squares,S., Stevens,K., Unwin,L.,
            Whitehead,S., Barrell,B.G. and Maskell,D.J.
  TITLE     Comparative analysis of the genome sequences of Bordetella
            pertussis, Bordetella parapertussis and Bordetella bronchiseptica
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 347356)
  AUTHORS   Sebaihia,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-2003) Submitted on behalf of the Pathogen
            Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA, E-mail: ms5@sanger.ac.uk
FEATURES             Location/Qualifiers
     source          1. .347356
                     /organism="Bordetella bronchiseptica RB50"
                     /mol_type="genomic DNA"
                     /strain="RB50"
                     /db_xref="taxon:257310"
     gene            1. .1920
                     /gene="gidA"
                     /locus_tag="BB0001"
     CDS             1. .1920
                     /gene="gidA"
                     /locus_tag="BB0001"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0001"
                     /codon_start=1
                     /transl_table=11
                     /product="glucose inhibited division protein A"
                     /protein_id="CAE30503.1"
                     /db_xref="GI:33575040"
                     /db_xref="GOA:Q7WRF1"
                     /db_xref="InterPro:IPR000205"
                     /db_xref="InterPro:IPR002218"
                     /db_xref="InterPro:IPR004416"
                     /db_xref="UniProt/TrEMBL:Q7WRF1"
                     /translation="MDFPREFDVIVVGGGHAGTEAALAAARAGAQTLLLTHNIETLGQ
                     MSCNPSIGGIGKGHLVKEVDALGGAMAIATDEAGIQFRILNSSKGPAVRATRAQADRV
                     LYRNAIRARLENQPNLWLFQQAVDDLMVQGDQVVGAVTQIGLRFRARTVVLTAGTFLN
                     GLIHVGLQNYSGGRAGDPPANSLGQRLKELQLPQGRLKTGTPPRIDGRSINYSVLEEQ
                     PGDLDPVPVFSFLGKASMHPRQLPCWITHTNARTHEIIRGGLDRSPMYSGVIEGVGPR
                     YCPSIEDKIHRFADKASHQVFLEPEGLNTHEIYPNGVSTSLPFDVQYELIHSLPGLEN
                     AHILRPGYAIEYDYFDPRALKSTLETKAISGLFFAGQINGTTGYEEAAAQGLLAGANA
                     ALQAQGKEPWVPRRDEAYLGVLVDDLVTRGVTEPYRMFTSRAEYRLSLREDNADLRLT
                     EIGRRLGLVDDVRWDAFSRKRDAVAQEVERLKSTWVNPRVLPAHAAEALLGKAIEREY
                     SLSDLLKRPNVSYEALMQARTDEGELLAGPGVLEDDVLAEQVETQVKYAGYIARQQDE
                     VQKHLLSHEQQPIPADIDYDAVTSLSFEVRQKLKTHRPETIGQAARVSGVTPAAISLLL
                     IHLKRLHYGSRKQAA"
     misc_feature    19. .1893
                     /gene="gidA"
                     /locus_tag="BB0001"
                     /note="HMMPfam hit to PF01134, DE Glucose inhibited
                     division protein A, score 0"
     misc_feature    820. .864
                     /gene="gidA"
                     /locus_tag="BB0001"
                     /note="PS01280 Glucose inhibited division protein A family
                     signature 1."
     misc_feature    1810. .1875
                     /gene="gidA"
                     /locus_tag="BB0001"
                     /note="Predicted helix-turn-helix motif with score
                     1129.000, SD 3.03 at aa 604-625, sequence
                     ETIGQAARVSGVTPAAISLLLI"
     gene            1917. .2549
                     /gene="gidB"
                     /locus_tag="BB0002"
     CDS             1917. .2549
                     /gene="gidB"
                     /locus_tag="BB0002"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0002"
                     /codon_start=1
                     /transl_table=11
                     /product="glucose inhibited division protein B"
                     /protein_id="CAE30504.1"
                     /db_xref="GI:33575041"
                     /db_xref="GOA:Q7WRF0"
                     /db_xref="InterPro:IPR003682"
                     /db_xref="UniProt/Swiss-Prot:Q7WRF0"
                     /translation="MSAVPDIPGGPAQRLAQACDALRLPADAGQQQKLLRYIEQMQRW
                     NRTYNLTAIRDPGQMLVQHLFDSLSVVAPLERGLPGVVLAIMRAHWDVTCVDAVEKKT
                     AFVRQMAGALGLPNLQAAHTRIEQLEPAQCDVVISRAFASLQDFAKLAGRHVREGGTL
                     VAMKGKVPDDEIQALQQHGHWTVERIEPLVVPALDAQRCLIWMRRSQGNI"
     misc_feature    1917. .2135
                     /gene="gidB"
                     /locus_tag="BB0002"
                     /note="PS00430 TonB-dependent receptor proteins signature
                     1."
     misc_feature    2004. .2519
                     /gene="gidB"
                     /locus_tag="BB0002"
                     /note="HMMPfam hit to PF02527, DE Glucose inhibited
                     division protein, score 4.7e-42"
     gene            2546. .3343
                     /gene="parA"
                     /locus_tag="BB0003"
                     /note="synonym: soj"
     CDS             2546. .3343
                     /gene="parA"
                     /locus_tag="BB0003"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0003"
                     /codon_start=1
                     /transl_table=11
                     /product="ParA family protein"
                     /protein_id="CAE30505.1"
                     /db_xref="GI:33575042"
                     /db_xref="GOA:Q7WRE9"
                     /db_xref="InterPro:IPR002586"
                     /db_xref="UniProt/TrEMBL:Q7WRE9"
                     /translation="MKNIPPSKSARVPCIANQKGGVGKTTTAINLAAGLATHKQRVLL
                     VDLDPQGNATMGSGIDKSTLESNLYQVLIGEAGIEQTRVRSESGGYDVLPANRELSGA
                     EIDLVQMDERERQLKAAIDKIAGEYDFVLIDCPPTLSLLTLNGLAAAHGVIIPMQCEY
                     FALEGLSDLVNTIKRVHRNINNELRVIGLLRVMFDPRMTLQQQVSAQLESHFGDKVFT
                     TVVPRNVRLAEAPSYGMPGVVYDRASRGAQAYIAFGAEMIERVKELD"
     misc_feature    2819. .3154
                     /gene="parA"
                     /locus_tag="BB0003"
                     /note="HMMPfam hit to PF00991, DE ParA family ATPase,
                     score 1.5e-33"
     gene            3350. .4156
                     /locus_tag="BB0004"
     CDS             3350. .4156
                     /locus_tag="BB0004"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0004"
                     /codon_start=1
                     /transl_table=11
                     /product="Putative acetyltransferase, GnaT family"
                     /protein_id="CAE30506.1"
                     /db_xref="GI:33575043"
                     /db_xref="GOA:Q7WRE8"
                     /db_xref="InterPro:IPR000182"
                     /db_xref="UniProt/TrEMBL:Q7WRE8"
                     /translation="MPAVLAVPAAHPAPAAGGELARLEEAALNATTVREQMLYEGWLV
                     RWAPSRAKRARCVNVLGLSTRPLDERLAYCTRWYAGHGVPLIVRLTSIGPDFSLDAQL
                     HERGFVAFDRTCVMAAPVVPFTAQAPSGLVFERTTPGRFAQEAGRLRGHDDAATAEHA
                     ARLQAAPLENLPLLVRDASGECIAAGLAAFDAELVGLFDIVTAPQRRRQGCGAAVLKY
                     LLEQASHAGARQAYLQVEPQNTAARALYARFGFVDRYTYWYRSMSEATPV"
     misc_feature    3869. .4105
                     /locus_tag="BB0004"
                     /note="HMMPfam hit to PF00583, DE Acetyltransferase (GNAT)
                     family, score 2.5e-14"
     gene            4200. .5117
                     /gene="parB"
```

```
                     /locus_tag="BB0005"
     CDS             4200. .5117
                     /gene="parB"
                     /locus_tag="BB0005"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0005"
                     /codon_start=1
                     /transl_table=11
                     /product="probable chromosome partitioning protein ParB"
                     /protein_id="CAE30507.1"
                     /db_xref="GI:33575044"
                     /db_xref="GOA:Q7WRE7"
                     /db_xref="InterPro:IPR003115"
                     /db_xref="InterPro:IPR004437"
                     /db_xref="UniProt/TrEMBL:Q7WRE7"
                     /translation="MATKKPKGLGRGLDALLGADVPAIDNIGKAPAAAPEGPPATLPV
                     SKMRAGKYQPRTRMDEGALNELAESIRTQGIMQPILVRALGADAPGQYEIIAGERRFR
                     AAQLAGLKEVPVLVREVADENAAVMALIENIQREDLNPLEEAHGVRRLLDEFGLTHEQ
                     AAQAIGRSRSATSNLLRLLNLAAPVQTMLLAGDVDMGHARALLAVDAATQIQLANQVI
                     AKRLSVREAEKLVARAQKDVDAAPRKKGNGASRDVTRLEEALSDHLGTRVALKVGARE
                     KGQIVIDFHGWEHLNSLLERQGLSGVVDA"
     misc_feature    4317. .4598
                     /gene="parB"
                     /locus_tag="BB0005"
                     /note="HMMPfam hit to PF02195, DE ParB-like nuclease
                     domain, score 5.9e-34"
     misc_feature    4665. .4730
                     /gene="parB"
                     /locus_tag="BB0005"
                     /note="Predicted helix-turn-helix motif with score
                     1537.000, SD 4.42 at aa 156-177, sequence
                     LTHEQAAQAIGRSRSATSNLLR"
     gene            5110. .6111
                     /gene="ansB"
                     /locus_tag="BB0006"
     CDS             5110. .6111
                     /gene="ansB"
                     /locus_tag="BB0006"
                     /EC_number="3.5.1.1"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0006"

Alignment Scores:
Pred. No.:              9.37e+03        Length:         347356
Score:                  41.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  77.8%           Mismatches:     0
Query Match:            91.1%           Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BX640437 (1-347356)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::||||||||||||||||
Db     118230 CCGGACCTGGTCTTCCTGCAGGAGGTC 118256

RESULT 49
BX640423
LOCUS       BX640423               348251 bp    DNA     linear   BCT 17-APR-2005
DEFINITION  Bordetella parapertussis strain 12822, complete genome; segment
            1/14.
ACCESSION   BX640423 BX470249
VERSION     BX640423.1  GI:33564799
KEYWORDS    complete genome.
SOURCE      Bordetella parapertussis
  ORGANISM  Bordetella parapertussis
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Bordetella.
REFERENCE   1
  AUTHORS   Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
            Harris,D.E., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,
            Mungall,K.L., Cerdeno-Tarraga,A.M., Temple,L., James,K., Harris,B.,
            Quail,M.A., Achtman,M., Atkin,R., Baker,S., Basham,D., Bason,N.,
            Cherevach,I., Chillingworth,T., Collins,M., Cronin,A., Davis,P.,
            Doggett,J., Feltwell,T., Goble,A., Hamlin,N., Hauser,H.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Norberczak,H.,
            O'Neil,S., Ormond,D., Price,C., Rabbinowitsch,E., Rutter,S.,
            Sanders,M., Saunders,D., Seeger,K., Sharp,S., Simmonds,M.,
            Skelton,J., Squares,R., Squares,S., Stevens,K., Unwin,L.,
            Whitehead,S., Barrell,B.G. and Maskell,D.J.
  TITLE     Comparative analysis of the genome sequences of Bordetella
            pertussis, Bordetella parapertussis and Bordetella bronchiseptica
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 348251)
  AUTHORS   Sebaihia,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-2003) Submitted on behalf of the Pathogen
            Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA, E-mail: ms5@sanger.ac.uk
FEATURES             Location/Qualifiers
     source          1. .348251
                     /organism="Bordetella parapertussis"
                     /mol_type="genomic DNA"
                     /strain="12822"
                     /db_xref="taxon:519"
     gene            1. .1920
                     /gene="gidA"
                     /locus_tag="BPP0001"
     CDS             1. .1920
                     /gene="gidA"
                     /locus_tag="BPP0001"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0001
                     Similar to Escherichia coli glucose inhibited division
                     protein A GidA or B3741 SWALL:GIDA_ECOLI (SWALL:P17112)
                     (629 aa) fasta scores: E(): 6.8e-164, 68.52% id in 629 aa"
                     /codon_start=1
                     /transl_table=11
                     /product="glucose inhibited division protein A"
                     /protein_id="CAE39742.1"
                     /db_xref="GI:33564800"
                     /db_xref="GOA:Q7W2I1"
                     /db_xref="InterPro:IPR000205"
                     /db_xref="InterPro:IPR002218"
                     /db_xref="InterPro:IPR004416"
                     /db_xref="UniProt/TrEMBL:Q7W2I1"
                     /translation="MDFPREFDVIVVGGGHAGTEAALAAARAGAQTLLLTHNIETLGQ
                     MSCNPSIGGIGKGHLVKEVDALGGAMAIATDEAGIQFRILNSSKGPAVRATRVQADRV
                     LYRNAMRARLENQPNLWLFQQAVDDLMVQGDQVVGAVTQIGLRFRARTVVLTAGTFLN
                     GLIHVGLQNYSGGRAGDPPANSLGQRLKELQLPQGRLKTGTPPRIDGRSINYSVLEEQ
                     PGDLDPVPVFSFLGKASMHPRQLPCWITHTNARTHEIIRGGLDRSPMYSGVIEGVGPR
                     YCPSIEDKIHRFADKASHQVFLEPEGLNTHEIYPNGVSTSLPFDVQYELIHSLPGLEN
                     AHILRPGYAIEYDYFDPRALKSTLETKAISGLFFAGQINGTTGYEEAAAQGLLAGANA
                     ALQAQGKEPWVPRRDEAYLGVLVDDLVTRGVTEPYRMFTSRAEYRLSLREDNADLRLT
                     EIGRRLGLVDDVRWDAFSRKRDAVAQEVERLKSTWVNPRVLPAHAAEALLGKAIEREY
                     SLSDLLKRPNVSYEALMQARTDEGELLAGPGVLEDDVLAEQVETQVKYAGYIARQQDE
                     VQKHLSHEQQPIPADIDYDAVTSLSFEVRQKLKTHRPETIGQAARVSGVTPAAISLLL
                     IHLKRLHYGSRKQAA"
     misc_feature    19. .1893
                     /gene="gidA"
                     /locus_tag="BPP0001"
                     /note="HMMPfam hit to PF01134, Glucose inhibited division
                     protein A, score 0"
     misc_feature    820. .864
                     /gene="gidA"
                     /locus_tag="BPP0001"
                     /note="PS01280 Glucose inhibited division protein A family
                     signature 1."
     gene            1917. .2609
                     /gene="gidB"
                     /locus_tag="BPP0002"
     CDS             1917. .2609
                     /gene="gidB"
                     /locus_tag="BPP0002"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0002
                     Similar to Pseudomonas putida glucose inhibited division
                     protein B GidB SWALL:GIDB_PSEPU (SWALL:P25757) (216 aa)
                     fasta scores: E(): 1.5e-24, 42.59% id in 216 aa, and to
                     Escherichia coli glucose inhibited division protein B GidB
```

```
                     or B3740 or Z5240 or Ecs4682 SWALL:GIDB_ECOLI
                     (SWALL:P17113) (207 aa) fasta scores: E(): 1.6e-12, 39% id
                     in 200 aa"
                     /codon_start=1
                     /transl_table=11
                     /product="glucose inhibited division protein B"
                     /protein_id="CAE39743.1"
                     /db_xref="GI:33564801"
                     /db_xref="GOA:Q7W2I0"
                     /db_xref="InterPro:IPR003682"
                     /db_xref="UniProt/Swiss-Prot:Q7W2I0"
                     /translation="MSAVPDIPGGPAQRLAQACDALRLPADAGQQQKLLRYIEQMQRW
                     NRTYNLTAIRDPGQMLVQHLFDSLSVVAPLERGLPAAGSGARVKLFDVGSGGGLPGVV
                     LAIMRAHWDVTCVDAVEKKTAFVRQMAGALGLPNLQAAHTRIEQLEPAQCDVVISRAF
                     ASLQDFAKLAGRHVREGGTLVAMKGKVPDDEIQALQQHGHWTVERIEPLVVPALDAQR
                     CLIWMRRSQGNI"
misc_feature         1917. .2135
                     /gene="gidB"
                     /locus_tag="BPP0002"
                     /note="PS00430 TonB-dependent receptor proteins signature
                     1."
misc_feature         2004. .2579
                     /gene="gidB"
                     /locus_tag="BPP0002"
                     /note="HMMPfam hit to PF02527, Glucose inhibited division
                     protein, score 1.7e-50"
gene                 2606. .3403
                     /gene="parA"
                     /locus_tag="BPP0003"
                     /note="synonym: soj"
CDS                  2606. .3403
                     /gene="parA"
                     /locus_tag="BPP0003"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0003
                     Similar to Caulobacter crescentus chromosome partitioning
                     protein ParA or Cc3753 SWALL:PARA_CAUCR (SWALL:O05189)
                     (267 aa) fasta scores: E(): 9.6e-46, 49.41% id in 255 aa,
                     and to Bacillus subtilis sporulation initiation inhibitor
                     protein Soj SWALL:SOJ_BACSU (SWALL:P37522) (253 aa) fasta
                     scores: E(): 1.3e-43, 48.98% id in 247 aa"
                     /codon_start=1
                     /transl_table=11
                     /product="ParA family protein"
                     /protein_id="CAE39744.1"
                     /db_xref="GI:33564802"
                     /db_xref="GOA:Q7W2H9"
                     /db_xref="InterPro:IPR002586"
                     /db_xref="UniProt/TrEMBL:Q7W2H9"
                     /translation="MKNIPPSKSARVFCIANQKGGVGKTTTAINLAAGLATHKQRVLL
                     VDLDPQGNATMGSGIDKSTLESNLYQVLIGEAGIEQTRVRSESGGYDVLPANRELSGA
                     EIDLVQMDEREROLKAAIDKIAGEYDFVLIDCPPTLSLLTLNGLAAAHGVIIPMQCEY
                     FALEGLSDLVNTIKRVHRNINNELRVIGLLRVMFDPRMTLQQQVSAQLESHFGDKVFT
                     TVVPRNVRLAEAPSYGMPGVVYDRASRGAQAYIAFGAEMIERVKELD"
misc_feature         2879. .3214
                     /gene="parA"
                     /locus_tag="BPP0003"
                     /note="HMMPfam hit to PF00991, ParA family ATPase, score
                     1.5e-33"
gene                 3410. .4216
                     /locus_tag="BPP0004"
CDS                  3410. .4216
                     /locus_tag="BPP0004"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0004
                     Similar to Caulobacter crescentus acetyltransferase, GnaT
                     family Cc2209 SWALL:Q9A685 (EMBL:AE005892) (207 aa) fasta
                     scores: E(): 1.5e-11, 34.63% id in 205 aa, and to
                     Rhizobium meliloti putative acetyltransferase protein
                     R02624 or Smc02449 SWALL:Q92MJ2 (EMBL:AL591791) (308 aa)
                     fasta scores: E(): 5.5e-10, 31.42% id in 245 aa"
                     /codon_start=1
                     /transl_table=11
                     /product="Putative acetyltransferase, GnaT family"
                     /protein_id="CAE39745.1"
                     /db_xref="GI:33564803"
                     /db_xref="GOA:Q7W2H8"
                     /db_xref="InterPro:IPR000182"
                     /db_xref="UniProt/TrEMBL:Q7W2H8"
                     /translation="MPAVLAVLAAHPAPAAGGELARLEEAALNATTVREQMLYEGWLV
                     RWAPSRAKRARCVNVLGLSTRPLDERLAYCTRWYAGHGVPLIVRLTSIGPDFSLDAQL
                     HERGFVAFDRTCVMAAPVVPFTAQAPSGLVFERTTPGRFAQEAGRLRGHDDAATAEHA
                     ARLQAAPLENLPLLVRDASGECIAAGLAAFDAELVGLFDIVTAPQRRRQGCGAAVLKY
                     LLEQASHAGARQAYLQVEPQNTAARALYARFGFVDRYTYWYRSMSEATPV"
misc_feature         3929. .4165
                     /locus_tag="BPP0004"
                     /note="HMMPfam hit to PF00583, Acetyltransferase (GNAT)
                     family, score 2.5e-14"
gene                 4260. .5177
                     /gene="parB"
                     /locus_tag="BPP0005"
CDS                  4260. .5177
                     /gene="parB"
                     /locus_tag="BPP0005"
                     /note="ortholog of Bordetella pertussis (BX470248) BP0005
                     Similar to Pseudomonas putida probable chromosome
                     partitioning protein ParB SWALL:PARB_PSEPU (SWALL:P31857)
                     (290 aa) fasta scores: E(): 1.3e-40, 48.81% id in 295 aa"
                     /codon_start=1
                     /transl_table=11
                     /product="probable chromosome partitioning protein ParB"
                     /protein_id="CAE39746.1"
                     /db_xref="GI:33564804"
                     /db_xref="GOA:Q7W2H7"
                     /db_xref="InterPro:IPR003115"
                     /db_xref="InterPro:IPR004437"
                     /db_xref="UniProt/TrEMBL:Q7W2H7"
                     /translation="MATKKPKGLGRGLDALLGADVPAIDNIGKAPAAAPEGPPATLPV
                     SKMRAGKYQPRTRMDEGALNELAESIRTQGIMQPILVRALGADAPGQYEIIAGERRFR
                     AAQLAGLKEVPVLVREVADENAAVMALIENIQREDLNPLEEAHGVRRLLDEFGLTHEQ
                     AAQAIGRSRSATSNLLRLLNLAAPVQTMLLAGDVDMGHARALLAVDAATQIQLANQVI
                     AKRLSVREAEKLVARAQKDVDAAPRKKGNGASRDVTRLEEALSDHLGTRVALKVGARE

Alignment Scores:
Pred. No.:              9.4e+03         Length:         348251
Score:                  41.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  77.8%           Mismatches:     0
Query Match:            91.1%           Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BX640423 (1-348251)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||||||||
Db     122425 CCGGACCTGGTCTTCCTGCAGGAGGTC 122451

RESULT 50
BX640445
LOCUS       BX640445              348706 bp    DNA     linear   BCT 17-APR-2005
DEFINITION  Bordetella bronchiseptica strain RB50, complete genome; segment
            9/16.
ACCESSION   BX640445 BX470250
VERSION     BX640445.1  GI:33576035
KEYWORDS    complete genome.
SOURCE      Bordetella bronchiseptica RB50
  ORGANISM  Bordetella bronchiseptica RB50
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Bordetella.
REFERENCE   1
  AUTHORS   Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
            Harris,D.E., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,
            Mungall,K.L., Cerdeno-Tarraga,A.M., Temple,L., James,K., Harris,B.,
            Quail,M.A., Achtman,M., Atkin,R., Baker,S., Basham,D., Bason,N.,
            Cherevach,I., Chillingworth,T., Collins,M., Cronin,A., Davis,P.,
            Doggett,J., Feltwell,T., Goble,A., Hamlin,N., Hauser,H.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Norberczak,H.,
            O'Neil,S., Ormond,D., Price,C., Rabbinowitsch,E., Rutter,S.,
```

```
            Sanders,M., Saunders,D., Seeger,K., Sharp,S., Simmonds,M.,
            Skelton,J., Squares,R., Squares,S., Stevens,K., Unwin,L.,
            Whitehead,S., Barrell,B.G. and Maskell,D.J.
  TITLE     Comparative analysis of the genome sequences of Bordetella
            pertussis, Bordetella parapertussis and Bordetella bronchiseptica
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 348706)
  AUTHORS   Sebaihia,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-2003) Submitted on behalf of the Pathogen
            Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA, E-mail: ms5@sanger.ac.uk
FEATURES             Location/Qualifiers
     source          1. .348706
                     /organism="Bordetella bronchiseptica RB50"
                     /mol_type="genomic DNA"
                     /strain="RB50"
                     /db_xref="taxon:257310"
     gene            160. .2523
                     /gene="ppsA"
                     /locus_tag="BB2622"
                     /note="synonym: pps"
     CDS             160. .2523
                     /gene="ppsA"
                     /locus_tag="BB2622"
                     /EC_number="2.7.9.2"
                     /note="ortholog of Bordetella pertussis (BX470248) BP1436"
                     /codon_start=1
                     /transl_table=11
                     /product="phosphoenolpyruvate synthase"
                     /protein_id="CAE33115.1"
                     /db_xref="GI:33576036"
                     /db_xref="GOA:Q7WJ77"
                     /db_xref="InterPro:IPR000121"
                     /db_xref="InterPro:IPR002192"
                     /db_xref="InterPro:IPR006319"
                     /db_xref="InterPro:IPR008279"
                     /db_xref="UniProt/TrEMBL:Q7WJ77"
                     /translation="MSYVVSFEQLRMTDVDSVGGKNASLGEMISQLAGAGVRVPGGFA
                     TTAQAFRDFLQASGLDKRIADRLATLNPEDVRELASAGAEIRQWVIDAPFSPEFEQAI
                     RTAFAELDADGKGSFAVRSSATAEDLPDASFAGQQETFLNVVGVDDVLDKIRHVFASL
                     YNDRAISYRVHKGYAHAEVALSAGIQRMVRSDKGSAGVMFTIDTESGFQDVVFITSSY
                     GLGETVVQGAVNPDEFYVFKPSLERGNYPIVSRRIGSKLIKMEFDPERPEGRAVRTVD
                     VPVSERNRYSLTDDEVIELARYAVIIEKHYQRPMDIEWGRDGVDGKIYILQARPETVK
                     SQQGVNDVQQRYRLKATGQVLVTGRAIGQKIGAGKVRVVADISEMDKVQAGDVLVTDM
                     TDPNWEPVMKRASAIITNRGGRTCHAAIIARELGIPAVVGCGDATDLLKDGQAVTASC
                     AEGDEGRIYDGLIETEVEEVRRGEMPPIDVKIMMNVGNPQLAFDFAQIPNAGVGLARL
                     EFIINNNIGIHPKAVLDYPNVDGELKKAVESAARGHASPRAFFVEKLAEGVATIGAAF
                     WPKPVIVRMSDFKSNEYRKLVGGSRYEPEEENPMLGFRGASRYIAADFEECFRMECEA
                     LKKVRDEMGLTNVEIMVPFVRTLGQARKVVELLAAHGLARGENGLKLIMMCEVPSNAI
                     LAEQFLEYFDGFSIGSNDMTQLTLGLDRDSGMELLAADFDERDDAVKFMLRRAIKACL
                     AANKYVGICGQGPSDHPDFAQWLKDEGILSMSLNPDTVVETWQQLAK"
     misc_feature    163. .1245
                     /gene="ppsA"
                     /locus_tag="BB2622"
                     /note="HMMPfam hit to PF01326, DE Pyruvate phosphate
                     dikinase, PEP/pyruvate binding domain, score 1.9e-129"
     misc_feature    1246. .1527
                     /gene="ppsA"
                     /locus_tag="BB2622"
                     /note="HMMPfam hit to PF00391, DE PEP-utilizing enzyme,
                     mobile domain, score 3.4e-40"
     misc_feature    1393. .1428
                     /gene="ppsA"
                     /locus_tag="BB2622"
                     /note="PS00370 PEP-utilizing enzymes phosphorylation site
                     signature."
     misc_feature    1549. .2517
                     /gene="ppsA"
                     /locus_tag="BB2622"
                     /note="This hit extended beyond the end of the feature by
                     1 aa and was clipped.
                     HMMPfam hit to PF02896, DE PEP-utilizing enzyme, TIM
                     barrel domain, score 7e-104"
     misc_feature    2233. .2289
                     /gene="ppsA"
                     /locus_tag="BB2622"
                     /note="PS00742 PEP-utilizing enzymes signature 2."
     gene            complement(2591. .3310)
                     /locus_tag="BB2623"
     CDS             complement(2591. .3310)
                     /locus_tag="BB2623"
                     /note="ortholog of Bordetella pertussis (BX470248) BP1437"
                     /codon_start=1
                     /transl_table=11
                     /product="probable class-I glutamine amidotransferase)"
                     /protein_id="CAE33116.1"
                     /db_xref="GI:33576037"
                     /db_xref="GOA:Q7U3A5"
                     /db_xref="InterPro:IPR000991"
                     /db_xref="UniProt/TrEMBL:Q7U3A5"
                     /translation="MTAALPVLILHTGNPAAALRAAHGSYAEMLRQAAGLDAAGTHVV
                     ASYKDEAPRAPGAYRAALITGSPAMVTDRAPWSEAAADWLRQAAGAGLPMFGICYGHQ
                     LLAHALGGVVGDNPAGRELGTLPVELLPAAATDPLLAGLPAAFDAQMMHEQAVLAAPP
                     GAAVLARSAQDAHQMLRLAPRIYTAQFHPEFTPGFVAAHLRHHAALYAAHGLDVERLA
                     RAAGPTPLAASLVRRFLACCA"
     misc_feature    complement(2693. .3259)
                     /locus_tag="BB2623"
                     /note="HMMPfam hit to PF00117, DE Glutamine
                     amidotransferase class-I, score 3.2e-09"
     misc_feature    complement(2999. .3034)
                     /locus_tag="BB2623"
                     /note="PS00442 Glutamine amidotransferases class-I active
                     site."
     gene            3434. .3859
                     /locus_tag="BB2624"
     CDS             3434. .3859
                     /locus_tag="BB2624"
                     /note="ortholog of Bordetella pertussis (BX470248) BP1438"
                     /codon_start=1
                     /transl_table=11
                     /product="putative membrane protein"
                     /protein_id="CAE33117.1"
                     /db_xref="GI:33576038"
                     /db_xref="UniProt/TrEMBL:Q7WJ76"
                     /translation="MWIWFGLAAVALIGEVGSGTFYLLLVALGLAAGGIAAWLGADLA
                     WQLVACGLVALLGLLALRRTGVLKKREVNAARNADVNLDIGQSVMVDAWSDAGTTQVY
                     YRGAQWQAELAHGQPRQAGRHIIAEIRGTRLVLAPAPAA"
     misc_feature    3434. .3526
                     /locus_tag="BB2624"
                     /note="Signal peptide predicted for BB2624 by SignalP 2.0
                     HMM (Signal peptide probabilty 0.622) with cleavage site
                     probability 0.235 between residues 31 and 32"
     misc_feature    join(3452. .3520,3563. .3616)
                     /locus_tag="BB2624"
                     /note="2 probable transmembrane helices predicted for
                     BB2624 by TMHMM2.0 at aa 7-29 and 44-61"
     gene            3913. .4839
                     /locus_tag="BB2625"
     CDS             3913. .4839
                     /locus_tag="BB2625"
                     /note="ortholog of Bordetella pertussis (BX470248) BP1440"
                     /codon_start=1
                     /transl_table=11
                     /product="putative membrane protein"
                     /protein_id="CAE33118.1"
                     /db_xref="GI:33576039"
                     /db_xref="GOA:Q7WJ75"
                     /db_xref="InterPro:IPR001107"
                     /db_xref="UniProt/TrEMBL:Q7WJ75"
                     /translation="MIDVSTVVLIVIVILALMIVVKAIAIVPQQHAWVVERLGKFDRV
                     LSPGAGFVIPFIERVSYKHSLKEIPLDVPSQVCITRDNTQLQVDGVLYFQVTDPMRAS
                     YGSSNYISAITQLAQTTLRSVIGKLELDRTFEEREFINSTIVASLDEAALNWGVKVLR
                     YEIKDLTPPNEILRAMQAQITAEREKRALIAASEGRRQEQINIATGEREAAIARSEGE
                     KQAQINQAQGEAAAVLAIAEATAKAIEQVGEAVRQPGGMEAVNLKVAERYVDAFSNVA
                     KEGNTLILPSNLSDVGGLIASAMTIVKSTKSA"
```

```
     misc_feature    3913. .3987
                     /locus_tag="BB2625"
                     /note="Signal peptide predicted for BB2625 by SignalP 2.0
                     HMM (Signal peptide probabilty 0.709) with cleavage site
                     probability 0.611 between residues 25 and 26"
     misc_feature    3931. .3999
                     /locus_tag="BB2625"
                     /note="1 probable transmembrane helix predicted for BB2625
                     by TMHMM2.0 at aa 7-29"
     misc_feature    3979. .4503
                     /locus_tag="BB2625"
                     /note="HMMPfam hit to PF01145, DE SPFH domain / Band 7
                     family, score 1.1e-70"
     gene            4905. .6062
                     /locus_tag="BB2626"
     CDS             4905. .6062
                     /locus_tag="BB2626"
                     /note="ortholog of Bordetella pertussis (BX470248) BP1441"
                     /codon_start=1
                     /transl_table=11
                     /product="probable membrane protein"
                     /protein_id="CAE33119.1"
                     /db_xref="GI:33576040"
                     /db_xref="GOA:Q7WJ74"
                     /db_xref="InterPro:IPR000160"
                     /db_xref="UniProt/TrEMBL:Q7WJ74"
                     /translation="MLAPVTLLVISTLSSLVTLGVLGSLLRSGIAGIRELMWANVLVL
                     CSLLLFTLQANPAMPPWLAIQAPNVLISSGVVLFCAGVFRFMQQRPPWRLLLLGLALS
                     IGANIWFHYVDPSVNARVVAASGLHALLWGGSAWTIYRNMPLRRSRYSYWFAGVTATV
                     AALGHALRTAVYGLQIEQTQGLLQSSVWNVAFLAIGVLVMPSMTLGMIMMIHDRMLAE
                     RERERANQDFLTGLLSRKAWWREGERLCARALRTGRPLTLLALDIDRFKQVNDQHGHAA
                     GDVVLRHFGSLATALLRSGDVIGRLGGEEFVALLPDTDGDTGVSVAERLLASVRATPC
                     SHGGKALTYTFSGGVAQYRAGDTLPVLVERADAALYAAKQAGRDRIRNRAG"
     misc_feature    4905. .4973
                     /locus_tag="BB2626"

Alignment Scores:
Pred. No.:                9.41e+03        Length:         348706
Score:                    41.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    77.8%           Mismatches:     0
Query Match:              91.1%           Indels:         0
DB:                       1               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BX640445 (1-348706)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||::::::|||||||||||||||
Db     215820 CCGGATCTCGTGTTCCTGCAGGAAGTG 215846

RESULT 51
BX640413
LOCUS       BX640413               349028 bp    DNA     linear   BCT 17-APR-2005
DEFINITION  Bordetella pertussis strain Tohama I, complete genome; segment
            3/12.
ACCESSION   BX640413 BX470248
VERSION     BX640413.1  GI:33571514
KEYWORDS    complete genome.
SOURCE      Bordetella pertussis Tohama I
  ORGANISM  Bordetella pertussis Tohama I
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Bordetella.
REFERENCE   1
  AUTHORS   Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
            Harris,D.E., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,
            Mungall,K.L., Cerdeno-Tarraga,A.M., Temple,L., James,K., Harris,B.,
            Quail,M.A., Achtman,M., Atkin,R., Baker,S., Basham,D., Bason,N.,
            Cherevach,I., Chillingworth,T., Collins,M., Cronin,A., Davis,P.,
            Doggett,J., Feltwell,T., Goble,A., Hamlin,N., Hauser,H.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Norberczak,H.,
            O'Neil,S., Ormond,D., Price,C., Rabbinowitsch,E., Rutter,S.,
            Sanders,M., Saunders,D., Seeger,K., Sharp,S., Simmonds,M.,
            Skelton,J., Squares,R., Squares,S., Stevens,K., Unwin,L.,
            Whitehead,S., Barrell,B.G. and Maskell,D.J.
  TITLE     Comparative analysis of the genome sequences of Bordetella
            pertussis, Bordetella parapertussis and Bordetella bronchiseptica
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 349028)
  AUTHORS   Sebaihia,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-2003) Submitted on behalf of the Pathogen
            Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA, E-mail: ms5@sanger.ac.uk
FEATURES             Location/Qualifiers
     source          1. .349028
                     /organism="Bordetella pertussis Tohama I"
                     /mol_type="genomic DNA"
                     /strain="Tohama I"
                     /db_xref="taxon:257313"
     gene            86. .1168
                     /gene="prfA"
                     /locus_tag="BP0678"
                     /note="synonyms: sueB, uar"
     CDS             86. .1168
                     /gene="prfA"
                     /locus_tag="BP0678"
                     /note="Similar to Escherichia coli peptide chain release
                     factor 1 PrfA or SueB or Uar or B1211 SW:RF1_ECOLI
                     (P07011) (360 aa) fasta scores: E(): 9.8e-89, 65.922% id
                     in 358 aa, and to Neisseria meningitidis peptide chain
                     release factor 1 Nmb1686 TR:Q9JY93 (EMBL:AE002519) (358
                     aa) fasta scores: E(): 2e-96, 71.989% id in 357 aa. Also
                     similar to BP1104, 40.571% identity in 350 aa overlap."
                     /codon_start=1
                     /transl_table=11
                     /product="peptide chain release factor 1"
                     /protein_id="CAE40989.1"
                     /db_xref="GI:33571515"
                     /db_xref="GOA:Q7W023"
                     /db_xref="InterPro:IPR000352"
                     /db_xref="InterPro:IPR004373"
                     /db_xref="InterPro:IPR005139"
                     /db_xref="UniProt/Swiss-Prot:Q7W023"
                     /translation="MKPSMRSRLEQLAHRLIEVDALLAEPETAADMDRFRKLSRERAE
                     LEPVVEAFNAFLGVEADVATAQEMLSDPDMKAMAEDEIKTGRARIEEMEAALQLLLLP
                     RDPDDGRSLFLEIRAGTGGDESALFSGDLLRMYTRYAETRGWRVEIMSESESELGGYK
                     EVIARIDGDGAYGRLKFESGAHRVQRVPATEAQGRIHTSACTVAVMPEADAMSDIVIN
                     PSDLRIDTFRASGAGGQHINKTDSAVRITHVPTGLVVECQDDRSQHRNKDKAMQVLAA
                     RLKDKEMRERQSKEAAERKSLIGSGDRSERIRTYNYPQGRVTDHRINLTLYKLQQIME
                     GDLDELTGALLAEHQAEQLAALGHDL"
     misc_feature    707. .1048
                     /gene="prfA"
                     /locus_tag="BP0678"
                     /note="HMMPfam hit to PF00472, Peptidyl-tRNA hydrolase
                     domain"
     gene            1179. .1991
                     /gene="hemK"
                     /locus_tag="BP0679"
     CDS             1179. .1991
                     /gene="hemK"
                     /locus_tag="BP0679"
                     /EC_number="2.1.1.-"
                     /note="Similar to Escherichia coli HemK protein or B1212
                     SW:HEMK_ECOLI (P37186) (277 aa) fasta scores: E():
                     1.2e-32, 42.578% id in 256 aa, and to Neisseria
                     meningitidis HemK protein or Nma0369 TR:Q9JWH6
                     (EMBL:AL162753) (273 aa) fasta scores: E(): 7.5e-41,
                     48.846% id in 260 aa"
                     /codon_start=1
                     /transl_table=11
                     /product="heme biosynthesis protein"
                     /protein_id="CAE40990.1"
                     /db_xref="GI:33571516"
                     /db_xref="GOA:Q7W022"
                     /db_xref="InterPro:IPR000051"
                     /db_xref="InterPro:IPR002052"
```

```
                     /db_xref="InterPro:IPR002345"
                     /db_xref="InterPro:IPR004556"
                     /db_xref="UniProt/TrEMBL:Q7W022"
                     /translation="MTLLKDLLADPGLPRLEARMLAEHVLGRSRAWLLAHDTDPVEPA
                     HEAAWRQLAARRLAGEPMAYLLGGREFMGHWYALTPDVLIPRPDTELLVETALHWLQG
                     RAAPRVLDLGTGSGAIAVSVALGCPQAEVTATDLSAAALAVAEGNAQRLGARVRCLAG
                     DWYEALPAQDRYDLIVSNPPYIAREDAHLAQGDLRFEPRGALTDENDGLAALARIAGG
                     APGRLLPGGAIWMEHGWDQAEAARALLRQAGLREVHSRRDLAGIERISGGYL"
misc_feature         1377. .1412
                     /gene="hemK"
                     /locus_tag="BP0679"
                     /note="ScanRegExp hit to PS00213, Lipocalin signature."
misc_feature         1701. .1721
                     /gene="hemK"
                     /locus_tag="BP0679"
                     /note="ScanRegExp hit to PS00092, N-6 Adenine-specific DNA
                     methylases signature."
gene                 2028. .2354
                     /locus_tag="BP0680"
CDS                  2028. .2354
                     /locus_tag="BP0680"
                     /note="Similar to Neisseria meningitidis hypothetical
                     protein Nma0984 or Nmb0773 TR:Q9JR73 (EMBL:AL162754) (103
                     aa) fasta scores: E(): 2.8e-24, 66.337% id in 101 aa, and
                     to Rhizobium meliloti conserved hypothetical protein
                     Smc00538 TR:CAC46360 (EMBL:AL591788) (111 aa) fasta
                     scores: E(): 9.6e-21, 57.009% id in 107 aa"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="CAE40991.1"
                     /db_xref="GI:33571517"
                     /db_xref="GOA:Q7W021"
                     /db_xref="InterPro:IPR002109"
                     /db_xref="InterPro:IPR004480"
                     /db_xref="UniProt/TrEMBL:Q7W021"
                     /translation="MSDVQEFIRETVTQHPVVLFMKGTTQFPQCGFSGRAIQILKGCG
                     VKKLVTVNVLEDDEVRQAIKTFSNWPTIPQLYVGGEFIGGSDIMSEMNESGELKQLLD
                     DSGATA"
gene                 2359. .2919
                     /locus_tag="BP0681"
CDS                  2359. .2919
                     /locus_tag="BP0681"
                     /note="Similar to Streptomyces sp VdcB TR:Q9X696
                     (EMBL:AF134589) (200 aa) fasta scores: E(): 4.8e-33,
                     51.892% id in 185 aa, and to Escherichia coli
                     phenylacrylic acid decarboxylase-like protein Pad1
                     TR:Q9F8R0 (EMBL:AF242211) (197 aa) fasta scores: E():
                     1.5e-31, 49.189% id in 185 aa"
                     /codon_start=1
                     /transl_table=11
                     /product="probable flavoprotein"
                     /protein_id="CAE40992.1"
                     /db_xref="GI:33571518"
                     /db_xref="GOA:Q7W020"
                     /db_xref="InterPro:IPR003382"
                     /db_xref="InterPro:IPR004507"
                     /db_xref="UniProt/TrEMBL:Q7W020"
                     /translation="MRRLVVGITGATGAIYAVRLLQALRGVSDVESHLIVSASGVLNI
                     KHELDVGRHDVQALADHVHSVRDVGATLASGAFQTAGMVIVPCSMRTLAAVAHGLSDN
                     LITRAADVTLKERRRLVMMVRETPFNLAHLRNMTAVTEMGGIIFPPLPAFYHRPASIE
                     EMVDHTVARVLELFDIEVPGPHWAGM"
misc_feature         2362. .2727
                     /locus_tag="BP0681"
                     /note="HMMPfam hit to PF02441, Flavoprotein"
gene                 3160. .4152
                     /locus_tag="BP0682"
CDS                  3160. .4152
                     /locus_tag="BP0682"
                     /note="Similar to Variovorax paradoxus 2,4-diphenoxyacetic
                     acid gene cluster TR:Q9RHQ7 (EMBL:AB028643) (337 aa) fasta
                     scores: E(): 1.4e-38, 36.577% id in 298 aa, and to
                     Rhizobium meliloti conserved hypothetical protein Sma1927
                     TR:AAK65716 (EMBL:AE007292) (330 aa) fasta scores: E():
                     2.3e-42, 38.769% id in 325 aa"
                     /codon_start=1
                     /transl_table=11
                     /product="putative exported protein"
                     /protein_id="CAE40993.1"
                     /db_xref="GI:33571519"
                     /db_xref="GOA:Q7W019"
                     /db_xref="InterPro:IPR005064"
                     /db_xref="InterPro:IPR006311"
                     /db_xref="UniProt/TrEMBL:Q7W019"
                     /translation="MQRRSFLAKVAHIAAAVGVSACAATSSFAATGDFPNKPLDIIVT
                     FPPGGGTDMLARLIGNYLTESLGQTAVVENRPGASGNVGARLVADRAPDGYSLLMVNS
                     SFAVNPGVFRNLPFDPKKDFAAVINVAYVPSVFVVPAGSKYKTLGELMAAAKQTNTQV
                     TYGSCGNGTPQHLAGELLNVSAKTHMVHVPYKGCGPALNDVLGSQIGLAVVTASSAIP
                     FIKAGKLQALAVTSKERSALLPEVPTVAEQGVAGYELNQWHGLLVPGATPMAVRQKLY
                     DGIAKVMQRDDVQKKLADLGYSTASDGPEVFQKMVETDIDRFSALTKQIGLKVD"
misc_feature         3160. .3246
                     /locus_tag="BP0682"
                     /note="Signal peptide predicted for BP0682 by SignalP 2.0
                     HMM (Signal peptide probabilty 1.000) with cleavage site
                     probability 0.982 between residues 29 and 30"
misc_feature         3175. .3243
                     /locus_tag="BP0682"
                     /note="1 probable transmembrane helix predicted for BP0682
                     by TMHMM2.0 at aa 13-35"
misc_feature         3193. .3225
                     /locus_tag="BP0682"
                     /note="PS00013 Prokaryotic membrane lipoprotein lipid
                     attachment site."
gene                 4171. .5166
                     /locus_tag="BP0683"
CDS                  4171. .5166

Alignment Scores:
Pred. No.:                9.42e+03        Length:          349028
Score:                    41.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    2
Best Local Similarity:    77.8%           Mismatches:      0
Query Match:              91.1%           Indels:          0
DB:                       1               Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BX640413 (1-349028)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||:::::|||||||||||||||||
Db      30998 CCGGACCTGGTCTTCCTGCAGGAGGTC 31024

RESULT 52
AP000599/c
LOCUS       AP000599               18284 bp    DNA     linear   PLN 14-FEB-2004
DEFINITION  Arabidopsis thaliana genomic DNA, chromosome 3, TAC clone: K9I22.
ACCESSION   AP000599 BA000014
VERSION     AP000599.1  GI:6045154
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.
  TITLE     Structural analysis of Arabidopsis thaliana chromosome 3. II.
            Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
            and BAC clones
  JOURNAL   DNA Res. 7 (3), 217-221 (2000)
   PUBMED   10907853
REFERENCE   2  (bases 1 to 18284)
  AUTHORS   Kaneko,T., Kato,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-1999) Yasukazu Nakamura, Kazusa DNA Research
            Institute, Department of Plant Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
```

```
            Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT     Address for correspondence: kaos@kazusa.or.jp
            For the latest information on annotation of this clone, please see
            http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=K9I22
            Genes with similarity to proteins in the databases are described in
            'product' or 'note' qualifiers.  Genes that have no significant
            protein similarity are described as 'unknown protein'.
            The software programs used to predict genes include: Grail
            (Informatics Group, Oak Ridge National Laboratory,
            http://compbio.ornl.gov/Grail-1.3/),
            GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
            NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
            Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
            SplicePredictor (Volker Brendel, Stanford University,
            http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).
            Genes encoding tRNAs are predicted by tRNAscan-SE
            (Sean Eddy, Washington University School of Medicine, St. Louis,
            http://genome.wustl.edu/eddy/tRNAscan-SE/).
            This sequence may not be the entire insert of this clone. It may be
            shorter because we remove overlaps between neighboring submissions.
            The 5' clone is K13N2 and the 3' clone is MPE11.
FEATURES                Location/Qualifiers
     source             1. .18284
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:3702"
                        /chromosome="3"
                        /clone="K9I22"
                        /clone_lib="Mitsui TAC"
                        /ecotype="Columbia"
     CDS                complement(join(3798. .3913,4386. .4801,5077. .5897))
                        /note="gene_id:K9I22.2"
                        /pseudo
                        /codon_start=1
                        /evidence=not_experimental
                        /product="non-LTR retroelement reverse transcriptase-like
                        protein"
     CDS                join(6758. .6967,7093. .7369,7466. .7847,7941. .8159,
                        8625. .8763,8836. .9087,9206. .9529)
                        /note="gene_id:K9I22.3"
                        /codon_start=1
                        /evidence=not_experimental
                        /product="limonene cyclase"
                        /protein_id="BAB01181.1"
                        /db_xref="GI:11994152"
                        /translation="MATLRISSALIYQNTLTHHFRLRRPHRFVCKSMTKTTPDTTLVE
                        LSRRSGNYQPSPWNHCYLLSIENKYASETEVITRDVLKKKVKSMLDDEKKSRLEQLEL
                        IDDLQKLGVSYHFEIEINDTLTDLHLKMGRNCWKCDKEEDLHATSLEFRLLRQHGFDV
                        SENIFDVIIDQIESNTFKTNNINGIISLYEASYLSTKSDTKLHKVIRPFATEQIRKFV
                        DDEDTKNIEVREKAYHALEMPYHWRMRRLDTRWYIDAYEKKHDMNLVLIEFAKIDFNI
                        VQAAHQEDLKYVSRWWKDTCLTNQLPFVRDRIVENYFWTVGLIYEPQFGYIRRIMTIV
                        NALVTTIDDIYDIYGTLEELELFTSMVENWDVNRLGELPEYMRLCFLILYNEINGIGC
                        DILKYKKIDVIPYLKKSWADLCRTYLVEAKWYKRGYKPSLEEYMQNAWISISAPTILI
                        HFYCVFSDQISVQNLETLSQHRQHIVRCSATVLRLANDLGTSPTELARGDVLKSVQCY
                        MHETGASEERARDHVHQMISDMWDDMNSETKTACNSSSRSRGFKEAAMNLARMSQCMY
                        QYGDGHGCPEKAKTIDRVQSLLVDPIPLDVNRLG"
     CDS                complement(join(12206. .12394,12483. .12557,12732. .12895,
                        12979. .13093,13169. .13460,13561. .13717,13814. .13885,
                        14093. .14134,14308. .14502,14763. .14950,15045. .15964,
                        16303. .16659))
                        /note="gene_id:K9I22.4"
                        /codon_start=1
                        /evidence=not_experimental
                        /product="protein kinase-like protein"
                        /protein_id="BAB01182.1"
                        /db_xref="GI:11994153"
                        /translation="MANDKQIESHHRKHHRSSSSSDDTEKASKRHKHRHHKHHHHRHR
                        HHRDKKRENEIPSAGDETEILDVTPAAPIVVSNGCEEEDVEEGEILEEDGIGDVLMKT
                        ADSDGESGEIKFEDNNLPPLGEKGRQGEEKSSNGVLTRESEREDKRWDKEAGGPSERV
                        SKLSYDNGRSSFSPSNSRQSNEGRARSRSKSHDRERETSRSRFAEDEFSDRGRHHDSS
                        RDYHHDRVDSVRTEERYHRRGRYEENDRQYSREVLERDRSKERDMDREGSIRDRDSEG
                        SKRRDKDSDRRRERERERKRREIESDRERRKEKERERSIDRDRRKEREGDYLRDRANER
                        GRSRDRTRYNSRERKREKEREGEKDWERGRETQKDREKYISSDVDYGEMRHKQSRNTR
                        YDAVDDLEMRKPNSLKAHNSKGDKLEETWANDERSRNEDGQDDNDEGLTWKSPEEEEE
                        ELLNRIKEESRKRMEAILEKHKRKPEQQNELLTQDNGKDIVPETGAPVSTSPAVVIAA
                        NVGQAKTNLDFDTVAAKAPLIAGGPPTMSGISDSEKNQAQAGLGEADMFHDDIFGESP
                        AGKGDGVPMVRSGLHDNWDDAEGYYSYQFGELLDGRYEVIATHGKGVFSTVVRAKDLK
                        AGPAEPEEVAIKIIRNNETMHKAGKIEVQILKKLAGADREDRRHCVRFLSSFKYRNHL
                        CLVFESLHLNLREVLKKFGRNIGLQLSAVRAYSKQLFIALKHLKNCGVLHCDIKPDNM
                        LVNEGKNVLKLCDFGNAMFAGKNEVTPYLVSRFYRSPEIILGLTYDHPLDIWSVGCCL
                        YELYSGKVLFPGATNNDMLRLHMELKGPFPKKMLRKGAFIDQHFDHDLNFYATEEDTV
                        SGKLIKRMIVNVKPKDFGSIIKGYPGEDPKILAHFRDLLDKMFILDPERRLTVSQALA
                        HPFITGK"
ORIGIN

Alignment Scores:
Pred. No.:                719             Length:         18284
Score:                    40.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   3
Best Local Similarity:    66.7%           Mismatches:     0
Query Match:              88.9%           Indels:         0
DB:                       15              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AP000599 (1-18284)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||::::::
Db      11792 CCTGATGTTGTTTTTCTCCAGGACATC 11766

RESULT 53
AC090334/c
LOCUS       AC090334               61555 bp    DNA     linear   HTG 17-FEB-2001
DEFINITION  Homo sapiens chromosome 18 clone RP11-863M17 map 18, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC090334
VERSION     AC090334.1  GI:12957970
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 61555)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 18, clone RP11-863M17
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 61555)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
            Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
            Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
            Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
            Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
            Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ---------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
```

Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
---------- Project Information
Center project name: L12607
Center clone name: 863_M_17
----------.
* NOTE: This record contains 75 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 739: contig of 739 bp in length
* 740 839: gap of 100 bp
* 840 1558: contig of 719 bp in length
* 1559 1658: gap of 100 bp
* 1659 2380: contig of 722 bp in length
* 2381 2480: gap of 100 bp
* 2481 3195: contig of 715 bp in length
* 3196 3295: gap of 100 bp
* 3296 4008: contig of 713 bp in length
* 4009 4108: gap of 100 bp
* 4109 4834: contig of 726 bp in length
* 4835 4934: gap of 100 bp
* 4935 5658: contig of 724 bp in length
* 5659 5758: gap of 100 bp
* 5759 6474: contig of 716 bp in length
* 6475 6574: gap of 100 bp
* 6575 7293: contig of 719 bp in length
* 7294 7393: gap of 100 bp
* 7394 8129: contig of 736 bp in length
* 8130 8229: gap of 100 bp
* 8230 8962: contig of 733 bp in length
* 8963 9062: gap of 100 bp
* 9063 9777: contig of 715 bp in length
* 9778 9877: gap of 100 bp
* 9878 10593: contig of 716 bp in length
* 10594 10693: gap of 100 bp
* 10694 11426: contig of 733 bp in length
* 11427 11526: gap of 100 bp
* 11527 12244: contig of 718 bp in length
* 12245 12344: gap of 100 bp
* 12345 13070: contig of 726 bp in length
* 13071 13170: gap of 100 bp
* 13171 13888: contig of 718 bp in length
* 13889 13988: gap of 100 bp
* 13989 14717: contig of 729 bp in length
* 14718 14817: gap of 100 bp
* 14818 15537: contig of 720 bp in length
* 15538 15637: gap of 100 bp
* 15638 16363: contig of 726 bp in length
* 16364 16463: gap of 100 bp
* 16464 17195: contig of 732 bp in length
* 17196 17295: gap of 100 bp
* 17296 18016: contig of 721 bp in length
* 18017 18116: gap of 100 bp
* 18117 18815: contig of 699 bp in length
* 18816 18915: gap of 100 bp
* 18916 19640: contig of 725 bp in length
* 19641 19740: gap of 100 bp
* 19741 20466: contig of 726 bp in length
* 20467 20566: gap of 100 bp
* 20567 21289: contig of 723 bp in length
* 21290 21389: gap of 100 bp
* 21390 22117: contig of 728 bp in length
* 22118 22217: gap of 100 bp
* 22218 22933: contig of 716 bp in length
* 22934 23033: gap of 100 bp
* 23034 23752: contig of 719 bp in length
* 23753 23852: gap of 100 bp
* 23853 24582: contig of 730 bp in length
* 24583 24682: gap of 100 bp
* 24683 25402: contig of 720 bp in length
* 25403 25502: gap of 100 bp
* 25503 26234: contig of 732 bp in length
* 26235 26334: gap of 100 bp
* 26335 27052: contig of 718 bp in length
* 27053 27152: gap of 100 bp
* 27153 27879: contig of 727 bp in length
* 27880 27979: gap of 100 bp
* 27980 28697: contig of 718 bp in length
* 28698 28797: gap of 100 bp
* 28798 29520: contig of 723 bp in length
* 29521 29620: gap of 100 bp
* 29621 30349: contig of 729 bp in length
* 30350 30449: gap of 100 bp
* 30450 31161: contig of 712 bp in length
* 31162 31261: gap of 100 bp
* 31262 31982: contig of 721 bp in length
* 31983 32082: gap of 100 bp
* 32083 32816: contig of 734 bp in length
* 32817 32916: gap of 100 bp
* 32917 33632: contig of 716 bp in length
* 33633 33732: gap of 100 bp
* 33733 34441: contig of 709 bp in length
* 34442 34541: gap of 100 bp
* 34542 35239: contig of 698 bp in length
* 35240 35339: gap of 100 bp
* 35340 36054: contig of 715 bp in length
* 36055 36154: gap of 100 bp
* 36155 36881: contig of 727 bp in length
* 36882 36981: gap of 100 bp
* 36982 37705: contig of 724 bp in length
* 37706 37805: gap of 100 bp
* 37806 38536: contig of 731 bp in length
* 38537 38636: gap of 100 bp
* 38637 39364: contig of 728 bp in length
* 39365 39464: gap of 100 bp
* 39465 40193: contig of 729 bp in length
* 40194 40293: gap of 100 bp
* 40294 41009: contig of 716 bp in length
* 41010 41109: gap of 100 bp
* 41110 41823: contig of 714 bp in length
* 41824 41923: gap of 100 bp
* 41924 42641: contig of 718 bp in length
* 42642 42741: gap of 100 bp
* 42742 43466: contig of 725 bp in length
* 43467 43566: gap of 100 bp
* 43567 44284: contig of 718 bp in length
* 44285 44384: gap of 100 bp
* 44385 45122: contig of 738 bp in length
* 45123 45222: gap of 100 bp
* 45223 45953: contig of 731 bp in length
* 45954 46053: gap of 100 bp
* 46054 46772: contig of 719 bp in length
* 46773 46872: gap of 100 bp
* 46873 47595: contig of 723 bp in length
* 47596 47695: gap of 100 bp
* 47696 48419: contig of 724 bp in length
* 48420 48519: gap of 100 bp
* 48520 49198: contig of 679 bp in length
* 49199 49298: gap of 100 bp
* 49299 50034: contig of 736 bp in length
* 50035 50134: gap of 100 bp
* 50135 50855: contig of 721 bp in length
* 50856 50955: gap of 100 bp
* 50956 51684: contig of 729 bp in length
* 51685 51784: gap of 100 bp
* 51785 52490: contig of 706 bp in length
* 52491 52590: gap of 100 bp

```
            *    52591    53307: contig of 717 bp in length
            *    53308    53407: gap of 100 bp
            *    53408    54127: contig of 720 bp in length
            *    54128    54227: gap of 100 bp
            *    54228    54958: contig of 731 bp in length
            *    54959    55058: gap of 100 bp
            *    55059    55796: contig of 738 bp in length
            *    55797    55896: gap of 100 bp
            *    55897    56612: contig of 716 bp in length

Alignment Scores:
Pred. No.:              2.58e+03        Length:         61555
Score:                  40.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  77.8%           Mismatches:     0
Query Match:            88.9%           Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AC090334 (1-61555)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||:::|||:::
Db       1018 CCAGATGTTATTTTTCTCAGGGAAATT 992

RESULT 54
AL663077
LOCUS       AL663077               96128 bp    DNA     linear   ROD 07-JUN-2002
DEFINITION  Mouse DNA sequence from clone RP23-22K4 on chromosome 2, complete
            sequence.
ACCESSION   AL663077
VERSION     AL663077.8  GI:21387219
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 96128)
  AUTHORS   Kay,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 10, 2002 this sequence version replaced gi:21214134.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-22K4 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .96128
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
                     /clone="RP23-22K4"
                     /clone_lib="RPCI-23"
ORIGIN

Alignment Scores:
Pred. No.:              4.11e+03        Length:         96128
Score:                  40.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  87.5%           Mismatches:     0
Query Match:            88.9%           Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AL663077 (1-96128)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||:::|||||||||||
Db      34742 CCCGATGTTGTGTTCCTGCAGGAA 34765

RESULT 55
BX545850/c
LOCUS       BX545850              137836 bp    DNA     linear   VRT 04-SEP-2004
DEFINITION  Zebrafish DNA sequence from clone RP71-14F13, complete sequence.
ACCESSION   BX545850
VERSION     BX545850.5  GI:51890099
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 137836)
  AUTHORS   Johnson,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 4, 2004 this sequence version replaced gi:51889334.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml RP71-14F13
            is from a Zebrafish mixed sex BAC library VECTOR: pTARBAC2.
FEATURES             Location/Qualifiers
     source          1. .137836
```

```
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="RP71-14P13"
                     /clone_lib="RPCI-71"
ORIGIN

Alignment Scores:
Pred. No.:                 6.01e+03        Length:          137836
Score:                     40.00           Matches:         7
Percent Similarity:        100.0%          Conservative:    1
Best Local Similarity:     87.5%           Mismatches:      0
Query Match:               88.9%           Indels:          0
DB:                        5               Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BX545850 (1-137836)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||||
Db     133308 CCTGACATTATATTCCTTCAGGAA 133285

RESULT 56
AC023394
LOCUS       AC023394               158050 bp    DNA     linear   PRI 11-DEC-2001
DEFINITION  Homo sapiens chromosome 18, clone RP11-389J22, complete sequence.
ACCESSION   AC023394
VERSION     AC023394.11  GI:17488705
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 158050)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 18, clone RP11-389J22
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 158050)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
            Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., Landers,T., Largocque,K., Lehoczky,J., Levine,R.,
            Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
            McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M.,
            Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
            Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
            Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
            Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 158050)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P.,
            FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S.,
            Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T.,
            Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J.,
            Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
            Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
            Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 158050)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-DEC-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Dec 11, 2001 this sequence version replaced gi:14589631.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L5537
                Center clone name: 389_J_22
            --------------.
FEATURES             Location/Qualifiers
     source          1. .158050
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="18"
                     /map="18"
                     /clone="RP11-389J22"
                     /clone_lib="RPCI-11 Human Male BAC"
     repeat_region   1017. .1051
                     /rpt_family="AT_rich"
     repeat_region   1310. .1620
                     /rpt_family="AluSg"
     repeat_region   complement(1788. .1847)
                     /rpt_family="MER96"
     repeat_region   1849. .1868
                     /rpt_family="(T)n"
     repeat_region   complement(2622. .4128)
                     /rpt_family="L1HS"
```

```
     repeat_region   4129. .4237
                     /rpt_family="MER96"
     repeat_region   4600. .4627
                     /rpt_family="AT_rich"
     repeat_region   5087. .5167
                     /rpt_family="(TC)n"
     repeat_region   5168. .5347
                     /rpt_family="(TC)n"
     repeat_region   complement(5446. .5776)
                     /rpt_family="AluSx"
     repeat_region   5870. .6422
                     /rpt_family="MER82"
     repeat_region   complement(7136. .7240)
                     /rpt_family="L2"
     repeat_region   complement(7269. .7565)
                     /rpt_family="AluJb"
     repeat_region   complement(7573. .7774)
                     /rpt_family="L2"
     repeat_region   complement(9683. .9820)
                     /rpt_family="MER69"
     repeat_region   complement(10004. .10302)
                     /rpt_family="AluJo"
     repeat_region   11736. .11780
                     /rpt_family="GA-rich"
     repeat_region   complement(12488. .12668)
                     /rpt_family="L2"
     repeat_region   13110. .13180
                     /rpt_family="L2"
     repeat_region   13267. .13565
                     /rpt_family="AluY"
     repeat_region   13926. .14219
                     /rpt_family="AluSx"
     repeat_region   15505. .15530
                     /rpt_family="(T)n"
     repeat_region   16477. .16752
                     /rpt_family="AluJb"
     repeat_region   17399. .17544
                     /rpt_family="MER104"
     repeat_region   18064. .18293
                     /rpt_family="MIR"
     repeat_region   19060. .19356
                     /rpt_family="THE1C"
     repeat_region   19357. .19376
                     /rpt_family="(TAAA)n"
     repeat_region   19377. .19426
                     /rpt_family="THE1C"
     repeat_region   19725. .19745
                     /rpt_family="AT_rich"
     repeat_region   complement(19885. .20047)
                     /rpt_family="MIR3"
     repeat_region   20451. .20483
                     /rpt_family="(CAA)n"
     repeat_region   complement(21217. .21334)
                     /rpt_family="MER5B"
     repeat_region   complement(21806. .22173)
                     /rpt_family="LTR16A"
     repeat_region   22715. .23007
                     /rpt_family="AluY"
     repeat_region   complement(23392. .23697)
                     /rpt_family="AluSc"
     repeat_region   complement(23801. .23913)
                     /rpt_family="MER5A"
     repeat_region   24003. .24029
                     /rpt_family="AT_rich"
     repeat_region   24190. .24430
                     /rpt_family="L2"
     repeat_region   28019. .28044
                     /rpt_family="(TG)n"
     repeat_region   29222. .29243
                     /rpt_family="AT_rich"
     repeat_region   29978. .30007
                     /rpt_family="(TG)n"
     repeat_region   31291. .31465
                     /rpt_family="MIR"
     repeat_region   31469. .31519
                     /rpt_family="L1PA14"
     repeat_region   33596. .33617

Alignment Scores:
Pred. No.:                6.93e+03        Length:          158050
Score:                    40.00           Matches:         7
Percent Similarity:       100.0%          Conservative:    2
Best Local Similarity:    77.8%           Mismatches:      0
Query Match:              88.9%           Indels:          0
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AC023394 (1-158050)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||:::|||:::
Db      38332 CCAGATGTTATTTTTCTCAGGGAAATT 38358

RESULT 57
CR974453/c
LOCUS       CR974453              167610 bp    DNA     linear   HTG 23-JUN-2005
DEFINITION  Sus scrofa chromosome 17 clone CH242-276I24, WORKING DRAFT
            SEQUENCE.
ACCESSION   CR974453
VERSION     CR974453.2  GI:68162307
KEYWORDS    HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 167610)
  AUTHORS   Johnson,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 23, 2005 this sequence version replaced gi:67509405.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            -------------- Project Information
            Center project name: bE276I24
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 167477 bases at least Q40
            Consensus quality: 167554 bases at least Q30
            Consensus quality: 167602 bases at least Q20
            Insert size: 167610; sum-of-contigs
            Insert size: 154519; 10.0% error; agarose-fp
            Quality coverage: 14.23x in Q20 bases; sum-of-contigs Quality
            coverage: 15.56x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *         1   167610: contig of 167610 bp in length.
FEATURES                Location/Qualifiers
     source             1. .167610
                        /organism="Sus scrofa"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9823"
                        /chromosome="17"
```

```
                     /clone="CH242-276I24"
                     /clone_lib="CHORI-242"
     misc_feature    1. .167610
                     /note="assembly_fragment:00037
                     clone_end:SP6
                     vector_side:left
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:                7.38e+03        Length:         167610
Score:                    40.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    87.5%           Mismatches:     0
Query Match:              88.9%           Indels:         0
DB:                       14              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CR974453 (1-167610)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||||||
Db      24498 CCAGATATTATTTTCCTTCAGGAA 24475

RESULT 58
BX649417/c
LOCUS       BX649417               174290 bp    DNA     linear   VRT 02-DEC-2004
DEFINITION  Zebrafish DNA sequence from clone CH211-227P7 in linkage group 4
            Contains a novel gene and three CpG islands, complete sequence.
ACCESSION   BX649417
VERSION     BX649417.8  GI:38372953
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 174290)
  AUTHORS   Wood,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests:
            http://www.sanger.ac.uk/Projects/D_rerio/faqs.shtml#dataeight
COMMENT     On Nov 18, 2003 this sequence version replaced gi:38228955.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            CH211-227P7 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            Any regions longer than 1kb tagged as misc-feature 'unsure' are
            part of a tandem repeat of more than 10kb in length where it has
            not been possible to anchor the base differences between repeat
            copies.  The region has been built up based on the repeat element
            to match the total size of repeat indicated by restriction digest,
            but repeat copies may not be in the correct order and the usual
            finishing criteria may not apply.
FEATURES             Location/Qualifiers
     source          1. .174290
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /chromosome="4"
                     /clone="CH211-227P7"
                     /clone_lib="CHORI-211"
     misc_feature    51016. .51057
                     /note="Sequence from uni-directional dGTP big dye
                     terminator reads only."
     misc_feature    69614. .69746
                     /note="Sequence from uni-directional dGTP big dye
                     terminator reads only."
     misc_feature    69934. .69938
                     /note="1327 bases of IS2 transposon (V00610) removed here.
                     This sequence represents the duplicated flanking sequence
                     of the IS2."
     gene            127227. .141218
                     /locus_tag="CH211-227P7.1-001"
     mRNA            join(<127227. .127450,127545. .127673,135348. .135541,
                     136884. .137031,137455. .137659,140779. .>141218)
                     /locus_tag="CH211-227P7.1-001"
                     /product="novel gene"
                     /note="match: ESTs: Em:BQ263153.1 Em:BQ284802.1"
     CDS             join(<127227. .127450,127545. .127673,135348. .135541,
                     136884. .137031,137455. .137659,140779. .141083)
                     /locus_tag="CH211-227P7.1-001"
                     /standard_name="OTTDARP00000008644"
                     /codon_start=3
                     /product="novel gene"
                     /protein_id="CAI11834.1"
                     /db_xref="GI:55962596"
                     /translation="SMESKLSPYQETIESMIRNGCTDEYISNFLKFECGVSKGASKMS
                     VRRFCAVHGIQRQRIVDSQLEVEVAKAVTETGPAFGRKMMTGYLSSKGVRASEGRVGD
                     VLRRINLPYHMARYHGSRNLNPIPYAAEYVGHKLHIDQNEKMVMFGVTHVAAVDGYSR
                     KIVGHSTMPVKNNLTIYQNVYRSAVIEYGIWDQVRVDHGREFYLTLFVQEKLAQFRNN
                     TERQPYCQTPSTNNHIVERIWPEVNNRVNYPLKGALIQLVDQEELDMQDNLTQYCVST
                     LMCQVAEIGLERVVQSWNAHRIPGIGVPNHLAHGGCPKKVSVDVLPDAVDAANCYDRE
                     VGSSLTRESSFGRNPFSSLEDQEAAQREFALHFPDMCLLFESTIHFQPNYFQQAIRKL
                     IDVVRCHA"
     polyA_signal    139200. .139205
                     /locus_tag="CH211-227P7.1-001"
     polyA_site      139218
                     /locus_tag="CH211-227P7.1-001"
     misc_feature    172290
                     /note="Clone_right_end: CH211-227P7"
ORIGIN

Alignment Scores:
Pred. No.:                7.68e+03        Length:         174290
Score:                    40.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    87.5%           Mismatches:     0
Query Match:              88.9%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BX649417 (1-174290)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||||||
Db      89075 CCTGACATTATATTCCTTCAGGAA 89052

RESULT 59
AC135187/c
LOCUS       AC135187               176333 bp    DNA     linear   ROD 28-JAN-2005
DEFINITION  Mus musculus BAC clone RP24-263A3 from 5, complete sequence.
ACCESSION   AC135187
VERSION     AC135187.3  GI:55792364
KEYWORDS    HTG.
```

```
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 176333)
  AUTHORS   Nguyen,C., Bielicki,L. and Haakenson,W.
  TITLE     The sequence of Mus musculus BAC clone RP24-263A3
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 176333)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 176333)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 176333)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   5  (bases 1 to 176333)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Nov 17, 2004 this sequence version replaced gi:24580487.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: M_BB0263A03
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-24 BAC Library has been constructed by Pieter de Jong and
            coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
            and/or brain genomic DNA. The clone and detailed information can be
            obtained from Pieter de Jong and coworkers at http://www.chori.org

            NEIGHBORING SEQUENCE INFORMATION:
            This sequence is the entire insert of the clone.
FEATURES              Location/Qualifiers
     source           1. .176333
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:10090"
                      /chromosome="5"
                      /map="5"
                      /clone="RP24-263A3"
                      /clone_lib="RPCI-24"
     unsure           17228. .17284
                      /note="Sequence derived from one plasmid subclone."
     misc_feature     22886. .23168
                      /note="Sequence derived from PCR product of project DNA."
ORIGIN

Alignment Scores:
Pred. No.:                 7.78e+03        Length:          176333
Score:                     40.00           Matches:         7
Percent Similarity:        100.0%          Conservative:    2
Best Local Similarity:     77.8%           Mismatches:      0
Query Match:               88.9%           Indels:          0
DB:                        9               Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AC135187 (1-176333)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||||||||||||:::
Db     113940 CCAGAAGTCATCTTCTTGCAAGAAATA 113914

RESULT 60
BX000438/c
LOCUS       BX000438              195087 bp    DNA     linear   VRT 12-FEB-2003
DEFINITION  Zebrafish DNA sequence from clone CH211-209M20 in linkage group 20,
            complete sequence.
ACCESSION   BX000438
VERSION     BX000438.7  GI:28372094
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 195087)
  AUTHORS   Corby,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 12, 2003 this sequence version replaced gi:28300518.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Repeat names
            beginning 'Dr' were identified by the Recon repeat discovery system
            (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr'
            were identified by Rick Waterman (Stephen Johnson lab, WashU).  For
            further information see http://www/Projects/D_rerio/fishmask.shtml
            CH211-209M20 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES              Location/Qualifiers
     source           1. .195087
```

```
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-209M20"
                     /clone_lib="CHORI-211"
ORIGIN

Alignment Scores:
Pred. No.:                 8.65e+03        Length:          195087
Score:                     40.00           Matches:         7
Percent Similarity:        100.0%          Conservative:    1
Best Local Similarity:     87.5%           Mismatches:      0
Query Match:               88.9%           Indels:          0
DB:                        5               Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x BX000438 (1-195087)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||||
Db     116692 CCTGATATCATTTTTCTTCAAGAG 116669

RESULT 61
CR956635/c
LOCUS       CR956635               211493 bp    DNA     linear   HTG 28-JUL-2005
DEFINITION  Sus scrofa chromosome 17 clone CH242-184L20, WORKING DRAFT
            SEQUENCE, 3 unordered pieces.
ACCESSION   CR956635
VERSION     CR956635.5  GI:71142728
KEYWORDS    HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 211493)
  AUTHORS   Lovell,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUL-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 25, 2005 this sequence version replaced gi:68611525.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            -------------- Project Information
            Center project name: bE184L20
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 211166 bases at least Q40
            Consensus quality: 211242 bases at least Q30
            Consensus quality: 211272 bases at least Q20
            Insert size: 211293; sum-of-contigs
            Insert size: 190949; 13.9% error; agarose-fp
            Quality coverage: 16.93x in Q20 bases; sum-of-contigs Quality
            coverage: 19.06x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    33818: contig of 33818 bp in length
            *    33819    33918: gap of 100 bp
            *    33919    45162: contig of 11244 bp in length
            *    45163    45262: gap of 100 bp
            *    45263   211493: contig of 166231 bp in length.
FEATURES             Location/Qualifiers
     source          1. .211493
                     /organism="Sus scrofa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9823"
                     /chromosome="17"
                     /clone="CH242-184L20"
                     /clone_lib="CHORI-242"
     misc_feature    1. .33818
                     /note="assembly_fragment:01086
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:left"
     misc_feature    33919. .45162
                     /note="assembly_fragment:00193
                     fragment_chain:1"
     misc_feature    45263. .211493
                     /note="assembly_fragment:01088
                     fragment_chain:1
                     clone_end:T7
                     vector_side:right"
ORIGIN

Alignment Scores:
Pred. No.:                 9.42e+03        Length:          211493
Score:                     40.00           Matches:         7
Percent Similarity:        100.0%          Conservative:    1
Best Local Similarity:     87.5%           Mismatches:      0
Query Match:               88.9%           Indels:          0
DB:                        14              Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x CR956635 (1-211493)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||||
Db      66497 CCAGATATTATTTTCCTTCAGGAA 66474

RESULT 62
CR318663
LOCUS       CR318663               215046 bp    DNA     linear   VRT 29-JUN-2005
DEFINITION  Zebrafish DNA sequence from clone CH211-214J5 in linkage group 12,
            complete sequence.
ACCESSION   CR318663
VERSION     CR318663.12  GI:68162013
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 215046)
  AUTHORS   Dunn,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests:
            http://www.sanger.ac.uk/Projects/D_rerio/faqs.shtml#dataeight
COMMENT     On Jun 23, 2005 this sequence version replaced gi:67763629.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
```

on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
Zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Any regions longer than 1kb tagged as misc-feature 'unsure' are
part of a tandem repeat of more than 10kb in length where it has
not been possible to anchor the base differences between repeat
copies. The region has been built up based on the repeat element
to match the total size of repeat indicated by restriction digest,
but repeat copies may not be in the correct order and the usual
finishing criteria may not apply. CH211-214J5 is from a CHORI-211
BAC library
VECTOR: pTARBAC2.1.

```
FEATURES             Location/Qualifiers
     source          1. .215046
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /chromosome="12"
                     /clone="CH211-214J5"
                     /clone_lib="CHORI-211"
ORIGIN

Alignment Scores:
Pred. No.:              9.58e+03        Length:         215046
Score:                  40.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  87.5%           Mismatches:     0
Query Match:            88.9%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CR318663 (1-215046)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||||
Db      91112 CCTGACATTATATTCCTTCAGGAA 91135

RESULT 63
BX000452/c
LOCUS       BX000452              220420 bp    DNA     linear   VRT 02-JUN-2003
DEFINITION  Zebrafish DNA sequence from clone DKEY-12J12 in linkage group 14,
            complete sequence.
ACCESSION   BX000452
VERSION     BX000452.7  GI:31335492
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 220420)
  AUTHORS   Whitehead,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 2, 2003 this sequence version replaced gi:30024382.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Repeat names
            beginning 'Dr' were identified by the Recon repeat discovery system
            (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr'
            were identified by Rick Waterman (Stephen Johnson lab, WashU). For
            further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-12J12
            is from a Zebrafish BAC library
            VECTOR: pIndigoBAC-5.
FEATURES             Location/Qualifiers
     source          1. .220420
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-12J12"
                     /clone_lib="DanioKey"
ORIGIN

Alignment Scores:
Pred. No.:              9.83e+03        Length:         220420
Score:                  40.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  87.5%           Mismatches:     0
Query Match:            88.9%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BX000452 (1-220420)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||||
Db      33075 CCTGACATTATATTCCTTCAGGAA 33052

RESULT 64
CR848052/c
LOCUS       CR848052              250603 bp    DNA     linear   HTG 16-APR-2005
DEFINITION  Danio rerio clone DKEY-28F6, *** SEQUENCING IN PROGRESS ***, 16
            unordered pieces.
ACCESSION   CR848052
VERSION     CR848052.3  GI:62719075
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 250603)
  AUTHORS   McLay,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 17, 2005 this sequence version replaced gi:62551172.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zK28F6
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 243561 bases at least Q40
            Consensus quality: 245099 bases at least Q30
```

```
            Consensus quality: 246433 bases at least Q20
            Insert size: 249103; sum-of-contigs
            Insert size: 261573; 7.8% error; agarose-fp
            Quality coverage: 6.44x in Q20 bases; sum-of-contigs Quality
            coverage: 6.23x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 16 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3451: contig of 3451 bp in length
            *     3452     3551: gap of 100 bp
            *     3552     6186: contig of 2635 bp in length
            *     6187     6286: gap of 100 bp
            *     6287    10237: contig of 3951 bp in length
            *    10238    10337: gap of 100 bp
            *    10338    12700: contig of 2363 bp in length
            *    12701    12800: gap of 100 bp
            *    12801    19802: contig of 7002 bp in length
            *    19803    19902: gap of 100 bp
            *    19903    36045: contig of 16143 bp in length
            *    36046    36145: gap of 100 bp
            *    36146    67601: contig of 31456 bp in length
            *    67602    67701: gap of 100 bp
            *    67702    99792: contig of 32091 bp in length
            *    99793    99892: gap of 100 bp
            *    99893   119046: contig of 19154 bp in length
            *   119047   119146: gap of 100 bp
            *   119147   132710: contig of 13564 bp in length
            *   132711   132810: gap of 100 bp
            *   132811   167723: contig of 34913 bp in length
            *   167724   167823: gap of 100 bp
            *   167824   217028: contig of 49205 bp in length
            *   217029   217128: gap of 100 bp
            *   217129   228087: contig of 10959 bp in length
            *   228088   228187: gap of 100 bp
            *   228188   236007: contig of 7820 bp in length
            *   236008   236107: gap of 100 bp
            *   236108   246201: contig of 10094 bp in length
            *   246202   246301: gap of 100 bp
            *   246302   250603: contig of 4302 bp in length.
FEATURES             Location/Qualifiers
     source          1. .250603
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-28F6"
                     /clone_lib="DanioKey"
     misc_feature    1. .3451
                     /note="assembly_fragment:00037
                     fragment_chain:1"
     misc_feature    3552. .6186
                     /note="assembly_fragment:00051
                     fragment_chain:1"
     misc_feature    6287. .10237
                     /note="assembly_fragment:00067
                     fragment_chain:1"
     misc_feature    10338. .12700
                     /note="assembly_fragment:00023
                     fragment_chain:1"
     misc_feature    12801. .19802
                     /note="assembly_fragment:00090
                     fragment_chain:1"
     misc_feature    19903. .36045
                     /note="assembly_fragment:00791
                     fragment_chain:1"
     misc_feature    36146. .67601
                     /note="assembly_fragment:01244
                     fragment_chain:1"
     misc_feature    67702. .99792
                     /note="assembly_fragment:00962
                     fragment_chain:2"
     misc_feature    99893. .119046
                     /note="assembly_fragment:00623
                     fragment_chain:2"
     misc_feature    119147. .132710
                     /note="assembly_fragment:00502
                     fragment_chain:2"
     misc_feature    132811. .167723
                     /note="assembly_fragment:01573
                     fragment_chain:2"
     misc_feature    167824. .217028
                     /note="assembly_fragment:01904
                     fragment_chain:2"
     misc_feature    217129. .228087
                     /note="assembly_fragment:00292
                     fragment_chain:2"
     misc_feature    228188. .236007
                     /note="assembly_fragment:00213
                     fragment_chain:3"
     misc_feature    236108. .246201
                     /note="assembly_fragment:00389
                     fragment_chain:3"
     misc_feature    246302. .250603
                     /note="assembly_fragment:00145.0"
ORIGIN

Alignment Scores:
Pred. No.:                1.13e+04        Length:         250603
Score:                    40.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    87.5%           Mismatches:     0
Query Match:              88.9%           Indels:         0
DB:                       14              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x CR848052 (1-250603)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||||
Db     240974 CCTGACATCATATTCCTTCAGGAA 240951

RESULT 65
AE016918
LOCUS       AE016918               302178 bp    DNA     linear   BCT 14-OCT-2003
DEFINITION  Chromobacterium violaceum ATCC 12472 section 9 of 16 of the
            complete genome.
ACCESSION   AE016918 AE016825
VERSION     AE016918.1  GI:34103550
KEYWORDS    .
SOURCE      Chromobacterium violaceum ATCC 12472
  ORGANISM  Chromobacterium violaceum ATCC 12472
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Chromobacterium.
REFERENCE   1  (bases 1 to 302178)
  AUTHORS   .
  CONSRTM   Brazilian National Genome Project Consortium
  TITLE     The complete genome sequence of Chromobacterium violaceum reveals
            remarkable and exploitable bacterial adaptability
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11660-11665 (2003)
   PUBMED   14500782
REFERENCE   2  (bases 1 to 302178)
  AUTHORS   Vasconcelos,A.T.R., de Almeida,D.F., Almeida,F.C., de
            Almeida,L.G.P., de Almeida,R., Goncalves,J.A.A., Andrade,E.M.,
            Antonio,R.V., Araripe,J., de Araujo,M.F.F., Filho,S.A., Azevedo,V.,
            Batista,A.J., Bataus,L.A.M., Batista,J.S., Belo,A., vander Berg,C.,
            Blamey,J., Bogo,M., Bonato,S., Bordignon,J., Brito,C.A.,
            Brocchi,M., Burity,H.A., Camargo,A.A., Cardoso,D.D.P.,
            Carneiro,N.P., Carraro,D.M., Carvalho,C.M.B., Cascardo,J.C.M.,
            Cavada,B.S., Chueire,L.M.O., Pasa,T.B.C., Duran,N., Fagundes,N.,
            Falcao,C.L., Fantinatti,F., Farias,I.P., Felipe,M.S.S.,
            Ferrari,L.P., Ferro,J.A., Ferro,M.I.T., Franco,G.R.,
            Freitas,N.S.A., Furlan,L.R., Gazzinelli,R.T., Gomes,E.A.,
```

```
            Goncalves,P.R., Grangeiro,T.B., Grattapaglia,D., Grisard,E.C.,
            Guimaraes,C.T., Hanna,E.S., Hungria,M., Jardim,S.N., Laurino,J.,
            Leoi,L.C.T., Fassarella,L., Lima,A., Loureiro,M.F., Lyra,M.C.P.,
            Macedo,M., Madeira,H.M.F., Manfio,G.P., Maranhao,A.Q.,
            Martins,W.S., di Mauro,S.M.Z., de Medeiros,S.R.B., Meissner,R.V.,
            Menck,C.F.M., Moreira,M.A.M., Nascimento,F.F., Nicolas,M.F.,
            Oliveira,J.G., Oliveira,S.C., Paixao,R.F.C., Parente,J.A.,
            Pedrosa,F.O., Pena,S.J.D., Perreira,J.O., Perreira,M.,
            Pinto,L.S.R.C., Pinto,L.S., Porto,J.I.R., Potrich,D.P.,
            Neto,C.E.R., Reis,A.M.M., Rigo,L.U., Rondinelli,E., dos
            Santos,E.B.P., Santos,F.R., Schneider,M.P.C., Seuanez,H.N.,
            Silva,A.M.R., da Silva,A.L.C., Silva,D.W., Silva,R., Simoes,I.C.,
            Simon,D., Soares,C.M.A., Soares,R.B.A., Souza,E.M., Souza,K.R.L.,
            Souza,R.C., Steffens,M.B.R., Steindel,M., Teixeira,S.R.,
            Urmenyi,T., Vettore,A., Wassem,R., Zaha,A. and Simpson,A.J.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JAN-2003) Labinfo, LNCC - Laboratorio Nacional de
            Computacao Cientifica, Rua Getulio Vargas 333, Petropolis, RJ
            25651070, Brazil
FEATURES             Location/Qualifiers
     source          1. .302178
                     /organism="Chromobacterium violaceum ATCC 12472"
                     /mol_type="genomic DNA"
                     /strain="ATCC 12472"
                     /db_xref="taxon:243365"
     gene            162. .1322
                     /locus_tag="CV2240"
     CDS             162. .1322
                     /locus_tag="CV2240"
                     /note="identified by sequence similarity; putative; ORF
                     located using Glimmer/Blastx/COG0845/TC:8.A.1.6.1"
                     /codon_start=1
                     /transl_table=11
                     /product="probable multidrug efflux protein"
                     /protein_id="AAQ59912.1"
                     /db_xref="GI:34103551"
                     /translation="MVSVFAAGMHWNMRILTVVAMAASLAGCGAAERTPEAGKPPEVG
                     VVRLEAKDVAIRAELAGRTVAYRSAEVRPQVNGIVLKRHFTEGSLVKAGDLLYQIDPA
                     GYQAALRQAEATLASARAALGSLKGRAERYAELAKIDGVSRQEREEADADYRKGLAAA
                     KAGEAAVAAARVDVSRTRVLAPISGRIGRSAVSEGALVTSGQSAALATIQQLDPMYVD
                     IVQSSQDLLAFKRRQADGELEPAGAAARLKLEDGAPYRHPGKLRFTELNVDRVSGAVT
                     LRAEFPNPEGLLLPGMYVRAELEQGVRRQAVLAPQQGVARDDKGRATALVLGAGDKVE
                     QRLLTTEGTHGAYWLVTAGLKAGDRLIVDGLQRARPGEAAKPVALPAYAGRS"
     gene            1326. .4454
                     /gene="acrB"
                     /locus_tag="CV2241"
     CDS             1326. .4454
                     /gene="acrB"
                     /locus_tag="CV2241"
                     /note="identified by sequence similarity; putative; ORF
                     located using
                     Glimmer/GeneMark/Blastx/COG0841/TC:2.A.6.2.2"
                     /codon_start=1
                     /transl_table=11
                     /product="acriflavin resistance protein B"
                     /protein_id="AAQ59913.1"
                     /db_xref="GI:34103552"
                     /translation="MLARFFIDRPVFAWVLAIVIMLAGLFALRALPVSQYPDIAPPTV
                     MVSASYPGASAQVVETSVIQVLEQELKGLDNLMYFSSTSSSSGQAEIMLTFRQGANPD
                     TAQVQVQNKVSQMLARLPQQVQQNGLSVQKMQGDFLLIASFYDISGQRGDTDIADWLA
                     SNIQDPLSRVDGVGSVRAFGAPYAMRIWLDPHKLNSYGLIPSDVSAAVSKQNTEVSVG
                     ELGALPSPPGQQLNAMVTAWSRLSTPEQFRGIVVKALPDGGAVRLGDVARVELGQEDY
                     GSSSRLNGHPASGLAVMLAPGANALATAEAVKARVAELQRNFPPGIRVAYPEDTTHFV
                     RLSIRAVMETLLEAVALVVLVMFVFLRNWRATLIPALTVPVVLLGTFGVLAAAGFGIN
                     TLTLFGMVLAIGLLVDDAIVVVENVERVMHEQGVDAHAATVRSMGEIGGALVGIGLVL
                     GAVFLPMAFFGGSVGVIYRQFSITIVAAMALSVLTALTLTPALCATLLRPEKPSDNGP
                     LAAFERRFLRLQSAYAERWLPAWLRRPLRGGLVYLALAAGLAALYPQLPTAFIPDEDQ
                     GTVMVQFTLPTGATHERADKVARAVEQHFLVTEKANVESVYMLSGFSFGGSGQNAGMA
                     FVALKDWSLRPGAGNSAQAIADRATAALASQRDAEVFGMVLPPIGGLGQTNGFEFWLQ
                     DASGMGREGLAKARDQLLKAAQGDDRLSAVRSGAADDKPQLRLDVDQSKAAALGLDLA
                     NVADTLGAAWGGRYINDFVDRGRVKKVMMQADAPYRSKPEDLGLWFVRGADGAMTPFS
                     AFARARWEYGPAQLSRYNGLPALPLSGSAAAGISSGEAMRAVEAIAGKLPGTQYEWSG
                     LSYQDRLSSGQTPLLYAVSILFVFLCLAALYESWLVPCAVMTAIPLGVLGAVLAVMAR
                     GLANDIYFQVGLLATIGLSAKNAILIIEFAEAAIRRGDDALSAALAAARQRLRPILMT
                     SLAFGVGVAPLVWAGGPGAGSQNAIGAAVLGGVLSATLLTLFFVPLAHVLIRGLAGRV
                     AERRWKPAEARP"
     gene            4451. .5830
                     /gene="ibeB"
                     /locus_tag="CV2242"
     CDS             4451. .5830
                     /gene="ibeB"
                     /locus_tag="CV2242"
                     /note="identified by sequence similarity; putative; ORF
                     located using
                     Glimmer/GeneMark/Blastx/COG1538/TC:1.B.17.3.5"
                     /codon_start=1
                     /transl_table=11
                     /product="outer membrane drug efflux lipoprotein"
                     /protein_id="AAQ59914.1"
                     /db_xref="GI:34103553"
                     /translation="MKRAFCLLMGAALAGCASPQPDYQRPAAPVPAHWPAADAESGQA
                     DIAWRRFFVDDGLRETIAQALANNRDLRVAAANVQAARAQYRILDAARLPAVQLQGKS
                     DRQLTRGGDPSLAADSRGVSAGLGVSAFELDLFGRVRSLSDAQWRRYLASEAGAQAAR
                     VSLIAETANAYLTLAADRSQLAAARTTLRSAERSLELTRERLRLGASSALDVSAMETV
                     RQQARADVARYQGQAERDRLALQLLVGAELDQTRVPAALEDRPYTLAALPADAPADAL
                     LRRPDVRQAENLLRAAHADVDAARAGLWPKLGLSASSGLSGVTLAGLLNAPARALSLA
                     LDASLPALDGGAAKAGVDQAEAQRLAALAAYEKALQNGFREVAGALSQRQALAEEKQA
                     RQAALASAEGSLRLAEARYRQGLDGYLSLLEAQRTAQAARLQWIGAHLAEAENMAALY
                     RSLGGGLES"
     gene            complement(5844. .6299)
                     /locus_tag="CV2243"
     CDS             complement(5844. .6299)
                     /locus_tag="CV2243"
                     /note="identified by sequence similarity; putative; ORF
                     located using GeneMark/Blastx/COG0346"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAQ59915.1"
                     /db_xref="GI:34103554"
                     /translation="MNIQGIDHFTIRTADLEASAVFYQRVLGLSDGPRPSFRFAGKWL
                     YAGGRPVLHLVETAVDDAELEAYLGRRENRSGSGRVDHIALRGQNLVDMQMRLLSLGQ
                     DFHERVVPELGEHQLFIDDPDGVRIELIFPYAPDNRMAGGQLPRLDIQS"
     gene            complement(6338. .6991)
                     /gene="nfnB"
                     /locus_tag="CV2244"
     CDS             complement(6338. .6991)
                     /gene="nfnB"
                     /locus_tag="CV2244"
                     /EC_number="1.6.99.7"
                     /note="identified by sequence similarity; putative; ORF
                     located using Glimmer/GeneMark/Blastx/COG0778"
                     /codon_start=1
                     /transl_table=11
                     /product="oxygen-insensitive NAD(P)H nitroreductase"
                     /protein_id="AAQ59916.1"
                     /db_xref="GI:34103555"
                     /translation="MNIAHYAQTRYTTKAFDPGFRLSAGQIEQIETLLRHSPSSTNSQ
                     PWHFFIAGSDESKARVAKATADGYAFNQAKVLNASHVVVFCTRAALDDAYLRTLLDQE
                     ERDGRFASPEARAGQHKGRSHFADMHRFELRDAPHWMEKQVYLAVGTLLLGAAALEID
                     ACPIEGFDQRTLGEELGLREKGLIASVIVALGRRSDEDFNARLPKSRLPAEAVITRL"
     gene            complement(7013. .8194)
                     /gene="nemA2"
                     /locus_tag="CV2245"
     CDS             complement(7013. .8194)
                     /gene="nemA2"
                     /locus_tag="CV2245"
                     /EC_number="1.-.-.-"
                     /note="identified by sequence similarity; putative; ORF
                     located using Glimmer/Blastx/COG1902"
                     /codon_start=1
                     /transl_table=11
                     /product="flavoprotein NADH-dependent oxidoreductase"
                     /protein_id="AAQ59917.1"
                     /db_xref="GI:34103556"
                     /translation="MQSIMSSLERLPLRGGRIPRLNDMPAPDLFSPLQAGALPLPSRI
```

```
                     VMAPMTRCRTDQPGDIPNAMMARYYAQRAGAGLIVSEATQISRQGQGYSFTPGIYSQD
                     QIDGWRRVTDAVRQAGGRMVLQLWHVGRMSHPMFHDGAPPVAPSAIAPDAQVWVADET
                     GAGRMVDCPTPRALETAEIAAIVGDYRQAARNAMAAGFDGVEIHAANGYLIDQFLRST
                     SNRRDDAYGGSLENRLRFLMEVAAAVADEIGAARVGVRLAPYITARGMNCPEIVPAIL
                     EAAVRLNGLGIAYLHLSEADWDDAPQVDDAFRHALRARFRGAIIVAGRYDQPRGERLL
                     AEGLADLIAFGRPFIANPDLPERLARGWPLADFDPSALFGGDGRGYDDYPRYREMATA
                     R"
     gene            8777. .9652
                     /locus_tag="CV2246"
     CDS             8777. .9652
                     /locus_tag="CV2246"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAQ59918.1"
                     /db_xref="GI:34103557"
                     /translation="MGAMTSIDFQSDRFYATNASFTATDEVIEFTRPLRSLKISYFTY
                     DRHYADGSRVALTNAPDWIQHYWRREMFKLAVFEKDSDNFTDGYVFWNWLNREPIYSA
                     AAQFGIDNGVTLIERGDGHCDFFHFGSVHDNSVNNDYLAGKMAHLQQFVSLFKYKLHD
                     VIEAACNDRIYLPESARGKLILDSDINHDMLVSDFYSLMDRRDVSRIYLGDADRSAYL
                     TKKEVALMRELVSGASCAEAAVSLDITSDAVNKHIKNIKERLNCRSLCQLGFVIGKLS
                     ARNIYPFSVASGANK"
     gene            9649. .10323
                     /gene="talC"
                     /locus_tag="CV2247"

Alignment Scores:
Pred. No.:              1.37e+04        Length:         302178
Score:                  40.00           Matches:        7
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  77.8%           Mismatches:     0
Query Match:            88.9%           Indels:         0
DB:                     1               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AE016918 (1-302178)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||||||||
Db      55923 CCGGACATGCTGTTTCTGCAGGAGGTG 55949

RESULT 66
BV547170
LOCUS       BV547170                 587 bp    DNA     linear   STS 09-APR-2005
DEFINITION  rrj99b02.b1 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV547170
VERSION     BV547170.1  GI:62438190
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 587)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 587
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES                Location/Qualifiers
     source             1. .587
                        /organism="Pan troglodytes verus"
                        /mol_type="genomic DNA"
                        /sub_species="verus"
                        /db_xref="taxon:37012"
                        /clone_lib="Clint"
     STS                <1. .>587
ORIGIN

Alignment Scores:
Pred. No.:              32.9            Length:         587
Score:                  39.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   3
Best Local Similarity:  66.7%           Mismatches:     0
Query Match:            86.7%           Indels:         0
DB:                     10              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BV547170 (1-587)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||||:::
Db        159 CCTGACCTCCTCTTCCTTCAGGAAATC 185

RESULT 67
BV553589
LOCUS       BV553589                 725 bp    DNA     linear   STS 09-APR-2005
DEFINITION  S217P62326FB8.T0 Noemie Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV553589
VERSION     BV553589.1  GI:62444610
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 725)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
```

```
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 725
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .725
                     /organism="Pan troglodytes troglodytes"
                     /mol_type="genomic DNA"
                     /sub_species="troglodytes"
                     /db_xref="taxon:37011"
                     /clone_lib="Noemie"
     STS             <1. .>725
ORIGIN

Alignment Scores:
Pred. No.:              41              Length:         725
Score:                  39.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   3
Best Local Similarity:  66.7%           Mismatches:     0
Query Match:            86.7%           Indels:         0
DB:                     10              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BV553589 (1-725)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||||:::
Db        420 CCTGACCTCCTCTTCCTTCAGGAAATC 446

RESULT 68
BV542059
LOCUS       BV542059                  739 bp    DNA     linear   STS 08-APR-2005
DEFINITION  G591P53098RF2.T0 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV542059
VERSION     BV542059.1  GI:62419910
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 739
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="Pan troglodytes verus"
                     /mol_type="genomic DNA"
                     /sub_species="verus"
                     /db_xref="taxon:37012"
                     /clone_lib="Clint"
     STS             <1. .>739
ORIGIN

Alignment Scores:
Pred. No.:              41.9            Length:         739
Score:                  39.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   3
Best Local Similarity:  66.7%           Mismatches:     0
```

```
Query Match:            86.7%       Indels:        0
DB:                     10          Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x BV542059 (1-739)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||||||:::
Db        495 CCTGACCTCCTCTTCCTTCAGGAAATC 521

RESULT 69
BV518605/c
LOCUS       BV518605                748 bp    DNA     linear   STS 08-APR-2005
DEFINITION  ssw41e11.g1 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV518605
VERSION     BV518605.1  GI:62396363
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 748)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 748
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES              Location/Qualifiers
     source           1. .748
                      /organism="Pan troglodytes verus"
                      /mol_type="genomic DNA"
                      /sub_species="verus"
                      /db_xref="taxon:37012"
                      /clone_lib="Clint"
     STS              <1. .>748
ORIGIN

Alignment Scores:
Pred. No.:              42.4        Length:        748
Score:                  39.00       Matches:       6
Percent Similarity:     100.0%      Conservative:  3
Best Local Similarity:  66.7%       Mismatches:    0
Query Match:            86.7%       Indels:        0
DB:                     10          Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x BV518605 (1-748)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||||||:::
Db        114 CCTGACCTCCTCTTCCTTCAGGAAATC 88

RESULT 70
BV553590
LOCUS       BV553590                752 bp    DNA     linear   STS 09-APR-2005
DEFINITION  S217P62324FB8.T0 Noemie Pan troglodytes troglodytes STS genomic,
            sequence tagged site.
ACCESSION   BV553590
VERSION     BV553590.1  GI:62444611
KEYWORDS    STS.
SOURCE      Pan troglodytes troglodytes
  ORGANISM  Pan troglodytes troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 752
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
```

```
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES              Location/Qualifiers
     source           1. .752
                      /organism="Pan troglodytes troglodytes"
                      /mol_type="genomic DNA"
                      /sub_species="troglodytes"
                      /db_xref="taxon:37011"
                      /clone_lib="Noemie"
     STS              <1. .>752
ORIGIN

Alignment Scores:
Pred. No.:                42.6          Length:         752
Score:                    39.00         Matches:        6
Percent Similarity:       100.0%        Conservative:   3
Best Local Similarity:    66.7%         Mismatches:     0
Query Match:              86.7%         Indels:         0
DB:                       10            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BV553590 (1-752)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||||||:::
Db        413 CCTGACCTCCTCTTCCTTCAGGAAATC 439

RESULT 71
BV505641/c
LOCUS       BV505641                 801 bp    DNA     linear   STS 07-APR-2005
DEFINITION  qrf49e12.g1 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV505641
VERSION     BV505641.1  GI:62349383
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 801)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172580933
            Fax: 6172580903
            Email: mczody@broad.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 801
            Protocol:
            23,021,928 chimpanzee whole genome shotgun reads were aligned to
            the Human genome NCBI
            Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
            including Clint (Pan
            troglodytes verus), 3 other Pan troglodytes verus chimps
            (Donald,Karlien,Yvonne), 3 Pan
            troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
            of unknown origin
            (Gon,Unknown Chimp). Common names:  Pan troglodytes verus is the
            western chimp and Pan
            troglodytes troglodytes is the central chimp. To be included in
            chimpanzee SNP discovery, a
            read must be at least 500bp in length, at least 50% of its base
            calls must have Phred
            score >= 20, at least 30% of its base calls must satisfy
            SNQS(30,25)(single strand NQS, the
            base in question has Phred score >= 30, the surrounding 10 bases in
            the read have Phred
            score >= 25), and the read must have at least 200 bp SNQS(30,25)
            bases. Reads not uniquely
            placed in the genome and read pairs whose two ends were not
            consistently placed were
            discarded. After above filtering, NQS(30,25) standard was applied
            to all pairs of
            overlapping reads to call NQS bases and SNPs. Alignments (between
            two reads) with less
            than 100 NQS bases or with SNP rate > 0.01 were discarded. To
            exclude alignment between two
            copies of a single read, comparisons between two reads that share
            95% of their genome
            alignments (>=95% bases of read A and >=95% bases of read B were
            placed at the same locus
            of human genome) were discarded.
FEATURES              Location/Qualifiers
     source           1. .801
                      /organism="Pan troglodytes verus"
                      /mol_type="genomic DNA"
                      /sub_species="verus"
                      /db_xref="taxon:37012"
                      /clone_lib="Clint"
     STS              <1. .>801
ORIGIN

Alignment Scores:
Pred. No.:                45.5          Length:         801
Score:                    39.00         Matches:        6
Percent Similarity:       100.0%        Conservative:   3
Best Local Similarity:    66.7%         Mismatches:     0
Query Match:              86.7%         Indels:         0
DB:                       10            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x BV505641 (1-801)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||||||:::
Db        369 CCTGACCTCCTCTTCCTTCAGGAAATC 343

RESULT 72
BV547859/c
LOCUS       BV547859                 834 bp    DNA     linear   STS 09-APR-2005
DEFINITION  rnx59b01.g1 Clint Pan troglodytes verus STS genomic, sequence
            tagged site.
ACCESSION   BV547859
VERSION     BV547859.1  GI:62438879
KEYWORDS    STS.
SOURCE      Pan troglodytes verus
  ORGANISM  Pan troglodytes verus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
            Jaffe,D.B.
  TITLE     Initial Sequence of the Chimpanzee Genome and Comparison with the
            Human Genome
  JOURNAL   Unpublished (2005)
COMMENT
            Contact: Michael C. Zody
            Broad Institute of MIT and Harvard
```

320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 834
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base calls must have Phred
score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two
copies of a single read, comparisons between two reads that share 95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus
of human genome) were discarded.

FEATURES Location/Qualifiers
source 1. .834
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
STS <1. .>834
ORIGIN

Alignment Scores:

Pred. No.:	47.5	Length:	834
Score:	39.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	3
Best Local Similarity:	66.7%	Mismatches:	0
Query Match:	86.7%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x BV547859 (1-834)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::::::|||||||||||:::
Db        469 CCTGACCTCCTCTTCCTTCAGGAAATC 443
```

RESULT 73
BV543970
LOCUS BV543970 850 bp DNA linear STS 09-APR-2005
DEFINITION qpc96h04.g1 Clint Pan troglodytes verus STS genomic, sequence tagged site.
ACCESSION BV543970
VERSION BV543970.1 GI:62434990
KEYWORDS STS.
SOURCE Pan troglodytes verus
ORGANISM Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 850)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome
JOURNAL Unpublished (2005)
COMMENT

Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 850
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base calls must have Phred
score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two
copies of a single read, comparisons between two reads that share 95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus
of human genome) were discarded.
FEATURES Location/Qualifiers
source 1. .850
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
STS <1. .>850
ORIGIN

Alignment Scores:

Pred. No.:	48.5	Length:	850
Score:	39.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	3
Best Local Similarity:	66.7%	Mismatches:	0
Query Match:	86.7%	Indels:	0
DB:	10	Gaps:	0

```
US-10-757-745-2_COPY_145_153 (1-9) x BV543970 (1-850)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::||||||||||||:::
Db         50 CCTGACCTCCTCTTCCTTCAGGAAATC 76

RESULT 74
DME302712
LOCUS       DME302712               5395 bp    DNA     linear   INV 04-MAY-2004
DEFINITION  Drosophila melanogaster transposon You, gag and pol genes.
ACCESSION   AJ302712
VERSION     AJ302712.1  GI:11323017
KEYWORDS    endonuclease; gag-like protein; gal gene; pol gene; reverse
            transcriptase.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1
  AUTHORS   Berezikov,E., Bucheton,A. and Busseau,I.
  TITLE     A search for reverse transcriptase-coding sequences reveals new
            non-LTR retrotransposons in the genome of Drosophila melanogaster
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 5395)
  AUTHORS   Busseau,I.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2000) Busseau I., Institut de Genetique Humaine,
            Centre National de la Recherche Scientifique, 141 rue de la
            Cardonille, Montpellier, 34396 cedex 5, FRANCE
COMMENT     12-5384: You non-LTR retrotransposon.
FEATURES             Location/Qualifiers
     source          1. .5395
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /chromosome="X"
                     /map="1F"
     repeat_region   1. .5395
                     /transposon="You"
     misc_feature    1. .11
                     /note="target site"
     5'UTR           12. .191
     gene            192. .1493
                     /gene="gag"
     CDS             192. .1493
                     /gene="gag"
                     /codon_start=1
                     /protein_id="CAC16870.1"
                     /db_xref="GI:11323018"
                     /db_xref="FLYBASE:FBgn0043054"
                     /db_xref="GOA:Q9GP61"
                     /db_xref="UniProt/TrEMBL:Q9GP61"
                     /translation="MAPGPLNLGDNRFASLSASPKAKKKRPFQKLQVDFPDLPETQCK
                     NPRYLVVSSKNSPKTISDYNCFAVHRALKYISKDITSISNLRDGNLLLLVNSREISDK
                     FIKVTSLPGLCDIECKLHNTLNFVKGTIYAPCLINIPENEIVEELKSQNVHSVFKFTK
                     MIDGTPKPFGKVLVTFDRFTLPSKLTVSWHTVKVSEYIPNPMRCKSCQLLGHTSKHCK
                     NPPACVSCNLAPHLPVPCTRIFCANCTGQHPASSPECPQYQTQKQLLHIKTSKKCSFY
                     EARTILKQQQNTTNNPTLTYSTVANQSTASVMLVKKSRNPNTIITNTDLPTSSTKVPAP
                     LSPTYPTSRTNSLNLEENLTSSPSTSDKTDFSYLKKTCIERSRSLRAEANALLGQTYN
                     DDNNTLMTSRSNSNESIASQTSNLQTKINTTDSDTMDDPDS"
     gene            1589. .5272
                     /gene="pol"
     CDS             1589. .5272
                     /gene="pol"
                     /codon_start=1
                     /protein_id="CAC16871.1"
                     /db_xref="GI:11323019"
                     /db_xref="FLYBASE:FBgn0043053"
                     /db_xref="GOA:Q9GP60"
                     /db_xref="UniProt/TrEMBL:Q9GP60"
                     /translation="MTITILQWNIHGIFNNYNELTLLIKDHAPDIVFLQETKLPCNST
                     NFICPKEYSGYFHNFSYNTSAKQGIGVLIKRNVPHTYRNINSSILCSALQLNFEQVIN
                     IVNAYIPPSQIFSSSDISEILQNLNGSTILLGDLNSWSPLWGSPRTNTRGKKIETVIL
                     ENSLIVLNDGSPTHLSTHNTFTHIDISLISPQIAHMCSWSISDNLHGSDHFPITIHIN
                     TPTRPDNTLPLPKYKTDQANWKRFNESCEKSAAYWTVGCLNQQVAQMTKVIRSAANYS
                     IPQTKRVIHKAKVPWWNANLQQLRDQKQRLFSSYKANTNDTNLIRYKKANALFKKAVL
                     SAKRNSLEKPTSKISPVSSTKKVWSDIKRLAGIPPTPFKYIKSNSGTLTGSFDIAEEF
                     ALSWSKYSSDQNFSAEYIRVKNRYLLEPYAIDSLSPSATSLDSNFTLLEIENSVAKAK
                     GKSPGADRVSYPMLKNLSPHLKTKLLDIFNQILTTGKYPHTWRSAIIIPISKPNKPPS
                     DINSYRPISLLSCLGKTLEKIIAQRLTWFIKRHNLISHNQVAFKSNHSTMDALLRIQH
                     FASNALSTKNHVSILATDFERAFDRVGIHAVLCRLERWGIGPRLYNLIKAFMTNRSFR
                     VRINNVTSNSHILHNGIPQGSPLSVVLFMIAIEDINDIVTRHKDIYISLYADDAIIFT
                     KIKNINTVREKFLEILQEINSWGATSGASLAIEKCQTLHICRKQRCNLSDIVFNSRTI
                     KDVNFLKILGITFDPKLLFKQHCQTLRKQLETRFNIIKFLSSKYSYIHIKTLIDITRA
                     LMLSKIDYGLPIFGWCAKSHLKKLQVPYHGAVRRAIHAFPTSPVACTLAESGLPSIQS
                     RVEETTLMLIPKLYTTSNCLLTKDFGAIFKQKRKFKCISTLRRCANYIKLLDLPLPKP
                     RRPFKSPALWGSKQPNINLQIYNAAKKDTGRLEYQKRFMSAQEDLGVKNWIYTDGSKV
                     TGATTFAVVDSNRKIIAGGRLPSYNSIFTAEAFAILKACQFASKNAGKSVICTDSLSS
                     LSAIRNWNHNDPTTQEVRHILSSHPKKITLLWVPSHQGIHGNELADKAAQEMRLTPSI
                     LFTPFNSKDLKSRIKLYLKEKKLSEWALFMHRYQSINPNCIMFKPPTNVHKRECATFI
                     RLRIGHTQSTHQHLLMRSARPTCQLCGDELTVDHILNACSQLHSIRSHLFGTHSLSNC
                     LSIPSCENISKIYKFVQKAKFII"
     3'UTR           5273. .5384
     misc_feature    5385. .5395
                     /note="target site"
ORIGIN

Alignment Scores:
Pred. No.:                338              Length:          5395
Score:                    39.00            Matches:         6
Percent Similarity:       100.0%           Conservative:    2
Best Local Similarity:    75.0%            Mismatches:      0
Query Match:              86.7%            Indels:          0
DB:                       2                Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x DME302712 (1-5395)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||::::::|||||||||||||
Db       1673 CCCGATATTGTATTTTTGCAGGAA 1696

RESULT 75
AY183922/c
LOCUS       AY183922                7568 bp    DNA     linear   INV 27-FEB-2003
DEFINITION  Drosophila melanogaster centromeric satellite sequence containing
            transposon You.
ACCESSION   AY183922
VERSION     AY183922.1  GI:28193336
KEYWORDS    .
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 7568)
  AUTHORS   Sun,X., Le,H.D., Wahlstrom,J.M. and Karpen,G.H.
  TITLE     Sequence analysis of a functional Drosophila centromere
  JOURNAL   Genome Res. 13 (2), 182-194 (2003)
   PUBMED   12566396
REFERENCE   2  (bases 1 to 7568)
  AUTHORS   Sun,X., Le,H.D., Wahlstrom,J.M. and Karpen,G.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2002) MCBL, The Salk Institute, 10010 N. Torrey
            Pines Rd., La Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .7568
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /map="centromere"
                     /note="y;ry-506; Dp(1;f)1230"
     repeat_region   join(1. .679,6013. .6962)
                     /note="satellite"
                     /rpt_type=tandem
                     /rpt_unit="aatat"
```

```
     repeat_region   680. .6012
                     /transposon="You"
     repeat_region   6963. .7568
                     /note="satellite"
                     /rpt_type=tandem
                     /rpt_unit="ttctc"
ORIGIN
```

Alignment Scores:

Pred. No.:	482	Length:	7568
Score:	39.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	86.7%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AY183922 (1-7568)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||::::::||||||||||||
Db       4386 CCCGATATTGTATTTTTGCAGGAA 4363
```

Search completed: March 10, 2006, 22:19:39
Job time : 540.898 secs

THIS PAGE BLANK (USPTO

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 15:40:34 ; Search time 6079.65 Seconds
(without alignments)
1746.920 Million cell updates/sec

Title: US-10-757-745-2_COPY_121_347
Perfect score: 1192
Sequence: 1 IDGLDLNNLSERARGVCSYL.....HIIPRSLDLLGLEKLDCGRF 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170329_4058/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_12397_@runat_09032006_170329_4058 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1192	100.0	1089	11	DQ049205	DQ049205 Homo sapi
	2	1192	100.0	1168	4	CR601303	CR601303 full-leng
	3	1192	100.0	1620	4	CR592636	CR592636 full-leng
	4	1192	100.0	1743	4	CR602029	CR602029 full-leng
	5	1192	100.0	1894	4	CR597293	CR597293 full-leng
	6	1192	100.0	1909	4	CR595644	CR595644 full-leng
c	7	1189	99.7	952	5	BX433489	BX433489 BX433489
	8	1188	99.7	1089	11	DQ049206	DQ049206 Pan trogl
	9	1177	98.7	965	7	CN803230	CN803230 ILLUMIGEN
	10	1177	98.7	1087	7	CN642527	CN642527 ILLUMIGEN
	11	1131	94.9	942	5	BX358707	BX358707 BX358707
	12	1123	94.2	1046	3	BM926092	BM926092 AGENCOURT
	13	1122	94.1	883	5	BU179107	BU179107 AGENCOURT
	14	1110	93.1	749	7	CR767800	CR767800 DKFZp469C
	15	1105	92.7	972	3	BM468826	BM468826 AGENCOURT
	16	1087.5	91.2	1090	7	CN641671	CN641671 ILLUMIGEN
	17	1079	90.5	680	7	CK952220	CK952220 4091710 B
	18	1055	88.5	1081	5	BX444691	BX444691 BX444691
	19	1051	88.2	770	2	BG498689	BG498689 602544301
	20	1043	87.5	650	2	BG431303	BG431303 602499879
	21	1037	87.0	922	7	CN801806	CN801806 ILLUMIGEN
	22	1032	86.6	904	5	BX422491	BX422491 BX422491
	23	1017	85.3	581	3	BP262101	BP262101 BP262101
	24	1010	84.7	582	3	BP261789	BP261789 BP261789
	25	997	83.6	981	5	BX338160	BX338160 BX338160
	26	980	82.2	577	1	AW964871	AW964871 EST376839
	27	972	81.5	909	5	BU928818	BU928818 AGENCOURT
	28	971.5	81.5	932	2	BG740339	BG740339 602635289
	29	955	80.1	571	6	CB125492	CB125492 K-EST0174
	30	955	80.1	1067	5	BX337905	BX337905 BX337905
	31	950	79.7	546	2	BG541031	BG541031 602570190
	32	950	79.7	570	3	BP258309	BP258309 BP258309
	33	949	79.6	1110	3	BM476771	BM476771 AGENCOURT
	34	948	79.5	1018	5	BX337141	BX337141 BX337141
c	35	946	79.4	559	2	BE162174	BE162174 IL2-HT044
	36	941	78.9	846	7	CR765451	CR765451 DKFZp469E
	37	938	78.7	534	7	CN298921	CN298921 170005315
	38	921	77.3	870	2	BG740396	BG740396 602634171
c	39	906	76.0	811	8	DN122554	DN122554 1122781 M
	40	887	74.4	841	2	BG562430	BG562430 602581056
c	41	883	74.1	860	8	CX958090	CX958090 JGI_CAAO9
c	42	883	74.1	891	1	AL879624	AL879624 AL879624
c	43	883	74.1	894	5	BX776016	BX776016 BX776016
c	44	883	74.1	909	5	BX780455	BX780455 BX780455
	45	869	72.9	1289	4	AK009089	AK009089 Mus muscu
c	46	868	72.8	835	8	DR867245	DR867245 JGI_CABG9
	47	837	70.2	815	5	BU936308	BU936308 AGENCOURT
	48	835	70.1	760	5	BU204728	BU204728 604152621
	49	832.5	69.8	870	5	BU159911	BU159911 AGENCOURT
	50	829.5	69.6	948	1	AL555333	AL555333 AL555333
	51	829	69.5	1192	3	BM553049	BM553049 AGENCOURT
	52	828	69.5	617	7	CN694288	CN694288 E0348E09-
	53	826	69.3	578	1	AV602300	AV602300 AV602300
	54	822	69.0	883	5	BU169945	BU169945 AGENCOURT
	55	821	68.9	724	6	CB443087	CB443087 693950 MA
	56	821	68.9	810	8	CX166335	CX166335 HESC2_39_
c	57	819	68.7	850	6	CF378639	CF378639 AGENCOURT
c	58	809	67.9	737	3	BJ639656	BJ639656 BJ639656
	59	807	67.7	614	2	BG695251	BG695251 NISC_iv14
	60	806	67.6	580	3	BM503555	BM503555 ih33e07.y
	61	806	67.6	827	8	DN097147	DN097147 JGI_CABE7
	62	806	67.6	984	5	BX846715	BX846715 BX846715
c	63	802	67.3	749	5	BX753763	BX753763 BX753763
	64	792	66.4	919	6	CA973696	CA973696 AGENCOURT
	65	790	66.3	757	3	BI760756	BI760756 603044763
	66	790	66.3	1103	3	BM555041	BM555041 AGENCOURT
	67	788	66.1	900	6	CA980929	CA980929 AGENCOURT
	68	780.5	65.5	927	2	BG119064	BG119064 602347589
	69	777	65.2	896	6	CA971409	CA971409 AGENCOURT
	70	772	64.8	802	8	DN100293	DN100293 JGI_CABE8
	71	761	63.8	608	7	CK966606	CK966606 4081935 B
	72	758	63.6	581	3	BP349936	BP349936 BP349936
	73	756	63.4	826	5	BU943127	BU943127 AGENCOURT
c	74	752	63.1	714	2	BG070243	BG070243 H3085C02-
	75	751	63.0	565	3	BP226464	BP226464 BP226464

ALIGNMENTS

RESULT 1

```
DQ049205
LOCUS       DQ049205               1089 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Homo sapiens TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ049205
VERSION     DQ049205.1  GI:66902404
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1089)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 1089)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.  Translation starts at the beginning of
            alignment.
FEATURES             Location/Qualifiers
     source          1. .1089
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
     gene            <1. .>1089
                     /gene="TTRAP"
                     /locus_tag="HC17203"
ORIGIN

Alignment Scores:
Pred. No.:              4.84e-133       Length:         1089
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x DQ049205 (1-1089)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 660

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 720

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 780

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 960

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTT 1041

RESULT 2
CR601303
LOCUS       CR601303               1168 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DN005YN02 of Adult brain of Homo sapiens
            (human).
ACCESSION   CR601303
VERSION     CR601303.1  GI:50482110
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1168)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1168)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1168
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="Adult brain"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Alignment Scores:
Pred. No.:              5.37e-133       Length:         1168
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
```

```
DB:                    4              Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x CR601303 (1-1168)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066

RESULT 3
CR592636
LOCUS       CR592636               1620 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DM007YE11 of Fetal liver of Homo sapiens
            (human).
ACCESSION   CR592636
VERSION     CR592636.1  GI:50473443
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1620)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1620)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1620
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DM007YE11"
                     /tissue_type="Fetal liver"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Alignment Scores:
Pred. No.:              8.77e-133       Length:         1620
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     4               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CR592636 (1-1620)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 363

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 423

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 483

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 543

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 603

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 663

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 723

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 783

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 843

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        844 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 903

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 963
```

```
Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||
Db        964 AAACTGGACTGTGGTAGATTT 984

RESULT 4
CR602029
LOCUS       CR602029               1743 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DI042YL19 of Placenta Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR602029
VERSION     CR602029.1  GI:50482836
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1743)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1743)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1743
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI042YL19"
                     /tissue_type="Placenta Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Alignment Scores:
Pred. No.:              9.79e-133       Length:         1743
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     4               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CR602029 (1-1743)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 289

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 349

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 409

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 469

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 529

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        530 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 589

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        590 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 649

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        650 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 709

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        710 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 769

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        770 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 829

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        830 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 889

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||
Db        890 AAACTGGACTGTGGTAGATTT 910

RESULT 5
CR597293
LOCUS       CR597293               1894 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DK007YK05 of HeLa cells Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR597293
VERSION     CR597293.1  GI:50478100
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1894)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1894)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK007YK05"
                     /tissue_type="HeLa cells Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN

Alignment Scores:
```

Pred. No.:	1.11e-132	Length:	1894
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x CR597293 (1-1894)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 438

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 498

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 558

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 618

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 678

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 738

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 798

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 858

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 918

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 978

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1038

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1039 AAACTGGACTGTGGTAGATTT 1059
```

RESULT 6
CR595644

LOCUS	CR595644 1909 bp mRNA linear HTC 21-JUL-2004
DEFINITION	full-length cDNA clone CS0DI052YN13 of Placenta Cot 25-normalized of Homo sapiens (human).
ACCESSION	CR595644
VERSION	CR595644.1 GI:50476451
KEYWORDS	HTC; CNSLT_cDNA.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1909)
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue
REFERENCE	2 (bases 1 to 1909)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES	Location/Qualifiers
source	1. .1909 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DI052YN13" /tissue_type="Placenta Cot 25-normalized" /plasmid="pCMVSPORT_6"
ORIGIN	

Alignment Scores:

Pred. No.:	1.12e-132	Length:	1909
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x CR595644 (1-1909)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 444

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 504

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 564

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 624

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 684

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 744

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 804

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 864

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 924

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        925 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 984

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1044

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1045 AAACTGGACTGTGGTAGATTT 1065

RESULT 7
BX433489/c
LOCUS       BX433489                952 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX433489 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN005YN02 3-PRIME, mRNA sequence.
ACCESSION   BX433489
VERSION     BX433489.2  GI:47002503
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 952)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30775195.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAI056ZD12_CS05372_1&c=3474.r
            .
FEATURES             Location/Qualifiers
     source          1. .952
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

Alignment Scores:
Pred. No.:                9.13e-133       Length:          952
Score:                    1189.00         Matches:         226
Percent Similarity:       100.0%          Conservative:    1
Best Local Similarity:    99.6%           Mismatches:      0
Query Match:              99.7%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BX433489 (1-952)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 724

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        723 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 664

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 604

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 544

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 484

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 424

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db        423 TTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 364

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 304

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 244

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 184

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 124

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db        123 AAACTGGACTGTGGTAGATTT 103

RESULT 8
DQ049206
LOCUS       DQ049206                1089 bp    DNA     linear   GSS 02-JUN-2005
DEFINITION  Pan troglodytes TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   DQ049206
VERSION     DQ049206.1  GI:66902405
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 1089)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
  JOURNAL   (er) PLoS Biol. 3 (6), E170 (2005)
   PUBMED   15869325
REFERENCE   2  (bases 1 to 1089)
  AUTHORS   Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.  Translation starts at the beginning of
            alignment.
```

```
FEATURES             Location/Qualifiers
     source          1. .1089
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>1089
                     /gene="TTRAP"
                     /locus_tag="HC17203"
ORIGIN

Alignment Scores:
Pred. No.:                1.48e-132      Length:          1089
Score:                    1188.00        Matches:         226
Percent Similarity:       99.6%          Conservative:    0
Best Local Similarity:    99.6%          Mismatches:      1
Query Match:              99.7%          Indels:          0
DB:                       11             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x DQ049206 (1-1089)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 420

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 480

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 540

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 600

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 660

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db        661 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGTGGAACGAATGAATCAG 720

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAATCAGCTACAGTTATATTTGCA 780

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 840

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 900

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAGCTTCGTTTTGATCGAATATTTTTC 960

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1020

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1021 AAACTGGACTGTGGTAGATTT 1041

RESULT 9
CN803230
LOCUS       CN803230                 965 bp    mRNA    linear   EST 26-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_32551 Katze_MMBR Macaca mulatta cDNA clone
            IBIUW:12605 5' similar to Bases 1 to 965 highly similar to human
            TTRAP (Hs.210628), mRNA sequence.
ACCESSION   CN803230
VERSION     CN803230.1  GI:47699206
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 965)
  AUTHORS   Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
  TITLE     Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
  JOURNAL   Genome Biol. 6 (7), R60 (2005)
   PUBMED   15998449
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.04.02. 780 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 965    Std Error: 0.00
            Plate: CL000288  row: D  column: 11
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .965
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:12605"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

Alignment Scores:
Pred. No.:                2.64e-131      Length:          965
Score:                    1177.00        Matches:         223
Percent Similarity:       100.0%         Conservative:    4
Best Local Similarity:    98.2%          Mismatches:      0
Query Match:              98.7%          Indels:          0
DB:                       7              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x CN803230 (1-965)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 187

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 GCGTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 247

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::|||:::||||||||||||||||||||||||||||||||||||||||
Db        248 CTAAAGAAGAGAGCAAGTGATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 307

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 367
```

```
Qy          81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         368 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 427

Qy         101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         428 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 487

Qy         121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         488 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 547

Qy         141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
               ||||||||||||||||||||:::||||||||||:::||||||||||||||||||||||||
Db         548 GGAGATACAAATCTAAGGGATCAAGAGGTTACCAAATGTGGTGGTTTACCCAACAACATT 607

Qy         161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         608 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 667

Qy         181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         668 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAGCTTCGTTTTGATCGAATATTTTTC 727

Qy         201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         728 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTGGAA 787

Qy         221 LysLeuAspCysGlyArgPhe 227
               |||||||||||||||||||||
Db         788 AAACTCGACTGTGGTAGATTT 808
```

RESULT 10
CN642527

```
LOCUS       CN642527                1087 bp    mRNA    linear   EST 12-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_6530 Katze_MMPL2 Macaca mulatta cDNA clone IBIUW:4699
            5' similar to Bases 1 to 1084 highly similar to human TTRAP
            (Hs.210628), mRNA sequence.
ACCESSION   CN642527
VERSION     CN642527.1  GI:47153537
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 1087)
  AUTHORS   Magness,C.L., Pellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
  TITLE     Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
  JOURNAL   Genome Biol. 6 (7), R60 (2005)
   PUBMED   15998449
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2003.12.24. 840 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 1087   Std Error: 0.00
            Plate: CL000066  row: D  column: 10
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .1087
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:4699"
                     /sex="male"
                     /dev_stage="newborn infant"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMPL2"
                     /note="Organ: placenta; Vector: Uni-ZAP XR; Site_1: EcoR
                     I; Site_2: Xho I; Created from Stratagene ZAP-cDNA
                     Synthesis kit (catalog #200400)  and ZAP-CDNA Gigapack III
                     Gold Cloning Kit (Catalog #200450)"
ORIGIN
```

Alignment Scores:

Pred. No.:	3.16e-131	Length:	1087
Score:	1177.00	Matches:	223
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	98.2%	Mismatches:	0
Query Match:	98.7%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x CN642527 (1-1087)

```
Qy           1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          80 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 139

Qy          21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         140 GCGTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 199

Qy          41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
               ||||||||||||:::||:::||||||||||||||||||||||||||||||||||||||||||
Db         200 CTAAAGAAGAGAGCAAGTGATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 259

Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         260 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 319

Qy          81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         320 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 379

Qy         101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         380 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 439

Qy         121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         440 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 499

Qy         141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
               ||||||||||||||||||||:::||||||||||:::||||||||||||||||||||||||
Db         500 GGAGATACAAATCTAAGGGATCAAGAGGTTACCAAATGTGGTGGTTTACCCAACAACATT 559

Qy         161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         560 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 619

Qy         181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         620 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAGCTTCGTTTTGATCGAATATTTTTC 679

Qy         201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         680 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTGGAA 739

Qy         221 LysLeuAspCysGlyArgPhe 227
               |||||||||||||||||||||
Db         740 AAACTCGACTGTGGTAGATTT 760
```

RESULT 11
BX358707

```
LOCUS       BX358707                942 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX358707 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI042YL19 5-PRIME, mRNA sequence.
ACCESSION   BX358707
VERSION     BX358707.2  GI:46306560
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 942)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 5, 2003 this sequence version replaced gi:30378261.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI042CF10QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .942
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI042YL19"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:                9.49e-126       Length:         942
Score:                    1131.00         Matches:        221
Percent Similarity:       98.2%           Conservative:   0
Best Local Similarity:    98.2%           Mismatches:     4
Query Match:              94.9%           Indels:         2
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BX358707 (1-942)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 289

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 349

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 409

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 469

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 529

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        530 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 589

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        590 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 649

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        650 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 709

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        710 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 769

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        770 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 829

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        830 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACHT-CTTGGAWTAGAA 888

Qy        221 LysLeuAspCysGly 225
              |||||||||||||
Db        889 AAAWG-GGMTGTGGA 902

RESULT 12
BM926092
LOCUS       BM926092                1046 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6649780 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764440
            5', mRNA sequence.
ACCESSION   BM926092
VERSION     BM926092.1  GI:19376471
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1046)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12817  row: o  column: 01
            High quality sequence start: 4
            High quality sequence stop: 632.
FEATURES             Location/Qualifiers
     source          1. .1046
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5764440"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_114"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is destroyed
                     upon cloning). Average insert size 1.5 kb, insert size
                     range 1-3 kb. Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen). Research Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
```

```
ORIGIN

Alignment Scores:
Pred. No.:                1.03e-124       Length:          1046
Score:                    1123.00         Matches:          218
Percent Similarity:       96.9%           Conservative:       2
Best Local Similarity:    96.0%           Mismatches:         6
Query Match:              94.2%           Indels:             1
DB:                       3               Gaps:               0

US-10-757-745-2_COPY_121_347 (1-227) x BM926092 (1-1046)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 293

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 353

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 413

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 473

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 533

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 593

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db        594 TTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 653

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 713

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||      ||||||||||||||||||||||||||||||||||||||||||||
Db        714 GTGGATGTCTGGGGAGTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 773

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 833

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp-LeuLeuGlyLeuGl 220
              ||||||||||||||||||||||||||||||||||||||    |||||| |||||||||
Db        834 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGGTTGGACCCTTCTTGGATAANA 893

Qy        220 uLysLeuAspCysGlyArg 226
              |||||||||    |||:::
Db        894 AAAACTGGACTGGGGGAAA 912

RESULT 13
BU179107
LOCUS       BU179107                 883 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7984768 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6174958
            5', mRNA sequence.
ACCESSION   BU179107
VERSION     BU179107.1  GI:22693091
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13548  row: o  column: 23
            High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6174958"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

Alignment Scores:
Pred. No.:                1.06e-124       Length:          883
Score:                    1122.00         Matches:         219
Percent Similarity:       98.7%           Conservative:      1
Best Local Similarity:    98.2%           Mismatches:        1
Query Match:              94.1%           Indels:            2
DB:                       5               Gaps:              0

US-10-757-745-2_COPY_121_347 (1-227) x BU179107 (1-883)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 278

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 338

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 398

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 458

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 518

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 578

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 638

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 698

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        699 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 758

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg-PheAspArgIlePhePh 200
              ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        759 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTTGATCGAATATTTTT 818

Qy        200 eArgAlaAlaAlaGluGluGlyHisIleIle-ProArgSerLeuAspLeuLeuGlyLeuG 220
              |||||||||||||||||||||||||   ||| ||||||||||||||||||||||||||:
Db        819 CCGAGCAGCAGCAGAAGAGGGACACTTTATTCCCCCGAAGTTTGGACCTTCTTGGATTAA 878

Qy        220 luLys 221
              ::|||
Db        879 AAAAA 883

RESULT 14
CR767800
LOCUS       CR767800                749 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp469C058_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469C058 5', mRNA sequence.
ACCESSION   CR767800
VERSION     CR767800.1  GI:52610302
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,
            Deutschenbaur,S., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix
            (Martinsried/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp469C058) is available at
            the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
            Berlin, Germany. Please contact RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469C058
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp469C058"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="469 (synonym: pkid1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:              2.35e-123       Length:         749
Score:                  1110.00         Matches:        211
Percent Similarity:     99.1%           Conservative:   3
Best Local Similarity:  97.7%           Mismatches:     2
Query Match:            93.1%           Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CR767800 (1-749)

Qy         12 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 31
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 AGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAG 66

Qy         32 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 51
              |||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db         67 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGTTCAAGTAATTATGAGATTATT 126

Qy         52 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 71
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 ACAGGTCTTGAAGAAGGGTATTTCACAGCTATCATGTTGAAGAAATCAAGAGTGAAATTA 186

Qy         72 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTG 246

Qy         92 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 111
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 CATGTGAGTGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 306

Qy        112 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCT 366

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CCAGAGTCAGCTACAGTTATATTTGCAGGAGATACGAATCTAAGGGATCGAGAGGTTACC 426

Qy        152 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 AGATGTGGTGGTTTACCCAACAACATCGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAA 486

Qy        172 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCys 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGT 546

Qy        192 LysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        547 AAGCTTCGTTTTGATCGAATATTTTTCAGAGCAGCCGCAGAAGAGGGACACATTGTTCCC 606

Qy        212 ArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              :::|||||||||||||||||||||||||||||||||||||||||||||
Db        607 CAAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 654

RESULT 15
BM468826
LOCUS       BM468826                972 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6445782 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5587154
            5', mRNA sequence.
ACCESSION   BM468826
VERSION     BM468826.1  GI:18517868
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 972)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12356  row: d  column: 03
            High quality sequence stop: 707.
FEATURES             Location/Qualifiers
     source          1. .972
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

```
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5587154"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_92"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.4e-122       Length:         972
Score:                  1105.00        Matches:        221
Percent Similarity:     96.9%          Conservative:   1
Best Local Similarity:  96.5%          Mismatches:     3
Query Match:            92.7%          Indels:         4
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BM468826 (1-972)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 315

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 375

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 435

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 495

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 555

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 615

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 TTAAAAATGGTTTTGAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 675

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 735

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 795

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCys-LysLeuArgPheAspArgIlePhePh 200
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        796 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAAACTTCGTTTTGATCGAATATTTTT 855

Qy        200 eArgAlaAlaAlaGlu-GluGlyHisIleIleProArgSerLeu-AspLeuLeuGlyLeu 219
              |||||||||||||||| |||||||||||||||   |||:::||| |||        ||||||
Db        856 CAGAGCAGCAGCAGAAAGAGGGACACATTATTTCCCGGAATTTGGGACCCTTCTGGATTA 915

Qy        220 Glu-LysLeuAspCysGly 225
              ||| ||||||||||||||
Db        916 GAAAAAACTGGACTGTGGG 934

RESULT 16
CN641671
LOCUS       CN641671                1090 bp    mRNA    linear   EST 12-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_5159 Katze_MMBR Macaca mulatta cDNA clone IBIUW:5970
            5' similar to Bases 11 to 990 highly similar to human TTRAP
            (Hs.210628), mRNA sequence.
ACCESSION   CN641671
VERSION     CN641671.1  GI:47152681
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 1090)
  AUTHORS   Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
  TITLE     Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
  JOURNAL   Genome Biol. 6 (7), R60 (2005)
   PUBMED   15998449
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2003.12.04. 732 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 1090   Std Error: 0.00
            Plate: CL000043  row: E  column: 07
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .1090
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:5970"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

Alignment Scores:
Pred. No.:              2.19e-120      Length:         1090
Score:                  1087.50        Matches:        212
Percent Similarity:     95.2%          Conservative:   4
Best Local Similarity:  93.4%          Mismatches:     9
Query Match:            91.2%          Indels:         2
DB:                     7              Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x CN641671 (1-1090)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 292

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GCGTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCTTATTATAGCTAC 352

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||:::|||:::||||||||||||||||||||||||||||||||||||||||
Db        353 CTAAAGAAGAGAGCAAGTGATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 412

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
```

```
Db        413 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 472

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100

Db        473 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 532

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120

Db        533 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 592

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140

Db        593 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 652

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db        653 GGAGATACAAATCTAAGGGATCAAGAGGTTACCAAATGTGGTGGTTTACCCAACAACATT 712

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180

Db        713 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 772

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200

Db        773 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAGCTTCCTTTTGATCGAATATTTTTC 832

Qy        201 ArgAlaAlaAlaGluGlu-GlyHisIleIleProArgSerLeu---AspLeuLeuGlyLe 219

Db        833 AGAGCAGCAGCAGAAGAAGGGACACATTATTCCCCCGAAGTTTGGAACCTTCCTTGGATT 892

Qy        219 uGluLysLeuAspCysGly 225

Db        893 GGAAAAAACTCCGACTGGG 911
```

```
RESULT 17
CK952220
LOCUS       CK952220                680 bp    mRNA    linear   EST 15-MAR-2004
DEFINITION  4091710 BARC 10BOV Bos taurus cDNA clone 10BOV29_F21 5', mRNA
            sequence.
ACCESSION   CK952220
VERSION     CK952220.1  GI:45466600
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Sonstegard, T.S., Van Tassell, C.P., Matukumalli, L.K., Harhay,
            G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
  TITLE     Production of EST from cDNA libraries derived from immunologically
            activated bovine gut
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Tad S. Sonstegard
            Bovine Functional Genomics Laboratory
            Animal and Natural Resources Institute
            Bdlg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
            Tel: 3015048416
            Fax: 3015048414
            Email: tads@anri.barc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            0.000925 using options -trim_alt " -trim_fasta. Vector identified
            by cross_match using options -minmatch 12 -minscore 18
            Plate: 29  row: F  column: 21
            Seq primer: CCCAGTCACGACGTTGTAAAACG
            High quality sequence stop: 680.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /strain="Holstein"
                     /db_xref="taxon:9913"
                     /clone="10BOV29_F21"
                     /sex="Male"
                     /tissue_type="Pooled"
                     /dev_stage="Multiple"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="BARC 10BOV"
                     /note="Organ: Small Intestine; Vector: pAgen-1; Site_1:
                     EcoRV; Site_2: NotI; Equimolar amounts of mRNA extracted
                     from proximal jejunums of 18 and 21 wk old steers, and
                     distal ileums of 14 day old calves. proximal jejunum
                     exposed to C. oncophora for 3 and 6 weeks, and distal
                     ileum exposed to C. parvum for 7 days"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              1.15e-119       Length:         680
Score:                  1079.00         Matches:        202
Percent Similarity:     96.0%           Conservative:   14
Best Local Similarity:  89.8%           Mismatches:     9
Query Match:            90.5%           Indels:         0
DB:                     7               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CK952220 (1-680)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20

Db          4 ATTGATGGATTGGACATGAACAATCTCCTAGAGAGGGCTCGAGGGGTGTGTTCCTATTTA 63

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40

Db         64 ACTTTGTACAGCCCTGATGTGATATTTCTACAGGAAGTCATTCCCCCATATTATGCCTAC 123

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60

Db        124 CTAAAGAAGAAAGCAAGTAGTTACAAGATTATTACAGGTCGTGAAGAAGGATATTTCACA 183

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80

Db        184 GCTATAATGTTGAAGAAATCAAGAGTGAAGTTTAAAAGCCAAGAAATTATTCCTTTTCCA 243

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100

Db        244 AATACGCAAATGATGAGAAACCTTTTGTGTGTTCATGTGAGTGTGTCAGGAAATGAACTT 303

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120

Db        304 TGCCTTATGACTTCTCATTTGGAAAGCACCAGAGGGCACGCTAAGGAACGAATGAATCAG 363

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140

Db        364 TTGAAAATGGTTTTAGAGAAAATGCAAGAGGCTCCAGGATCAGCTACAGTTATATTTGCT 423

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db        424 GGAGATACGAATCTCAGGGATCAAGAAGTTACCAAATGTGGTGGTTTACCCAACAACATT 483

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180

Db        484 TTGGATGTCTGGGAGTTTTTGGGCAAACCTAAGCATTGCCAGTATACATGGGATACACAA 543

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200

Db        544 ATGAACTCTAATCTTGGAATAGCTGCTAATTGTAAACTTCGTTTTGATCGAATATTTTTT 603

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220

Db        604 AGAGCTGCAGCAGAAGGTGGCCACATTATTCCCCAAAGTTTGGACCTGCTTGGGCTGGAA 663

Qy        221 LysLeuAspCysGly 225

Db        664 AAACTGGACTGTGGT 678

RESULT 18
```

```
BX444691
LOCUS       BX444691                1081 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX444691 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN005YN02 5-PRIME, mRNA sequence.
ACCESSION   BX444691
VERSION     BX444691.2  GI:47009162
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1081)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30780264.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DN005DG01QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .1081
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

Alignment Scores:
Pred. No.:                1.86e-116      Length:          1081
Score:                    1055.00        Matches:         218
Percent Similarity:       96.5%          Conservative:    1
Best Local Similarity:    96.0%          Mismatches:      8
Query Match:              88.5%          Indels:          4
DB:                       5              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BX444691 (1-1081)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 446

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 506

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 566

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 626

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        627 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 686

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 746

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        747 TTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 806

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCMAC-AACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAA-CCTAAACATTGCCAGTATACATGGGATACACAC 924

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 984

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||  ||||||||||||||||||||||||||||  |||||||||  |||||||||||||
Db        985 AGARCAGCAGCAGAAGAGGGACACATTATYCCCGA-AGTTTGGACCYTCTTGGATTAGA- 1042

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||  ||||||
Db       1043 AAACTGGACTGTWGTAGATTY 1063

RESULT 19
BG498689
LOCUS       BG498689                  770 bp    mRNA    linear   EST 27-MAR-2001
DEFINITION  602544301F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4666633 5',
            mRNA sequence.
ACCESSION   BG498689
VERSION     BG498689.1  GI:13460206
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 770)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1474  row: a  column: 02
            High quality sequence stop: 668.
FEATURES             Location/Qualifiers
     source          1. .770
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4666633"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_60"
                     /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI
                     (ggccattatggcc); Double-stranded cDNA was prepared from
                     cell line RNA.  5' and 3' adaptors were used in cloning as
                     follows: 5' adaptor sequence:  5'-CACGGCCATTATGGCC-3' and
                     3' adaptor sequence:
```

5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	3.42e-116	Length:	770
Score:	1051.00	Matches:	207
Percent Similarity:	98.6%	Conservative:	1
Best Local Similarity:	98.1%	Mismatches:	0
Query Match:	88.2%	Indels:	3
DB:	2	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x BG498689 (1-770)

```
Qy         20 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSer 39
              |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTTAGCTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGC 61

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 59
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTC 121

Qy         60 ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 79
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTT 181

Qy         80 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 99
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAG 241

Qy        100 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 119
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAAT 301

Qy        120 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 139
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTT 361

Qy        140 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsn 159
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAAC 421

Qy        160 IleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThr 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACA 481

Qy        180 GlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe 199
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTCGATCGAATATTT 541

Qy        200 PheArgAlaAlaAlaGluGluGlyHisIleIleProArg-SerLeuAspLeuLeuGly-L 219
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||| |
Db        542 TTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGGAAGTTTGGACCTTCTTGGCAT 601

Qy        219 euGluLysLeuAsp-CysGlyArgPhe 227
              |||||||||||||| ||||||||||||
Db        602 TAGAAAAACTGGACCTGTGGTAGATTT 628
```

RESULT 20
BG431303
LOCUS BG431303 650 bp mRNA linear EST 14-MAR-2001
DEFINITION 602499879F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4613587 5', mRNA sequence.
ACCESSION BG431303
VERSION BG431303.1 GI:13337809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 650)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1363 row: n column: 20
High quality sequence stop: 648.
FEATURES Location/Qualifiers
source 1. .650
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4613587"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_75"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	2.47e-115	Length:	650
Score:	1043.00	Matches:	202
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	87.5%	Indels:	1
DB:	2	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x BG431303 (1-650)

```
Qy         25 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGA 61

Qy         45 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTG 121

Qy         65 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATG 181

Qy         85 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACA 241

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTT 301

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAAT 361

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        362 CTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGG 421

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAAT 481

Qy        185 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCA 541

Qy        205 GluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCys 224
              |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        542 GAAGAGGGACACATTATTCCCCGAAG-TTGGACCTTCTTGGATTAGAAAAACTGGACTGT 600

Qy        225 GlyArgPhe 227
              |||||||||
Db        601 GGTAGATTT 609
```

```
RESULT 21
CN801806
LOCUS       CN801806                922 bp    mRNA    linear   EST 26-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_37066 Katze_MMBR Macaca mulatta cDNA clone
            IBIUW:15450 5' similar to Bases 18 to 922 highly similar to human
            TTRAP (Hs.210628), mRNA sequence.
ACCESSION   CN801806
VERSION     CN801806.1  GI:47697782
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 922)
  AUTHORS   Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
  TITLE     Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
  JOURNAL   Genome Biol. 6 (7), R60 (2005)
  PUBMED    15998449
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.05.13. 514 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 922   Std Error: 0.00
            Plate: CL000291  row: H  column: 09
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .922
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:15450"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

Alignment Scores:
Pred. No.:            2.22e-114      Length:        922
```

```
Score:                 1037.00       Matches:       200
Percent Similarity:    93.7%         Conservative:  7
Best Local Similarity: 90.5%         Mismatches:    14
Query Match:           87.0%         Indels:        0
DB:                    7             Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x CN801806 (1-922)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db        243 ATTGATGGATTAAATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 302

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GCGTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 362

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::|||:::|||||||||||||||||||||||||||||||||||||||
Db        363 CTAAAGAAGAGAGCAAGTGATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 422

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 482

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 542

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 602

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||||
Db        603 TTAAAAATGGTTTTAAAGAGAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 662

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              |||||||||   |||||||||:::|||||||||:::||||||   ||||||||||||||||
Db        663 GGAGATACATATCTAAGGGATCAAGAGGTTACCAAATGTGGTCGTTTACCCAACAACATT 722

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||   ||||||||||||   ||||||||||||||||||||||
Db        723 GTGGATGTCTGGGAGTTTTTGAGCAAACCTAAACATTTGCAGTATACATGGGATACATCA 782

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||   |||   ||||||||||||||||||||||||||||||||||||
Db        783 ATGAACTCTAATCTGGAAATACCTGCTGCTTGTAAGCTTCGTTTTGATCGAATATTTTTC 842

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              :::|||       |||   |||   |||   ||||||||||||      ||||||||||||
Db        843 AAAGCACCAACAGAACGAGGAACCATTTATCCCCGAAGCTTGTACTTTCTTGGATTGGAA 902

Qy        221 Lys 221
              |||
Db        903 AAA 905
```

```
RESULT 22
BX422491
LOCUS       BX422491                904 bp    mRNA    linear   EST 03-MAY-2004
DEFINITION  BX422491 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
            CS0DM007YE11 5-PRIME, mRNA sequence.
ACCESSION   BX422491
VERSION     BX422491.2  GI:46955237
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
```

```
COMMENT     On May 15, 2003 this sequence version replaced gi:30766188.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DM007AC06QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .904
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DM007YE11"
                     /tissue_type="FETAL LIVER"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL LIVER"
                     /note="Organ: liver; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

Alignment Scores:
Pred. No.:              8.69e-114       Length:         904
Score:                  1032.00         Matches:        196
Percent Similarity:     98.5%           Conservative:   1
Best Local Similarity:  98.0%           Mismatches:     3
Query Match:            86.6%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BX422491 (1-904)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        304 ATTGATGGATTAGATCTAAACAAWCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 363

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 423

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 483

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 543

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 603

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 663

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db        664 TTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 723

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        724 GGAGATACAAATCTAAGGGATCGAGAGGTTACCARATGTGGTGGTTTACCCAACAACATT 783

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 843

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db        844 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAAWATTTTTC 903

RESULT 23
BP262101
LOCUS       BP262101                581 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP262101 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI05893, mRNA sequence.
ACCESSION   BP262101
VERSION     BP262101.1  GI:52177332
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI05893"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:              2.94e-112       Length:         581
Score:                  1017.00         Matches:        193
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            85.3%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BP262101 (1-581)

Qy          5 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGC 60

Qy         25 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGA 120

Qy         45 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTG 180

Qy         65 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATG 240

Qy         85 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACA 300

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 TCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTT 360

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAAT 420

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGG 480

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAAT 540

Qy        185 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArg 197
              |||||||||||||||||||||||||||||||||||||||
Db        541 CTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGA 579

RESULT 24
BP261789
LOCUS       BP261789                 582 bp    mRNA    linear   EST 16-SEP-2004
DEFINITION  BP261789 Sugano cDNA library, small intestine Homo sapiens cDNA
            clone HSI05013, mRNA sequence.
ACCESSION   BP261789
VERSION     BP261789.1  GI:52177020
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
  PUBMED    15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HSI05013"
                     /tissue_type="small intestine"
                     /clone_lib="Sugano cDNA library, small intestine"
ORIGIN

Alignment Scores:
Pred. No.:              2.07e-111       Length:         582
Score:                  1010.00         Matches:        192
Percent Similarity:     99.5%           Conservative:   1
Best Local Similarity:  99.0%           Mismatches:     1
Query Match:            84.7%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BP261789 (1-582)

Qy          5 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 24
              |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATCTAAACAATCTGTGAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGC 60

Qy         25 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGA 120

Qy         45 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||

Db        121 TCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATGATGTTG 180

Qy         65 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATG 240

Qy         85 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACA 300

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTT 360

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAAT 420

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGG 480

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAAT 540

Qy        185 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 198
              ||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATA 582

RESULT 25
BX338160
LOCUS       BX338160                 981 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX338160 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI056YC23 5-PRIME, mRNA sequence.
ACCESSION   BX338160
VERSION     BX338160.2  GI:46283046
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 981)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30345671.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI056AB12QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .981
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI056YC23"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

```
Alignment Scores:
Pred. No.:                1.7e-109        Length:          981
Score:                    997.00          Matches:         198
Percent Similarity:       99.0%           Conservative:    0
Best Local Similarity:    99.0%           Mismatches:      2
Query Match:              83.6%           Indels:          2
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BX338160 (1-981)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 433

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 493

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 553

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 613

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 673

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        674 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 733

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        734 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 793

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        794 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 853

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        854 GTGGATGTCTGGGAGTTTTTGGGCAA-CCTAAACATTGCCAGTATACATGGGATACACAA 912

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        913 ATGAACTCTAATCTTGGAATAMT-GCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 971

RESULT 26
AW964871
LOCUS       AW964871                577 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST376839 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW964871
VERSION     AW964871.1  GI:8154602
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
            Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 206
            Seq primer: Reverse.
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGH"
                     /note="Vector: pBluescriptSKm"
ORIGIN

Alignment Scores:
Pred. No.:                8.79e-108       Length:          577
Score:                    980.00          Matches:         186
Percent Similarity:       97.4%           Conservative:    1
Best Local Similarity:    96.9%           Mismatches:      5
Query Match:              82.2%           Indels:          0
DB:                       1               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x AW964871 (1-577)

Qy         25 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 44
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCAAATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGA 60

Qy         45 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTG 120

Qy         65 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATG 180

Qy         85 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACA 240

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTT 300

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAAT 360

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGG 420

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACACATGAACTCTAAT 480

Qy        185 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAla 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCT 540

Qy        205 GluGluGlyHisIleIleProArgSerLeuAspLeu 216
              ||||||||||||||||||||  ||||||||||||||
Db        541 AGAGAAGGACACATTATTCCCCCGAAGTTTGACCTT 576

RESULT 27
BU928818
LOCUS       BU928818                909 bp    mRNA    linear   EST 18-OCT-2002
DEFINITION  AGENCOURT_10422426 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:6654806
            5', mRNA sequence.
ACCESSION   BU928818
VERSION     BU928818.1  GI:24117548
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2907  row: a  column: 14
            High quality sequence stop: 543.
FEATURES             Location/Qualifiers
     source          1. .909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6654806"
                     /tissue_type="glioblastoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_57"
                     /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.55 kb (range 0.9-4.0 kb).  12/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
ORIGIN

Alignment Scores:
Pred. No.:                1.62e-106       Length:          909
Score:                    972.00          Matches:         185
Percent Similarity:       99.5%           Conservative:    0
Best Local Similarity:    99.5%           Mismatches:      1
Query Match:              81.5%           Indels:          0
DB:                       5               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BU928818 (1-909)

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCT 62

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAAGT 122

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCys 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 ACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGC 182

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTA 242

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 AAAATGGTTTTGAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGA 302

Qy        142 AspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleVal 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        303 GATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTG 362

Qy        162 AspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATG 422

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 AACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGA 482

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAANAA 542

Qy        222 LeuAspCysGlyArgPhe 227
              ||||||||||||||||||
Db        543 CTGGACTGTGGTAGATTT 560

RESULT 28
BG740339
LOCUS       BG740339                 932 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602635289F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4780318 5',
            mRNA sequence.
ACCESSION   BG740339
VERSION     BG740339.1  GI:14050992
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 932)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10639  row: a  column: 23
            High quality sequence stop: 759.
FEATURES             Location/Qualifiers
     source          1. .932
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4780318"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn3"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:                1.93e-106       Length:          932
Score:                    971.50          Matches:         203
Percent Similarity:       91.3%           Conservative:    7
Best Local Similarity:    88.3%           Mismatches:      11
Query Match:              81.5%           Indels:          9
DB:                       2               Gaps:            4

US-10-757-745-2_COPY_121_347 (1-227) x BG740339 (1-932)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 251
```

```
Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 311

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 371

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 431

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 491

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 551

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 611

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 671

Qy        161 ValAspValTrpGluPhe-LeuGlyLysProLysHisCysGln-TyrThrTrpAspThrG 180
              |||||||||||||||||| |||||||||||||||||||||| |||||||||||||||||
Db        672 GTGGATGTCTGGGAGTTCCTTGGGCAAACCTAAACATTGCCAGTTATACATGGGATACAC 731

Qy        180 lnMetAsnSerAsn-LeuGlyIleThrAlaAlaCys---LysLeuArgPhe---AspArg 197
              ||||||||||||| ||||||||||||||          ||||||          |||||
Db        732 AAATGAACTCTAATCTTGGGAATAACTGCTGACTGGTAAAACTTTCGATTTTGGATCGAA 791

Qy        198 IlePhePheArg------AlaAlaAlaGluGluGlyHisIle---IleProArgSerLeu 214
              |||||||||:::         |||::::::::|||||         ||||||   |||
Db        792 ATATTTTTCAAGAAGCAGACAGCCAGCAAAGACGGACACCATTAATTCCCCCGAAAGTTG 851

Qy        215 -AspLeuLeuGlyLeuGluLys 221
               |||||||||   :::::::::
Db        852 GGACCTTCTTGTGATTCAACAA 873
```

RESULT 29
CB125492

```
LOCUS       CB125492                571 bp    mRNA    linear   EST 29-JAN-2003
DEFINITION  K-EST0174330 C1SNU17 Homo sapiens cDNA clone C1SNU17-16-A01 5',
            mRNA sequence.
ACCESSION   CB125492
VERSION     CB125492.1  GI:28086564
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 571)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 16  row: A  column: 01
            High quality sequence stop: 571.
```

```
FEATURES             Location/Qualifiers
     source          1. .571
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="C1SNU17-16-A01"
                     /sex="F"
                     /tissue_type="Uterine"
                     /cell_type="Epithelial"
                     /cell_line="SNU-17"
                     /lab_host="Top10F'"
                     /clone_lib="C1SNU17"
                     /note="Organ: Cervix; Vector: pCNS-D2; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
ORIGIN
```

Alignment Scores:

Pred. No.:	9.22e-105	Length:	571
Score:	955.00	Matches:	181
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	80.1%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x CB125492 (1-571)

```
Qy         47 AsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAA 62

Qy         67 SerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGA 122

Qy         87 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCAT 182

Qy        107 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 TTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAG 242

Qy        127 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 AAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGG 302

Qy        147 AspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPhe 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTT 362

Qy        167 LeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGly 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGA 422

Qy        187 IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 ATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAG 482

Qy        207 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 226
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 GGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGA 542

Qy        227 Phe 227
                |||
Db        543 TTT 545

RESULT 30
BX337905
LOCUS       BX337905                1067 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX337905 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI052YN13 5-PRIME, mRNA sequence.
ACCESSION   BX337905
VERSION     BX337905.2  GI:46273926
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1067)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30339657.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI052CG07QP1&c=3474.r.
FEATURES              Location/Qualifiers
     source           1. .1067
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DI052YN13"
                      /tissue_type="PLACENTA COT 25-NORMALIZED"
                      /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                      /note="1st strand cDNA was primed with a NotI-oligo(dT)
                      primer. Five prime end enriched, double-strand cDNA was
                      digested with Not I and cloned into the Not I and EcoR V
                      sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:              2.35e-104       Length:         1067
Score:                  955.00          Matches:        194
Percent Similarity:     94.2%           Conservative:   2
Best Local Similarity:  93.3%           Mismatches:     10
Query Match:            80.1%           Indels:         4
DB:                     5               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x BX337905 (1-1067)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 444

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 504

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 564

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 624

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 684

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 744

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        745 TTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 804

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCAAC-AACATT 863

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        864 GTGGATGTCTGGGAGTTTTTKGGCAAACCTAAACATTGCCAGTATACATGGGRTACACAA 923

Qy        181 MetAsnSer-AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePh 200
              ||||||||| |||||||||   ||||||   |||   |||||||||||||      |||||
Db        924 ATGAACTCTHAATCTTGGASTAACTGCTGC-TGTTAACTTCGTTTTGAT---CRATTTTT 979

Qy        200 eArgAlaAlaAlaGluGluGly 207
              |   :::   ||||||   |||
Db        980 TCHGAGCTGCGCAGARGGGGGC 1001

RESULT 31
BG541031
LOCUS       BG541031                546 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602570190F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4694374 5',
            mRNA sequence.
ACCESSION   BG541031
VERSION     BG541031.1  GI:13533264
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1518  row: d  column: 23
            High quality sequence stop: 545.
FEATURES              Location/Qualifiers
     source           1. .546
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4694374"
                      /lab_host="DH10B (T1 phage-resistant)"
                      /clone_lib="NIH_MGC_77"
                      /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                      SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                      3' adaptors were used in cloning as follows: 5' adaptor
                      sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                      5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
```

C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	3.48e-104	Length:	546
Score:	950.00	Matches:	180
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	99.4%	Mismatches:	0
Query Match:	79.7%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x BG541031 (1-546)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 62

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 122

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 182

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 242

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 302

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 362

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db        363 GAGTCAGCTACAGTTATATTTGCAGGAGATACAGATCTAAGGGATCGAGAGGTTACCAGA 422

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 482

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 542

Qy        193 Leu 193
              |||
Db        543 CTT 545
```

RESULT 32
BP258309
LOCUS BP258309 570 bp mRNA linear EST 16-SEP-2004
DEFINITION BP258309 Sugano cDNA library, heart Homo sapiens cDNA clone HRT03149, mRNA sequence.
ACCESSION BP258309
VERSION BP258309.1 GI:52173539
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 570)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRT03149"
/tissue_type="heart"
/clone_lib="Sugano cDNA library, heart"

ORIGIN

Alignment Scores:

Pred. No.:	3.71e-104	Length:	570
Score:	950.00	Matches:	184
Percent Similarity:	98.4%	Conservative:	1
Best Local Similarity:	97.9%	Mismatches:	3
Query Match:	79.7%	Indels:	1
DB:	3	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x BP258309 (1-570)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 65

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 125

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 185

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 245

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 305

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 365

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        366 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGGTACAGTTATATTTGCA 425

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 485

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||   ||||||||||||||||||||||||||||:::|||||||||
Db        486 GTGGATGTCTGGGAGTTTTT-GGGAAACCTAAACATTGCCAGTATACATGGGATTCACAA 544

Qy        181 MetAsnSerAsnLeuGlyIleThr 188
              ||||||||||| |||||||||
Db        545 ATGAACTCTAATTTTGGAATAACT 568
```

RESULT 33
BM476771
LOCUS BM476771 1110 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6481623 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5555311

```
            5', mRNA sequence.
ACCESSION   BM476771
VERSION     BM476771.1  GI:18525813
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1110)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12275  row: e  column: 08
            High quality sequence stop: 582.
FEATURES             Location/Qualifiers
     source          1. .1110
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5555311"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

Alignment Scores:
Pred. No.:              1.33e-103       Length:         1110
Score:                  949.00          Matches:        180
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            79.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BM476771 (1-1110)

Qy         48 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 67
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 TATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCA 64

Qy         68 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 AGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAAC 124

Qy         88 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 CTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTG 184

Qy        108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 GAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAA 244

Qy        128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGAT 304

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 CGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTG 364

Qy        168 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 GGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATA 424

Qy        188 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 ACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGA 484

Qy        208 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 CACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 544

RESULT 34
BX337141
LOCUS       BX337141                1018 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX337141 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI036YJ02 5-PRIME, mRNA sequence.
ACCESSION   BX337141
VERSION     BX337141.2  GI:46271144
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1018)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30337585.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI036DE01QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .1018
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI036YJ02"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:              1.54e-103       Length:         1018
Score:                  948.00          Matches:        182
Percent Similarity:     92.5%           Conservative:   2
Best Local Similarity:  91.5%           Mismatches:     15
Query Match:            79.5%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BX337141 (1-1018)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 410

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 470
```

```
Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 530

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 590

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 650

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        651 TGCCTTATGACATCCCATTTGGAGAGCAMCAGAGGGCATGCTGCGGAAMGAATGAATCAG 710

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        711 TTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 770

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTAMAMAAAAAMATT 830

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||  |||  ||||||  ||||||||||  |||
Db        831 GTGGATGTATGGGAAGTTTTTGGGMAAMCTAAACATTGMMAGTATACATGGGATAMACAA 890

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe 199
              ||||||  ||||||||||||  |||  |||  :::||||||||||||||||||
Db        891 ATGAACTMTAATCTTGGRATAAMTGCTGMTTGTKAAATTCGTTTTGATMGAATATTT 947

RESULT 35
BE162174/c
LOCUS       BE162174                559 bp    mRNA    linear   EST 21-JUN-2000
DEFINITION  IL2-HT0449-250300-046-A04 HT0449 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE162174
VERSION     BE162174.1  GI:8624895
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 559)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL2-HT0449-250
            300-046-A04&t3=2000-03-25&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 558.
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT0449"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:              1.1e-103        Length:         559
Score:                  946.00          Matches:        182
Percent Similarity:     98.4%           Conservative:   1
Best Local Similarity:  97.8%           Mismatches:     3
Query Match:            79.4%           Indels:         0
DB:                     2               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BE162174 (1-559)

Qy         20 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSer 39
              ||||||||||||||||  ||||||||||||||||||:::|||||||||  |||||||||
Db        559 TTAGCTTTGTACAGCCAAGATGTGATATTTCTACAGAAAGTTATTCCNNCATATTATAGC 500

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 59
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 TACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTC 440

Qy         60 ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 ACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTT 380

Qy         80 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 CCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAG 320

Qy        100 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAAT 260

Qy        120 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTT 200

Qy        140 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsn 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAAC 140

Qy        160 IleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThr 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 ATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACA 80

Qy        180 GlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe 199
              ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db         79 CAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTAATGATCGAATATTT 20

Qy        200 PheArgAlaAlaAlaGlu 205
              |||||||||||||||||
Db         19 TTCAGAGCAGCAGCAGAA 2

RESULT 36
CR765451
LOCUS       CR765451                846 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp469E2434_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469E2434 5', mRNA sequence.
ACCESSION   CR765451
VERSION     CR765451.1  GI:52605526
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Hominidae; Pongo.
REFERENCE  1  (bases 1 to 846)
  AUTHORS  Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,
           Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
           Wiemann,S.
  TITLE    Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,
           Deutschenbaur,S., et al.)
  JOURNAL  Unpublished (2004)
COMMENT    Contact: MIPS
           MIPS
           Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
           This is the 5' sequence of the clone insert. Clone from S. Wiemann,
           Molecular Genome Analysis, German Cancer Research Center (DKFZ);
           Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix
           (Martinsried/Germany) within the cDNA sequencing consortium of the
           German Genome Project. This clone (DKFZp469E2434) is available at
           the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
           Berlin, Germany. Please contact RZPD for ordering:
           http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469E2434
           Further information about the clone and the sequencing project is
           available at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp469E2434"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="469 (synonym: pkid1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:                8.24e-103      Length:          846
Score:                    941.00         Matches:         186
Percent Similarity:       97.4%          Conservative:    1
Best Local Similarity:    96.9%          Mismatches:      3
Query Match:              78.9%          Indels:          2
DB:                       7              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x CR765451 (1-846)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 323

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 383

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CTAAAGAAGAGTTCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGGTATTTCACA 443

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 GCTATCATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 503

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        504 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAGTGTGTCAGGAAATGAGCTT 563

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 623

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 683

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATC 743

Qy        161 ValAspValTrp-GluPheLeuGlyLysProLysHisCys-GlnTyrThrTrpAspThrG 180
              |||||||||||| |||||||||||||||||||||||||    |||  ||||||||||||
Db        744 GTGGATGTCTGGGGAGTTTTTGGGCAAACCTAAACATTGGCCAGTNTACATGGGATACAC 803

Qy        180 lnMetAsnSerAsnLeuGlyIleThrAlaAla 190
              |||||||||||||||||||||||||||||||
Db        804 AAATGAACTCTAATCTTGGAATAACTGCTGCT 835

RESULT 37
CN298921
LOCUS       CN298921                 534 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000531590248 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN298921
VERSION     CN298921.1  GI:47315335
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 534)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
  PUBMED    15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 534   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, embryoid bodies
                     derived from H1, H7 and H9 cells"
                     /clone_lib="GRN_EB"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from embryoid body outgrowths derived from hES cell lines
                     H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                     conditions."
ORIGIN

Alignment Scores:
Pred. No.:                9.54e-103      Length:          534
Score:                    938.00         Matches:         178
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              78.7%          Indels:          0
DB:                       7              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x CN298921 (1-534)

Qy         12 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 31
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAG 60

Qy         32 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 51
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 120

Qy         52 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 71
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTA 180

Qy         72 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 91
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTG 240

Qy         92 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 111
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 300

Qy        112 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 131
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCT 360

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 151
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACC 420

Qy        152 ArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLys 171
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAA 480

Qy        172 HisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 189
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCT 534
```

RESULT 38
BG740396

```
LOCUS       BG740396                870 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602634171F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4779449 5',
            mRNA sequence.
ACCESSION   BG740396
VERSION     BG740396.1  GI:14051049
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 870)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10636  row: m  column: 18
            High quality sequence stop: 826.
FEATURES             Location/Qualifiers
     source          1. .870
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4779449"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn3"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:              2.27e-100       Length:         870
Score:                  921.00          Matches:        191
Percent Similarity:     95.5%           Conservative:   2
Best Local Similarity:  94.6%           Mismatches:     4
Query Match:            77.3%           Indels:         6
DB:                     2               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BG740396 (1-870)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 331

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 391

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 451

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 511

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 571

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 631

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle-PheAl 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        632 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATAATTCGC 691

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeu--ProAsnAsn 159
              |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db        692 AGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCCAAACAAC 751

Qy        160 IleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThr 179
              ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        752 ATTGTGGATGTCTGGGA-TTTTTGGGCAAACCTAAACATTGCCAGTATACCTGGGATACC 810

Qy        180 GlnMetAsnSerAsn-LeuGlyIleThrAlaAla-CysLysLeuArgPheAspArg 197
              ::::::|||||||| |||||||||||||||| ||| |||||||||
Db        811 AAATTGAACTCTAATCCTTGGAATAACTGCTGCTTTGTTAACCTTCGTTTGATCGG 866
```

RESULT 39
DN122554/c

```
LOCUS       DN122554                811 bp    mRNA    linear   EST 15-FEB-2005
DEFINITION  1122781 MARC 4PIG Sus scrofa cDNA 3', mRNA sequence.
ACCESSION   DN122554
VERSION     DN122554.1  GI:59816833
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 811)
  AUTHORS   Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
            Nonneman,D.J., Wray,J.E. and Keele,J.W.
  TITLE     Porcine EST collection using a normalized library constructed from
            embryos representing early developmental stages
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
```

```
            Plate: HHY8024  row: L  column: 17
            Seq primer: TAGAAGGCACAGTCGAGG.
FEATURES              Location/Qualifiers
     source           1. .811
                      /organism="Sus scrofa"
                      /mol_type="mRNA"
                      /db_xref="taxon:9823"
                      /tissue_type="pooled"
                      /lab_host="DH10B"
                      /clone_lib="MARC 4PIG"
                      /note="Vector: pcDNA3.1; Site_1: EcoRI; Site_2: NotI;
                      Library made with combined RNA from day-10, day-13,
                      day-15, day-25, and day-30 whole embryos."
ORIGIN

Alignment Scores:
Pred. No.:                  1.34e-98        Length:          811
Score:                      906.00          Matches:         170
Percent Similarity:         96.8%           Conservative:    9
Best Local Similarity:      91.9%           Mismatches:      6
Query Match:                76.0%           Indels:          0
DB:                         8               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x DN122554 (1-811)

Qy          43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
               |||:::::|||:::|||||||||||||||||||:::|||||||||||||||||||||||||
Db         809 AAGAAAGCAAGTAGTTATGAGATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATA 750

Qy          63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
               |||||||||||||||||||||||    |||||||||||||||||||||||||||:::|||
Db         749 ATGTTGAAGAAATCGAGAGTGAAATTTAAAAGCCAAGAGATTATTCCTTTTCCAAATACA 690

Qy          83 LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu 102
               |||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db         689 AAAATGATGAGAAACCTTTTGTGTGTGCATGTGAGTGTGTCAGGAAATGAACTTTGCCTT 630

Qy         103 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 122
               |||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db         629 ATGACTTCTCACTTGGAGAGCACCAGAGGGCATGCCGAGGAACGAATGAATCAGCTGAAA 570

Qy         123 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 142
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         569 ATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGAT 510

Qy         143 ThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
               |||||||||||||||:::|||||||||:::|||||||||||||||||||||||:::|||
Db         509 ACAAATTTAAGGGATCAAGAAGTTACCAAATGTGGTGGCTTACCCAACAACATTTTGGAT 450

Qy         163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 182
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         449 GTCTGGGAGTTTTTGGGCAAACCTAAGCATTGTCAGTATACATGGGATACCCAAATGAAC 390

Qy         183 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAla 202
               |||||||||||||||   |||  ||||||||||||||||||||||||||||||||||||
Db         389 TCTAATCTTGGAATAGCTGCTACTTGTAAGCTTCGTTTTGATCGAATATTTTTTAGAGCA 330

Qy         203 AlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 222
               |||||||||      ||||||||||||||||||||||||||||||||||||||||||||
Db         329 GCAGCAGAAGGCAGCCACATCATTCCCCGAAGCTTGGACCTCCTTGGGTTGGAAAAACTG 270

Qy         223 AspCysGlyArgPhe 227
               |||||||||||||||
Db         269 GACTGTGGTAGATTT 255

RESULT 40
BG562430
LOCUS       BG562430                841 bp    mRNA    linear   EST 10-APR-2001
DEFINITION  602581056F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4709220 5',
            mRNA sequence.
ACCESSION   BG562430
VERSION     BG562430.1  GI:13570082
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 841)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1545  row: o  column: 13
            High quality sequence stop: 801.
FEATURES              Location/Qualifiers
     source           1. .841
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4709220"
                      /lab_host="DH10B (T1 phage-resistant)"
                      /clone_lib="NIH_MGC_76"
                      /note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1:
                      SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                      3' adaptors were used in cloning as follows: 5' adaptor
                      sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                      5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                      C, or G and N = A, C,  G, or T).  Average insert size 1.85
                      kb (range 1.0-4.0 kb). 15/15 colonies contained  inserts
                      by PCR. This library was enriched for full-length clones
                      and was constructed by Clontech  Laboratories (Palo Alto,
                      CA). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                  2.83e-96        Length:          841
Score:                      887.00          Matches:         169
Percent Similarity:         99.4%           Conservative:    0
Best Local Similarity:      99.4%           Mismatches:      1
Query Match:                74.4%           Indels:          0
DB:                         2               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BG562430 (1-841)

Qy          58 TyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle 77
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           3 TATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATT 62

Qy          78 ProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly 97
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 CCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGA 122

Qy          98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 AATGAGCTNTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGA 182

Qy         118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         183 ATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTT 242

Qy         138 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro 157
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         243 ATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCC 302

Qy         158 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         303 AACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGG 362
```

```
Qy        178 AspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArg 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTNTGATCGA 422

Qy        198 IlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu 217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 ATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTT 482

Qy        218 GlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||
Db        483 GGATTAGAAAAACTGGACTGTGGTAGATTT 512
```

RESULT 41
CX958090/c

```
LOCUS       CX958090                 860 bp    mRNA    linear   EST 08-AUG-2005
DEFINITION  JGI_CAAO9595.rev NIH_XGC_tropTe5 Xenopus tropicalis cDNA clone
            IMAGE:7706138 3', mRNA sequence.
ACCESSION   CX958090
VERSION     CX958090.2  GI:71943141
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 860)
  AUTHORS   Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C.,
            Brokstein,P. and Lindquist,E.A.
  TITLE     DOE Joint Genome Institute Xenopus tropicalis EST project
  JOURNAL   Unpublished (2004)
COMMENT     On Feb 8, 2005 this sequence version replaced gi:58771617.
            Other_ESTs: JGI_CAAO9595.fwd
            Contact: Lindquist,E.A., Richardson,P.
            DOE Joint Genome Institute
            2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            Tel: 925 296 5600
            Fax: 925 296 5710
            Email: cdna@jgi-psf.org
            Tissue Procurement: Timothy Grammer (Richard M. Harland Laboratory,
            University of California, Berkeley:
            http://tropicalis.berkeley.edu/home)
            cDNA Library Preparation: DOE Joint Genome Institute:
            http://www.jgi.doe.gov
            DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
            Clone Distribution: I.M.A.G.E. Consortium/LLNL:
            http://image.llnl.gov
            Naming Conventions: EST name is generated by the concatenation of
            the JGI Clone Id and the direction of sequencing. The suffix '.rev'
            indicates a reverse sequencing read of the insert. It does not
            necessarily reflect the orientation of the insert.
            Plate: CAAO 0097  row: f  column: 24
            High quality sequence stop: 855
            POLYA=Yes.
FEATURES              Location/Qualifiers
     source           1. .860
                      /organism="Xenopus tropicalis"
                      /mol_type="mRNA"
                      /db_xref="taxon:8364"
                      /clone="IMAGE:7706138"
                      /sex="male"
                      /tissue_type="Testes"
                      /dev_stage="Adult"
                      /lab_host="ElectroMAX DH10B"
                      /clone_lib="NIH_XGC_tropTe5"
                      /note="Vector: pCMVsport6; Site_1: SalI; Site_2: NotI;
                      This library was made from dT primed cDNA and cloned into
                      Invitrogen pCMVsport6 vector. The work was done at DOE
                      Joint Genome Institute. Poly A RNA were primed with 5'
                      GACTAGTTCTAGATCGCGAG CGGCCGCCCTTTTTTTTTTTTTT 3'. CDNA
                      were ligated to SalI adapter (5' TCGACCCACGCGTCCG and
                      5'CGGACGCGTGGG), digested with NotI, size fractionated in
                      1.1% agarose gel electrophoresis and ligated into NotI and
                      SalI digested pCMVsport6 vector."
ORIGIN
```

Alignment Scores:

Pred. No.:	8.92e-96	Length:	860
Score:	883.00	Matches:	163
Percent Similarity:	85.0%	Conservative:	30
Best Local Similarity:	71.8%	Mismatches:	34
Query Match:	74.1%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x CX958090 (1-860)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||||||||||||||   :::|||::::::|||||||||||||||||||||||||||||||
Db        810 ATAGATGGCCTTGATGAATCAAATGTTGCAGAAAGGGCTCGTGGTGTGTGTTCCTATTTG 751

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||:::|||||||||||||||||||||||||||||||   |||
Db        750 GCCCTGTACAGTCCAGATGTAGTATTCTTACAGGAGGTTATTCCACCATACTATGAGTAC 691

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::   :::|||   ||||||||||||:::||:::   |||||||||
Db        690 CTGAAGAAGAGAGCTGTCAGCTACACAATTATTACAGGTAATGAAGATGAATACTTCACT 631

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||:::|||||||||||||||||||||||||||   ||||||||||||:::|||:::|||
Db        630 GCTATGATGCTGAAAAAGAGCAGGGTAAAGTTAATAAGCCAAGAAATTGTACCATACCCA 571

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||   ||||||||||||||||||      :::||||||:::|||||||||   :::
Db        570 TCCACGGTGATGATGAGAAATTTGCTTGTAGCAAATGTGAATATCTCTGGCAACAGTATA 511

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||:::   |||:::   ||||||:::   |||
Db        510 TGCTTAATGACTTCTCATCTAGAGAGCACAAAAGATCATTCAAAGGAGCGATTAAAGCAG 451

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||   :::|||||||||||||||      ||||||      ||||||||||||||||
Db        450 CTGGACATTGTGCTAAAAAAGATGATGGGTGCTCCCCCTTTAGCAACTGTAATATTTGGA 391

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||:::||||||   :::   ||||||:::||||||||||||||
Db        390 GGGGATACAAACCTGCGAGATCAGGAGGTTGCTAAAATTGGAGGCATGCCAAATAATATA 331

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              :::||||||||||||||||||:::||||||:::||||||:::||||||||||||||:::
Db        330 TTGGATGTTTGGGAGTTCTTGGGTAAGCCTGAGCACTGTCGCTATACCTGGGACACAAAA 271

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              :::|||   ||||||                ||||:::||||||||||||||||||||||
Db        270 GTGAACAAGAACCTGCGTGCTCCCTACACTTGCAGGTTGCGCTTTGACCGCATATTTTTT 211

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||:::   |||         :::||||||:::||||||   |||   |||   |||
Db        210 AGAGCTTCTCAGGAGGGAAGTCAAGTGATCCCCCAGTCTCTTTATCTTGCTGGCACTGAA 151

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db        150 AAGCTGGACTGTGGCAGATTT 130
```

RESULT 42
AL879624/c

```
LOCUS       AL879624                 891 bp    mRNA    linear   EST 15-SEP-2002
DEFINITION  AL879624 XGC-egg Xenopus tropicalis cDNA clone TEgg019p19 3', mRNA
            sequence.
ACCESSION   AL879624
VERSION     AL879624.1  GI:22899889
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
```

```
          Xenopodinae; Xenopus; Silurana.
REFERENCE 1  (bases 1 to 891)
  AUTHORS Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
  TITLE   Sanger Xenopus tropicalis EST project 2002
  JOURNAL Unpublished (2001)
COMMENT   Contact: Taylor R
          Sanger Institute
          Hinxton, Cambridgeshire, CB10 1SA, UK
          Email: trop@sanger.ac.uk
          Sanger Xenopus tropicalis EST project 2001
          TROPICALIS_SEQUENCE_ID: TEgg019p19.q1kT7
          Sequencing primer: T7
          This sequence is from a Xenopus Gene Collection (XGC) library
          constructed by Aaron M. Zorn.
FEATURES             Location/Qualifiers
     source          1. .891
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TEgg019p19"
                     /dev_stage="egg"
                     /lab_host="Escherichia coli XL1-blue"
                     /clone_lib="XGC-egg"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from egg.
                     EcoRI-NotI cut cDNA was then ligated into pCS107 with
                     EcoRI at the 5' end and NotI at the 3' end"
ORIGIN

Alignment Scores:
Pred. No.:                9.41e-96      Length:          891
Score:                    883.00        Matches:         162
Percent Similarity:       85.5%         Conservative:    32
Best Local Similarity:    71.4%         Mismatches:      33
Query Match:              74.1%         Indels:          0
DB:                       1             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x AL879624 (1-891)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||||||||||||||   :::|||:::::::|||||||||||||||||||||||||||||
Db        796 ATAGATGGCCTTGATGAATCAAATGTCGCAGAAAGGGCTCGTGGTGTGTGTTCCTATTTG 737

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||:::|||||||||||||||||||||||||||||||   |||
Db        736 GCCCTGTACAGTCCAGATGTAGTATTCTTACAGGAGGTTATTCCACCATACTATGAGTAC 677

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::   :::|||   |||||||||||||:::|||:::   |||||||||
Db        676 CTGAAGAAGAGAGCTGTCAGCTACACAATTATTACAGGTAATGAGGATGAATACTTCACT 617

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||:::|||||||||||||||||||||||||||   ||||||||||||:::|||:::|||
Db        616 GCTATGATGCTGAAAAAGAGCAGGGTAAAGTTAATAAGCCAAGAAATTGTACCATACCCA 557

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||  |||||||||||||||||||      :::||||||:::|||||||||  :::
Db        556 TCCACGGTGATGATGAGAAATTTGCTTGTAGCAAATGTGAATATCTCTGGCAACAGTATA 497

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||:::   |||:::   ||||||:::   |||
Db        496 TGCTTAATGACTTCTCATCTAGAGAGCACAAAAGATCATTCAAAGGAGCGATTAAAGCAG 437

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||   :::|||||||||||||||  :::||||||     ||||||||||||||||
Db        436 CTGGACATTGTGCTAAAAAAGATGATGGATGCTCCCCCTTTAGCAACTGTAATATTTGGA 377

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||:::|||||||   :::   |||||||:::|||||||||||||
Db        376 GGGGATACAAACCTGCGAGATCAGGAGGTTGCTAAAATTGGAGGCATGCCAAATAATATA 317

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              :::||||||||||:::||||||||||||||||||:::|||||||:::|||||||||||||:::

Db        316 TTGGATGTTTGGGACTTCTTGGGTAAGCCTGAGCACTGTCGCTATACCTGGGACACAAAA 257

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              :::|||   ||||||                  |||:::|||||||||||||||||||||||
Db        256 GTGAACAAGAACCTGCGTGCTCCCTACATTTGCAGGTTGCGCTTTGACCGCATATTTTTT 197

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuGluGlyLeuGlu 220
              |||||||:::   |||         :::|||||||:::||||||   |||   |||   |||
Db        196 AGAGCTTCGCAGGAGGGAAGTCAAGTGATCCCCCAGTCTCTTTATCTTGCTGGCACTGAA 137

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||
Db        136 AAGCTGGACTGTGGCAGATTT 116

RESULT 43
BX776016/c
LOCUS       BX776016                 894 bp    mRNA    linear   EST 10-DEC-2003
DEFINITION  BX776016 XGC-egg Xenopus tropicalis cDNA clone TEgg059p03 3', mRNA
            sequence.
ACCESSION   BX776016
VERSION     BX776016.1  GI:39683222
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (11_2003)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Croning MDR
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TEgg059p03.q1kT7
            Sequencing primer: T7
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            cDNA was oligo dT primed from 5ug of poly A+ RNA from egg.
            EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
            5' end and NotI at the 3' end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1-blue.
FEATURES             Location/Qualifiers
     source          1. .894
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TEgg059p03"
                     /dev_stage="egg"
                     /lab_host="Escherichia coli XL1-blue"
                     /clone_lib="XGC-egg"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from egg.
                     EcoRI-NotI cut cDNA was then ligated into pCS107 with
                     EcoRI at the 5' end and NotI at the 3' end"
ORIGIN

Alignment Scores:
Pred. No.:                9.45e-96      Length:          894
Score:                    883.00        Matches:         162
Percent Similarity:       85.5%         Conservative:    32
Best Local Similarity:    71.4%         Mismatches:      33
Query Match:              74.1%         Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BX776016 (1-894)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||||||||||||||   :::|||::::::|||||||||||||||||||||||||||||
Db        814 ATAGATGGCCTTGATGAATCAAATGTCGCAGAAAGGGCTCGTGGTGTGTGTTCCTATTTG 755
```

```
Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
Db        754 GCCCTGTACAGTCCAGATGTAGTATTCTTACAGGAGGTTATTCCACCATACTATGAGTAC 695

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
Db        694 CTGAAGAAGAGAGCTGTCAGCTACACAATTATTACAGGTAATGAGGATGAATACTTCACT 635

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
Db        634 GCTATGATGCTGAAAAAGAGCAGGGTAAAGTTAATAAGCCAAGAAATTGTACCATACCCA 575

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
Db        574 TCCACGGTGATGATGAGAAATTTGCTTGTAGCAAATGTGAATATCTCTGGCAACAGTATA 515

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
Db        514 TGCTTAATGACTTCTCATCTAGAGAGCACAAAAGATCATTCAAAGGAGCGATTAAAGCAG 455

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
Db        454 CTGGACATTGTGCTAAAAAAGATGATGGATGCTCCCCCTTTAGCAACTGTAATATTTGGA 395

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
Db        394 GGGGATACAAACCTGCGAGATCAGGAGGTTGCTAAAATTGGAGGCATGCCAAATAATATA 335

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
Db        334 TTGGATGTTTGGGACTTCTTGGGTAAGCCTGAGCACTGTCGCTATACCTGGGACACAAAA 275

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
Db        274 GTGAACAAGAACCTGCGTGCTCCCTACATTTGCAGGTTGCGCTTTGACCGCATATTTTTT 215

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
Db        214 AGAGCTTCGCAGGAGGGAAGTCAAGTGATCCCCCAGTCTCTTTATCTTGCTGGCACTGAA 155

Qy        221 LysLeuAspCysGlyArgPhe 227
Db        154 AAGCTGGACTGTGGCAGATTT 134
```

```
RESULT 44
BX780455/c
LOCUS       BX780455                909 bp    mRNA    linear   EST 10-DEC-2003
DEFINITION  BX780455 XGC-egg Xenopus tropicalis cDNA clone TEgg074a05 3', mRNA
            sequence.
ACCESSION   BX780455
VERSION     BX780455.1  GI:39687661
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (11_2003)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Croning MDR
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TEgg074a05.q1kT7
            Sequencing primer: T7
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            cDNA was oligo dT primed from 5ug of poly A+ RNA from egg.
            EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
            5' end and NotI at the 3' end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1-blue.
FEATURES             Location/Qualifiers
     source          1. .909
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TEgg074a05"
                     /dev_stage="egg"
                     /lab_host="Escherichia coli XL1-blue"
                     /clone_lib="XGC-egg"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from egg.
                     EcoRI-NotI cut cDNA was then ligated into pCS107 with
                     EcoRI at the 5' end and NotI at the 3' end"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                9.69e-96      Length:          909
Score:                    883.00        Matches:         162
Percent Similarity:       85.5%         Conservative:    32
Best Local Similarity:    71.4%         Mismatches:      33
Query Match:              74.1%         Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BX780455 (1-909)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
Db        814 ATAGATGGCCTTGATGAATCAAATGTCGCAGAAAGGGCTCGTGGTGTGTGTTCCTATTTG 755

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
Db        754 GCCCTGTACAGTCCAGATGTAGTATTCTTACAGGAGGTTATTCCACCATACTATGAGTAC 695

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
Db        694 CTGAAGAAGAGAGCTGTCAGCTACACAATTATTACAGGTAATGAGGATGAATACTTCACT 635

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
Db        634 GCTATGATGCTGAAAAAGAGCAGGGTAAAGTTAATAAGCCAAGAAATTGTACCATACCCA 575

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
Db        574 TCCACGGTGATGATGAGAAATTTGCTTGTAGCAAATGTGAATATCTCTGGCAACAGTATA 515

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
Db        514 TGCTTAATGACTTCTCATCTAGAGAGCACAAAAGATCATTCAAAGGAGCGATTAAAGCAG 455

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
Db        454 CTGGACATTGTGCTAAAAAAGATGATGGATGCTCCCCCTTTAGCAACTGTAATATTTGGA 395

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
Db        394 GGGGATACAAACCTGCGAGATCAGGAGGTTGCTAAAATTGGAGGCATGCCAAATAATATA 335

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
Db        334 TTGGATGTTTGGGACTTCTTGGGTAAGCCTGAGCACTGTCGCTATACCTGGGACACAAAA 275

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
Db        274 GTGAACAAGAACCTGCGTGCTCCCTACATTTGCAGGTTGCGCTTTGACCGCATATTTTTT 215

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
Db        214 AGAGCTTCGCAGGAGGGAAGTCAAGTGATCCCCCAGTCTCTTTATCTTGCTGGCACTGAA 155

Qy        221 LysLeuAspCysGlyArgPhe 227
```

```
                    ||||||||||||||||||||
Db        154 AAGCTGGACTGTGGCAGATTT 134

RESULT 45
AK009089
LOCUS       AK009089               1289 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male tongue cDNA, RIKEN full-length enriched
            library, clone:2310002D07 product:Traf and Tnf receptor associated
            protein, full insert sequence.
ACCESSION   AK009089
VERSION     AK009089.1  GI:12843662
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1289)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.jp/) for further
            details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved
            with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
            Host: SOLR.
FEATURES             Location/Qualifiers
     source          1. .1289
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2310002D07"
                     /db_xref="taxon:10090"
                     /clone="2310002D07"
                     /sex="male"
                     /tissue_type="tongue"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    55. .1160
                     /note="Traf and Tnf receptor associated protein
                     (MGD|MGI:1860486, GB|NM_019551, evidence: BLASTN, 99%,
                     match=1184)
                     putative"
     polyA_signal    1260. .1265
                     /note="putative"
     polyA_site      1289
                     /note="putative"
ORIGIN

Alignment Scores:
Pred. No.:              8.11e-94        Length:         1289
Score:                  869.00          Matches:        177
Percent Similarity:     88.1%           Conservative:   23
Best Local Similarity:  78.0%           Mismatches:     25
Query Match:            72.9%           Indels:         5
DB:                     4               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x AK009089 (1-1289)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||  ||||||  ||||||||||||||||||||   |||
Db        441 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTCCTGC-CTA 499

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||:::|||||||||||||||   |||||||||   :::|||
Db        500 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTAT-CCCCCATACTGTGCCTAC 558

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::::::::|||   ||||||||||||::||||||||||||||||||
Db        559 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 618

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||:::|||||||||   |||||||||   ||||||||||||||||||||||||||||
Db        619 GCTATACTATTGAAGAAAGA-AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 677

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::||||||||||||||||||||||||||:::|||:::::   |||||||||
Db        678 AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 737

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||   |||:::|||||||||:::   |||
Db        738 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 797

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||   ||||||   ||||||||||||||||||:::|||   ||||||||||||||||
```

```
Db        798 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 857

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db        858 GGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTT 917

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180

Db        918 TTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAA 977

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200

Db        978 GCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTC 1037

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220

Db       1038 AGA------GCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 1091

Qy        221 LysLeuAspCysGlyArgPhe 227

Db       1092 AAACTGGACTGTGGTAGATTT 1112
```

RESULT 46
DR867245/c

```
LOCUS       DR867245                 835 bp    mRNA    linear   EST 29-JUL-2005
DEFINITION  JGI_CABG9422.rev NIH_XGC_tropSto1 Xenopus tropicalis cDNA clone
            IMAGE:7840638 3', mRNA sequence.
ACCESSION   DR867245
VERSION     DR867245.1  GI:71523835
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 835)
  AUTHORS   Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C.,
            Brokstein,P. and Lindquist,E.A.
  TITLE     DOE Joint Genome Institute Xenopus tropicalis EST project
  JOURNAL   Unpublished (2004)
COMMENT     Other_ESTs: JGI_CABG9422.fwd
            Contact: Lindquist,E.A., Richardson,P.
            DOE Joint Genome Institute
            2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            Tel: 925 296 5600
            Fax: 925 296 5710
            Email: cdna@jgi-psf.org
            Tissue Procurement: Robert M. Grainger
            cDNA Library Preparation: Bruce Blumberg Laboratory, University of
            California, Irvine
            DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
            Clone Distribution: I.M.A.G.E. Consortium/LLNL:
            http://image.llnl.gov
            Naming Conventions: EST name is generated by the concatenation of
            the JGI Clone Id and the direction of sequencing. The suffix '.rev'
            indicates a reverse sequencing read of the insert. It does not
            necessarily reflect the orientation of the insert.
            Poly-A: Based upon the presence of a run of 14 or more T residues
            at the beginning of the sequence, this clone was polyadenylated.
            The resulting Poly-T sequence has been removed.
            Plate: CABG 0097  row: k  column: 4
            High quality sequence stop: 826
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .835
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /strain="N6 (Nigerian 6th generation inbred)"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:7840638"
                     /tissue_type="Stomach"
                     /dev_stage="Adult"
                     /lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
                     /clone_lib="NIH_XGC_tropSto1"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: XhoI; The
                     library was prepared from 5 ug of poly A+ RNA by oligo-dT
                     priming
                     (5'-ACTAGTGCGGCCGCCTAGGCCTCGAGTTTTTTTTTTTTTTTTV-3') and
                     Stratascript reverse transcriptase. After ligation of
                     EcoRI adapters (5'-AATTCGGCACGAGG-3') followed by kinasing
                     adapters and by XhoI digestion, the cDNA was size selected
                     by chromatography on Sepharose CL-2B columns and fractions
                     containing cDNAs larger than 1000 bp were ligated into
                     EcoRI/XhoI-digested pCS107. Reference for library
                     construction: Current Genomics 4, 635-644. Library
                     constructed by Michelle Tabb and Bruce Blumberg (Dept of
                     Developmental and Cell Biology, University of California,
                     Irvine)."
ORIGIN
```

Alignment Scores:

Pred. No.:	5.59e-94	Length:	835
Score:	868.00	Matches:	162
Percent Similarity:	81.9%	Conservative:	32
Best Local Similarity:	68.4%	Mismatches:	33
Query Match:	72.8%	Indels:	10
DB:	8	Gaps:	1

US-10-757-745-2_COPY_121_347 (1-227) x DR867245 (1-835)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20

Db        825 ATAGATGGCCTTGATGAATCAAATGTCGCAGAAAGGGCTCGTGGTGTGTGTTCCTATTTG 766

Qy         21 AlaLeu----------------------------TyrSerProAspValIlePheLeu 30

Db        765 GCCCTGTAAGTTCCACTACCATTTGCTACTACTAGGTACAGTCCAGATGTAGTATTCTTA 706

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50

Db        705 CAGGAGGTTATTCCACCATACTATGAGTACCTGAAGAAGAGAGCTGTCAGCTACACAATT 646

Qy         51 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 70

Db        645 ATTACAGGTAATGAGGATGAATACTTCACTGCTATGATGCTGAAAAAGAGCAGGGTAAAG 586

Qy         71 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 90

Db        585 TTAATAAGCCAAGAAATTGTACCATACCCATCCACGGTGATGATGAGAAATTTGCTTGTA 526

Qy         91 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 110

Db        525 GCAAATGTGAATATCTCTGGCAACAGTATATGCTTAATGACTTCTCATCTAGAGAGCACA 466

Qy        111 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 130

Db        465 AAAGATCATTCAAAGGAGCGATTAAAGCAGCTGGACATTGTGCTAAAAAAGATGATGGAT 406

Qy        131 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluVal 150

Db        405 GCTCCCCCTTTAGCAACTGTAATATTTGGAGGGGATACAAACCTGCGAGATCAGGAGGTT 346

Qy        151 ThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPro 170

Db        345 GCTAAAATTGGAGGCATGCCAAATAATATATTGGATGTTTGGGACTTCTTGGGTAAGCCT 286

Qy        171 LysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAla 190

Db        285 GAGCACTGTCGCTATACCTGGGACACAAAAGTGAACAAGAACCTGCGTGCTCCCTACATT 226

Qy        191 CysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210

Db        225 TGCAGGTTGCGCTTTGACCGCATATTTTTTAGAGCTTCGCAGGAGGGAAGTCAAGTGATC 166

Qy        211 ProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
```

```
Db        165 CCCCAGTCTCTTTATCTTGCTGGCACTGAAAAGCTGGACTGTGGCAGATTT 115

RESULT 47
BU936308
LOCUS       BU936308                 815 bp    mRNA    linear   EST 18-OCT-2002
DEFINITION  AGENCOURT_10522379 NIH_MGC_169 Mus musculus cDNA clone
            IMAGE:6704064 5', mRNA sequence.
ACCESSION   BU936308
VERSION     BU936308.1  GI:24125127
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 815)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jonathan Kuo, NIMH
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2999  row: e  column: 24
            High quality sequence stop: 600.
FEATURES             Location/Qualifiers
     source          1. .815
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6704064"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_169"
                     /note="Organ: Testicles; Vector: pDNR-LIB; Site_1: SfiI
                     (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); cDNA made
                     by oligo-dT priming and directionally cloned. 5' and 3'
                     adaptors were used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 0.5 kb
                     size fraction. Library created in thelaboratory of M.
                     Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              3.06e-90        Length:         815
Score:                  837.00          Matches:        156
Percent Similarity:     89.5%           Conservative:   23
Best Local Similarity:  78.0%           Mismatches:     19
Query Match:            70.2%           Indels:         2
DB:                     5               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x BU936308 (1-815)

Qy         28 IlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsn 47
              :::|||||||||||||||||||||||||||||   :::|||||||||||||||:::::::::
Db          4 GTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGT 63

Qy         48 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 67
              |||   ||||||||||||:::|||||||||||||||||||||||:::|||||||||
Db         64 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 123

Qy         68 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 87
              |||||||||   ||||||||||||||||||||||||||:::||||||||||||||||||
Db        124 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAAC 183

Qy         88 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 107
              |||||||||||:::|||::::::   |||||||||   ||||||||||||||||||||
Db        184 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 243

Qy        108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
              ||||||||||   |||:::|||||||:::   ||||||||   ||||||   |||
Db        244 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 303

Qy        128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 147
              ||||||||||||||:::|||   ||||||||||||||||||||||||||||||||||||
Db        304 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 363

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
              :::|||||   :::|||||||||||||||:::|||:::   |||   ||||||||||||
Db        364 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 423

Qy        168 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 187
              ||||||||||||||||||||||||||||||||||:::   |||:::||||||   |||
Db        424 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 483

Qy        188 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 207
                 ||||||   |||   |||||||||||||||||||||||||      ||||||||||||
Db        484 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGG 537

Qy        208 HisIleIleProArgSerLeuAspLeuGluGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||:::|||||||:::|||||||||||||:::|||||||||||||||||||||||||||||
Db        538 CACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTT 597

RESULT 48
BU204728
LOCUS       BU204728                 760 bp    mRNA    linear   EST 25-NOV-2002
DEFINITION  604152621F1 CSEQCHN03 Gallus gallus cDNA clone ChEST995j5 5', mRNA
            sequence.
ACCESSION   BU204728
VERSION     BU204728.1  GI:25369232
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .760
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="White Leghorn, Hisex"
                     /db_xref="taxon:9031"
                     /clone="ChEST995j5"
                     /tissue_type="whole embryo"
                     /dev_stage="20-21"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN03"
                     /note="Organ: whole embryo; Vector: pBluescript II KS(+);
                     Site_1: EcoRI; Site_2: Not1; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
```

compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores:

Pred. No.:	4.82e-90	Length:	760
Score:	835.00	Matches:	163
Percent Similarity:	82.9%	Conservative:	26
Best Local Similarity:	71.5%	Mismatches:	37
Query Match:	70.1%	Indels:	3
DB:	5	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x BU204728 (1-760)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||    ||||||   ||||||||||||||||||||||||||||||
Db         18 ATTGATGGGCTGGACCTGGGGAATCTGCAGGAGCGAGCTAGAGGTGTCTGTTCTTACCTG 77

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||
Db         78 GCATTATACAGCCCAGATGTGGTGTTTTTACAGGAGGTTATCCCACCGTATCTCTGTATT 137

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||::::::|||:::     |||   ||||||   |||:::   :::||||||||||||
Db        138 CTACAGAGGAGAGCAGGCGGTTACACCATTATTCCAGGTAACGTAGATGGCTATTTCACT 197

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||::::::|||||||||   |||||||||:::      ||||||||||||   ||||||
Db        198 GCCATGCTGTTGAAGAAACCAAGAGTGAAAGTACTAAAACAGGAGATAATACGTTTTCCA 257

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::|||   ||||||||||||||||||   ||||||||||||:::||||||||||||||||
Db        258 ACAACTTCCATGATGAGGAACCTTTTGGTTGTGCATGTGAACATATCTGGTAATGAACTT 317

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||   |||:::   |||||||||   |||
Db        318 TGCCTTATGACTTCTCATCTGGAGAGCACTAGAGATCACTCAAAGGAGCGTATGAAGCAA 377

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||::::::||||||   ||||||||||||      :::|||   ||||||||||||
Db        378 CTGCAAATAGTGTTAAACAAAATGCAGGAGGAGTCTCAGTCCACCACTGTTATATTTGGG 437

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||   ||||||   :::   ||||||||||||   ||||||
Db        438 GGGGATACAAACCTCAGAGACAGCGAGGTGGCTAAGCTAGGTGGCTTACCTAAAAACATC 497

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||:::||||||||||||||||||||||:::||||||:::||||||||||||||||||
Db        498 ACGGATATCTGGGAGTTTCTGGGCAAACCCCAGCACTGCCGCTATACGTGGGACACAAGC 557

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||:::||||||   |||   :::   ||||||||||||   ||||||:::||||||
Db        558 TCCAACACCAACCTGCGCATAGAGAGTAAATGCAAGTTGAGGTT-GATCGCCTTTTCTTT 616

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIlePro-ArgSerLeuAspLeuLeuGlyLeuGl 220
              |||   |||||||||   ||||||||||||||| |||::::::||||||:::||||||||
Db        617 CGGCCTGCAGCAGAAGGGGGACATATCATTCCACCGCAATATGGACTTGATCGGATTAGA 676

Qy        220 uLysLeu-AspCysGlyArg 226
              |||||||  |||||||||:::
Db        677 AAAGCTGGGACTGTGGGCAG 696
```

RESULT 49
BU159911
LOCUS BU159911 870 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7933863 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6144208 5', mRNA sequence.
ACCESSION BU159911
VERSION BU159911.1 GI:22673821
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 870)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13468 row: n column: 17
High quality sequence stop: 677.
FEATURES Location/Qualifiers
source 1. .870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6144208"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."
ORIGIN

Alignment Scores:

Pred. No.:	1.18e-89	Length:	870
Score:	832.50	Matches:	170
Percent Similarity:	94.5%	Conservative:	1
Best Local Similarity:	93.9%	Mismatches:	6
Query Match:	69.8%	Indels:	4
DB:	5	Gaps:	1

US-10-757-745-2_COPY_121_347 (1-227) x BU159911 (1-870)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 355

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 415

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 475

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 535

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 595

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        596 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 655

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        656 TTAAAAATGGTTTTAAAGAAANATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 715
```

```
Qy        141 GlyAspThrAsnLeuArgAspArgGluValThr-ArgCys-GlyGlyLeuProAsnAsnI 160
              ||||||||||||||||||||||||||||||||| ||||||  ||||||||||||||||
Db        716 GGAGATACAAATCTAAGGGATCGAGAGGTTACCCAGATGTGNGTGGTTTACCCAACAACA 775

Qy        160 leValAspValTrp-GluPheLeuGlyLysProLysHisCys---GlnTyrThrTrp 177
              |||||||||||||| ||||||||||:::                ||||||||||||
Db        776 TTGTGGATGTCTGGGGAGTTTTTGGGGCAAACCCTANACATTGGCCAGTATACATGG 832

RESULT 50
AL555333
LOCUS       AL555333                948 bp    mRNA    linear   EST 30-MAR-2004
DEFINITION  AL555333 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DK007YK05 5-PRIME, mRNA sequence.
ACCESSION   AL555333
VERSION     AL555333.3  GI:45860070
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:31277141.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DK007AF03QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .948
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK007YK05"
                     /cell_type="HELA CELLS COT 25-NORMALIZED"
                     /cell_line="HELA"
                     /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Alignment Scores:
Pred. No.:              3.11e-89        Length:         948
Score:                  829.50          Matches:        172
Percent Similarity:     92.0%           Conservative:   0
Best Local Similarity:  92.0%           Mismatches:     14
Query Match:            69.6%           Indels:         3
DB:                     1               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x AL555333 (1-948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              |||||||||   |||   |||||||||||||||||||||||||||||||||      |||
Db        800 GGAGATACAWATCTARGGGATCGAGAGGTTACCAGATGTGGTGGKTTACCAACAAM-ATT 858

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||   ||||||||||||||||         ||||||
Db        859 GTKGATKTCTGGGAGTTTTTGGGCAAMCTTAA-CATTGCAGTATAAWB---GGATMAMAA 914

Qy        181 MetAsnSerAsnLeuGlyIle 187
              |||||||||||||||||||
Db        915 ATGAACTCTAATCTYGGAATA 935

RESULT 51
BM553049
LOCUS       BM553049                1192 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6542413 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742798
            5', mRNA sequence.
ACCESSION   BM553049
VERSION     BM553049.1  GI:18791437
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1192)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12761  row: i  column: 07
            High quality sequence stop: 747.
FEATURES             Location/Qualifiers
     source          1. .1192
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5742798"
                     /tissue_type="medulla"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_119"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source normal medulla from
                     anonymous male age 27.  Library  is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed  upon
                     cloning).  Average insert size 1.3 kb, insert size range
```

0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	5.04e-89	Length:	1192
Score:	829.00	Matches:	176
Percent Similarity:	83.9%	Conservative:	7
Best Local Similarity:	80.7%	Mismatches:	25
Query Match:	69.5%	Indels:	10
DB:	3	Gaps:	3

US-10-757-745-2_COPY_121_347 (1-227) x BM553049 (1-1192)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 477

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 537

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 597

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 657

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 717

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 777

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        778 TTAAAAATGGTTTTTAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 837

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThr-ArgCysGlyGlyLeuProAsnAsnIl 160
              |||||||||||||||||||||||||||||||| ||||||||||||||||||      ::
Db        838 GGAGATACAAATCTAAGGGATCGAGAGGTTACCACGATGTGGTGGTTTACCCACAACATT 897

Qy        160 eValAspVal-Trp---GluPheLeuGlyLysProLysHisCys--GlnTyrThrTrpAs 178
              :   |||||| |||     |||   |||   ||||||||||| :::||||||||
Db        898 GTGGGATGTCCTGGGGAGTTTTTGGGGGCAACCCTAAACATTGCCCAGGTATACATGGGG 957

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------IleThrAlaAlaCysLysLeuArgPh 195
                        ::::::                 :::   |||                ||
Db        958 ATACCACAAATGAAACTCCTATTCCTGGGAAAAACTGCCTGCTTGGAAAACCTTCCCTTT 1017

Qy        195 eAspArgIlePhePheArgAlaAlaAla------GluGluGlyHis 208
              ||||      ||||||      |||         :::||||||:::
Db       1018 TGATTCGAAATTTTTTTCCCAAGCAACCCACCCCAAAGAAGGGAAC 1063
```

RESULT 52
CN694288

LOCUS	CN694288 617 bp mRNA linear EST 18-MAY-2004
DEFINITION	E0348E09-5 NIA Mouse E10.5 whole embryo cDNA library (Long) Mus musculus cDNA clone NIA:E0348E09 IMAGE:30863576 5', mRNA sequence.
ACCESSION	CN694288
VERSION	CN694288.1 GI:47463037
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE	1 (bases 1 to 617)
AUTHORS	Sharov,A.A., Piao,Y., Matoba,R., Dudekula,D.B., Qian,Y., VanBuren,V., Falco,G., Martin,P.R., Stagg,C.A., Bassey,U.C., Wang,Y., Carter,M.G., Hamatani,T., Aiba,K., Akutsu,H., Sharova,L., Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S., Nagaraja,R., Boheler,K.R., Taub,D., Hodes,R.J., Longo,D.L., Schlessinger,D., Keller,J., Klotz,E., Kelsoe,G., Umezawa,A., Vescovi,A.L., Rossant,J., Kunath,T., Hogan,B.L., Curci,A., D'Urso,M., Kelso,J., Hide,W. and Ko,M.S.
TITLE	Transcriptome analysis of mouse stem cells and early embryos
JOURNAL	PLoS Biol. 1 (3), 410-419 (2003)
COMMENT	Contact: Dawood B. Dudekula Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@lgsun.grc.nia.nih.gov Plate: E0348 row: E column: 09 Seq primer: M13 Reverse High quality sequence stop: 617 POLYA=No.

FEATURES Location/Qualifiers

source 1. .617
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:E0348E09-5"
/db_xref="taxon:10090"
/clone="NIA:E0348E09 IMAGE:30863576"
/tissue_type="whole embryo including extraembryonic tissues at 10.5-days postcoitum"
/dev_stage="E10.5"
/lab_host="DH10B"
/clone_lib="NIA Mouse E10.5 whole embryo cDNA library (Long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were extracted from a pool of 8 embryos at 10.5-days postcoitum. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from 2ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.4Kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:

Pred. No.:	2.48e-89	Length:	617
Score:	828.00	Matches:	156
Percent Similarity:	88.4%	Conservative:	20
Best Local Similarity:	78.4%	Mismatches:	23
Query Match:	69.5%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x CN694288 (1-617)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||      ||||||  ||||||||||||||||||||||||   |||
Db         20 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 79
```

```
Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||:::|||||||||||||||||||||||||||  :::|||
Db         80 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTAC 139

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||:::::::::|||   ||||||||||||:::||||||||||||||||||||
Db        140 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 199

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||:::|||||||||   |||||||||   ||||||||||||||||||||||||||||
Db        200 GCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 259

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::|||||||||||||||||||||||||||||:::|||:::::::   |||||||||
Db        260 AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 319

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||   |||:::|||||||:::   |||
Db        320 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 379

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||   ||||||   ||||||||||||||||||:::|||   |||||||||||||||
Db        380 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 439

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||:::||||||   :::|||||||||||||||:::|||:::
Db        440 GGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTT 499

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        500 TTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAA 559

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe 199
              |||:::||||||   |||   ||||||   |||   ||||||||||||||||||
Db        560 GCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTT 616

RESULT 53
AV602300
LOCUS       AV602300                578 bp    mRNA    linear   EST 23-FEB-2005
DEFINITION  AV602300 Bos taurus kidney fetus Bos taurus cDNA clone E1KI009D11
            5', mRNA sequence.
ACCESSION   AV602300
VERSION     AV602300.1  GI:9724625
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
            and Sugimoto,Y.
  TITLE     Establishment of a high throughput EST sequencing system using
            poly(A) tail-removed cDNA libraries and determination of 36,000
            bovine ESTs
  JOURNAL   Nucleic Acids Res. 29 (22), E108 (2001)
  PUBMED    11713328
COMMENT     Contact: Yoshikazu Sugimoto
            Animal Genetics Division
            Shirakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
            Fax: 81-248-25-5725
            Email: kazusugi@siag.or.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
FEATURES             Location/Qualifiers
     source          1. .578
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /clone="E1KI009D11"
                     /tissue_type="kidney"
                     /dev_stage="fetus"
                     /lab_host="DH10B"
                     /clone_lib="Bos taurus kidney fetus"
                     /note="Vector: pZL1; Site_1: Sal1; Site_2: Not1; Poly A
                     was deleted from a Not1 site"
ORIGIN

Alignment Scores:
Pred. No.:                3.93e-89        Length:          578
Score:                    826.00          Matches:         156
Percent Similarity:       95.9%           Conservative:    8
Best Local Similarity:    91.2%           Mismatches:      7
Query Match:              69.3%           Indels:          0
DB:                       1               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x AV602300 (1-578)

Qy         57 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 76
              ||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||
Db         12 GGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAGTTTAAAAGCCAAGAAATT 71

Qy         77 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 96
              ||||||||||||:::|||:::|||||||||||||||||||||||||||||:::||||||
Db         72 ATTCCTTTTCCAAATACGCAAATGATGAGAAACCTTTTGTGTGTTCATGTGAGTGTGTCA 131

Qy         97 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        132 GGAAATGAACTTTGCCTTATGACTTCTCATTTGGAAAGCACCAGAGGGCACGCTANGGAA 191

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              ||||||||||||||||||||||||||:::|||||||||||||||||   ||||||||||
Db        192 CGAATGAATCAGTTGAAAATGGTTTTAGAGAAAATGCAAGAGGCTCCAGGATCAGCTACA 251

Qy        137 ValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeu 156
              ||||||||||||||||||||||||||||||||||:::||||||||::|||||||||||||
Db        252 GTTATATTTGCTGGAGATACGAATCTCAGGGATCAAGAAGTTACCAAATGTGGTGGTTTA 311

Qy        157 ProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 176
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        312 CCCAACAACATTTTGGATGTCTGGGAGTTTTTGGGCAAACCTAAGCATTGCCAGTATACA 371

Qy        177 TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAsp 196
              ||||||||||||||||||||||||||||||||||   |||   ||||||||||||||||||
Db        372 TGGGATACACAAATGAACTCTAATCTTGGAATAGCTGCTAATTGTAAACTTCGTTTTGAT 431

Qy        197 ArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeu 216
              ||||||||||||||||||||||||||||   ||||||||||||||:::||||||||||||
Db        432 CGAATATTTTTTAGAGCTGCAGCAGAAGGTGGCCACATTATTCCCCAAAGTTTGGACCTG 491

Qy        217 LeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||   |||||||||
Db        492 CTTGGGCTGGAAAAACTGGACTGNGGTAGATTT 524

RESULT 54
BU169945
LOCUS       BU169945                883 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7913097 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6024760
            5', mRNA sequence.
ACCESSION   BU169945
VERSION     BU169945.1  GI:22683929
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
```

```
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM13235  row: m  column: 17
           High quality sequence stop: 413.
FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6024760"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

Alignment Scores:
Pred. No.:                2.26e-88        Length:         883
Score:                    822.00          Matches:        168
Percent Similarity:       93.3%           Conservative:   0
Best Local Similarity:    93.3%           Mismatches:     5
Query Match:              69.0%           Indels:         7
DB:                       5               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x BU169945 (1-883)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 300

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 360

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 420

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 480

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 540

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 600

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 660

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCys-GlyGlyLeuProAsnAsnIl 160
              |||||||||||||||||||||||||  ||||||  ||||||||||||||||||||||
Db        661 GGAGATACAAATCTAAGGGATCGAGAGGGTACCAGATGGGGGGGGTTTACCCAACAACAT 720

Qy        160 eValAspValTrp--GluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              |   |||   |||     ||||||||||||||||||||||||||||||||||
Db        721 TGGGGATGGCTGGGGAGTTTTTTGGGGAAACCTAAA----------ACATGG 762

RESULT 55
CB443087
LOCUS       CB443087                724 bp    mRNA    linear   EST 25-MAR-2003
DEFINITION  693950 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   CB443087
VERSION     CB443087.1  GI:29231758
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
            Wray,J.E. and Keele,J.W.
  TITLE     A second set of bovine ESTs from pooled-tissue normalized libraries
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: FQY8050  row: J  column: 10
            Seq primer: GTAATACGACTCACTATAGGG.
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 6BOV"
                     /note="Vector: pcDNA3.1; Site_1: EcoRI; Site_2: NotI;
                     Library made with RNA pooled from multiple tissues
                     including liver, lung, hypothalamus, pituitary, and
                     placenta/endometrium."
ORIGIN

Alignment Scores:
Pred. No.:                2.22e-88        Length:         724
Score:                    821.00          Matches:        155
Percent Similarity:       93.2%           Conservative:   10
Best Local Similarity:    87.6%           Mismatches:     12
Query Match:              68.9%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CB443087 (1-724)

Qy         49 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 68
              :::|||      |||   ||||||   ||||||||||||||||||||||||||||||||||
Db        194 CAAATACCTGAAGGTCGTGAAGAAAGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 253

Qy         69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 88
              ||||||   ||||||||||||||||||||||||||:::|||:::|||||||||||||||
Db        254 GTGAAGTTTAAAAGCCAAGAAATTATTCCTTTTCCAAATACGCAAATGATGAGAAACCTT 313

Qy         89 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
              ||||||||||||||:::||||||   ||||||||||||||||||||||||||||||||||
Db        314 TTGTGTGTTCATGTGAGTGTGTCAAGAAATGAACTTTGCCTTATGACTTCTCATTTGGAA 373

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              ||||||||||||||||||   ||||||||||||||   ||||||||||||:::||||||
Db        374 AGCACCAGAGGGCACGCTAAGGAACGAATGAATCAGTTCAAAATGGTTTTAGAGAAAATG 433

Qy        129 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 148
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||:::
Db        434 CAAGAGGCTCCAGGATCAGCTACAGTTATATTTGCTGGAGATACAAATCTCAGGGATCAA 493

Qy        149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
              |||||||||:::||||||||||||||||||||||||:::||||||||||||||||||||
Db        494 GAAGTTACCAAATGTGGTGGTTTACCCAACAACATTTTGGATGTCTGGGAGTTTTTGGGC 553

Qy        169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 188
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 AAACCTAAGCATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAGCT 613

Qy        189 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 208
                 |||   ||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        614 GCTAATTGTAAACTTCGTTTTGATCGAATATTTTTTAGAGCTGCAGCAGAAAGTGGCCAC 673

Qy        209 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGly 225
                 ||||||||||:::|||||||||||||||||||||||:::|||||||||||||||
Db        674 ATTATTCCCCAAAGTTTGGACCTGCTTGGGCTGAAAAAACTGGACTGTGGT 724

RESULT 56
CX166335
LOCUS       CX166335                 810 bp    mRNA    linear   EST 23-DEC-2004
DEFINITION  HESC2_39_B09.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone
            IMAGE:7469923 5', mRNA sequence.
ACCESSION   CX166335
VERSION     CX166335.1  GI:56796415
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 810)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15771  row: d  column: 17
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 810.
FEATURES             Location/Qualifiers
     source          1. .810
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7469923"
                     /sex="male"
                     /tissue_type="human embryonic stem cells differentiated to
                     an early endodermal cell type"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_258"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early endodermal cell type. Cell line
                     id and NIH Registry designation is BG01. Positive for
                     GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP
                     expression. Passage number 40. cDNA primed using oligo-dT
                     primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and
                     cloned into the EcoRV/NotI sites of pExpress-1. This
                     primary library is non-normalized (normalized primary
                     library is NIH_MGC_259). It was constructed by Express
                     Genomics (Frederick, MD). Sequence ends have been trimmed
                     to exclude vector and regions below Phred quality 16.
                     Three-prime sequences are presented as their reverse
                     complement and have been trimmed to exclude polyA. Note:
                     this is a Mammalian Gene Collection library."

ORIGIN

Alignment Scores:
Pred. No.:                2.63e-88        Length:          810
Score:                    821.00          Matches:         163
Percent Similarity:       99.4%           Conservative:    0
Best Local Similarity:    99.4%           Mismatches:      1
Query Match:              68.9%           Indels:          1
DB:                       8               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x CX166335 (1-810)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 377

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 437

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 497

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 557

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 617

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 677

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 737

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db        738 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAC-AACATT 796

Qy        161 ValAspValTrp 164
              ||||||||||||
Db        797 GTGGATGTCTGG 808

RESULT 57
CF378639/c
LOCUS       CF378639                 850 bp    mRNA    linear   EST 27-AUG-2003
DEFINITION  AGENCOURT_15349090 NICHD_XGC_Swb1N Xenopus tropicalis cDNA clone
            IMAGE:7005669 5', mRNA sequence.
ACCESSION   CF378639
VERSION     CF378639.1  GI:34316083
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 850)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Rob Granger, University of Virginia
             cDNA Library Preparation: Open Biosystems
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM14704  row: p  column: 19
           High quality sequence stop: 594.
FEATURES             Location/Qualifiers
     source          1. .850
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:7005669"
                     /tissue_type="whole body"
                     /clone_lib="NICHD_XGC_Swb1N"
                     /note="Vector: pExpress-1; Site_1: EcoRV; Site_2: NotI;
                     Bulk tissue was collected from a whole 10 month old male
                     from the F6 strain. 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer, double-stranded cDNA was cloned into
                     the Not I and EcoRV sites of pExpress-1. Library was
                     size-selected for >1.5 kb fragments for an average insert
                     size of 1.92 kb. Library was normalized to Cot5 with a
                     180-fold reduction of actin. A non-normalized version of
                     this library is also available (NICHD_XGC_Swb1). Library
                     was constructed by Open Biosystems (Huntsville, AL).
                     PLEASE NOTE: This library contains high level of
                     contamination by worm"
ORIGIN

Alignment Scores:
Pred. No.:              4.93e-88        Length:         850
Score:                  819.00          Matches:        152
Percent Similarity:     84.6%           Conservative:   29
Best Local Similarity:  71.0%           Mismatches:     32
Query Match:            68.7%           Indels:         1
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CF378639 (1-850)

Qy          15 GlyValCysSerTyrLeuAlaLeuTyrSer-ProAspValIlePheLeuGlnGluValIl 34
               ||||||||||||||||||||||||||||||| |||||||||:::|||||||||||||||||
Db         782 GGTGTGTGTTCCTATTTGGCCCTGTACAGTCCCAGATGTAGTATTCTTACAGGAGGTTAT 723

Qy          34 eProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHi 54
               |||||||||||||   ||||||||||||||:::   :::|||  ||||||||||||::
Db         722 TCCACCATACTATGAGTACCTGAAGAAGAGAGCTGTCAGCTACACAATTATTACAGGTAA 663

Qy          54 sGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGl 74
               :|||:::   ||||||||||||:::|||||||||||||||||||||||||   |||||
Db         662 TGAGGATGAATACTTCACTGCTATGATGCTGAAAAAGAGCAGGGTAAAGTTAATAAGCCA 603

Qy          74 nGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAs 94
               |||||||:::|||:::|||||||||   ||||||||||||||||||      :::|||||
Db         602 AGAAATTGTACCATACCCATCCACGGTGATGATGAGAAATTTGCTTGTAGCAAATGTGAA 543

Qy          94 nValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAl 114
               |:::|||||||||   :::|||||||||||||||||||||||||||:::   |||::
Db         542 TATCTCTGGCAACAGTATATGCTTAATGACTTCTCATCTAGAGAGCACAAAAGATCATTC 483

Qy         114 aAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSe 134
               :   ||||||:::   ||||||   :::|||||||||||||||   :::|||||||
Db         482 AAAGGAGCGATTAAAGCAGCTGGACATTGTGCTAAAAAAGATGATGGATGCTCCCCCTTT 423

Qy         134 rAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGl 154
               |||||||||||||||   |||||||||||||||||||:::|||||||   :::   ||
Db         422 AGCAACTGTAATATTTGGAGGGGATACAAACCTGCGAGATCAGGAGGTTGCTAAAATTGG 363

Qy         154 yGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGl 174
               ||||:::|||||||||||:::|||||||:::|||||||||||||||:::|||||||::
Db         362 AGGCATGCCAAATAATATATTGGATGTTTGGGACTTCTTGGGTAAGCCTGAGCACTGTCG 303

Qy         174 nTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuAr 194
               :|||||||||||||||::::::|||   ||||||                 |||:::||||||
Db         302 CTATACCTGGGACACAAAAGTGAACAAGAACCTGCGTGCTCCCTACATTTGCAGGTTGCG 243

Qy         194 gPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLe 214
               |||||||||||||||||||||||||:::   |||          :::|||||:::||||||
Db         242 CTTTGACCGCATATTTTTTAGAGCTTCGCAGGAGGGAAGTCAAGTGATCCCCCAGTCTCT 183

Qy         214 uAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
               |   |||   |||   ||||||||||||||||||||||
Db         182 TTATCTTGCTGGCACTGAAAAGCTGGACTGTGGCAGATTT 143

RESULT 58
BJ639656/c
LOCUS       BJ639656                737 bp    mRNA    linear   EST 01-OCT-2003
DEFINITION  BJ639656 NIBB Mochii normalized Xenopus early gastrula library
            Xenopus laevis cDNA clone XL202p03 3', mRNA sequence.
ACCESSION   BJ639656
VERSION     BJ639656.1  GI:37299131
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 737)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
            Kohara,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            The information of this clone is available through the following
            URL.
            http://xenopus.nibb.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .737
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="XL202p03"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 10.5"
                     /clone_lib="NIBB Mochii normalized Xenopus early gastrula
                     library"
ORIGIN

Alignment Scores:
Pred. No.:              6.47e-87        Length:         737
Score:                  809.00          Matches:        149
Percent Similarity:     83.5%           Conservative:   33
Best Local Similarity:  68.3%           Mismatches:     36
Query Match:            67.9%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BJ639656 (1-737)

Qy          10 SerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePhe 29
               :::|||||||||||   |||||||||  ||||||||||||:::|||||||||:::|||
Db         736 GCAGAGAGGGCTCGTGCTGTATGTTCCTGTTTGGCCCTGTACACTCCAGATGTAGTATTT 677

Qy          30 LeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGlu 49
               ||||||||||||||||||||||||||   |||   |||||||||:::   :::|||:::
Db         676 TTGCAGGAAGTTATTCCACCATACTGTGAGTACNTGAAGAAGAGGGCCGTCAGCTACAAA 617

Qy          50 IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 69
               |||||||||||:::|||:::   ||||||||||||:::|||||||||||||||||||||
Db         616 ATTATTACAGGTAATGAGGATGAATACTTTACTGCTATGATGCTGAAAAAAAGCAGGGTG 557

Qy          70 LysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu 89
               ||||||   ||||||||||||:::|||:::|||||||||   ||||||||||||||||||
```

```
Db        556 AAGTTAATTAGTCAAGAAATTGTACCATACCCATCCACATTAATGATGAGAAATTTGCTT 497

Qy         90 CysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSer 109
              :::||||||:::||||||||||   :::||||||||||||||||||||||||||||||||
Db        496 GTAGCAAATGTGAACATCTCTGGCAATAGTATATGCTTAATGACTTCACATCTAGAAAGC 437

Qy        110 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 129
              |||:::   |||:::   ||||||:::   ||||||     ||||||||||||||||
Db        436 ACCAAAGATCATTCAAAGGAGCGATTAAAGCAGCTGGACACTGTGCTTAAAAAGATGATG 377

Qy        130 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 149
              :::||||||   ||||||||||||||||||   ||||||||||||||||||||:::|||
Db        376 GATGCTCCCCCTTCAGCTACTGTAATATTTGGAGGAGATACAAATCTGAGAGATCAGGAG 317

Qy        150 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLys 169
              |||   :::   ||||||||||||||||   |||:::||||||||||||||||||||||
Db        316 GTGGCTAAAATTGGAGGTCTGCCAAATACCATATTGGATGTTTGGGAGTTTTTGGGTAAG 257

Qy        170 ProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 189
              |||:::|||||:::|||||||||||||||||||:::::|||:::||||||
Db        256 CCTGAGCATTGTCGCTATACCTGGGACACAAAACTGAACAACAATCTGAGAGCTTGCTAT 197

Qy        190 AlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIle 209
              :::||||||||||||||||||   :::|||||:::   |||          :::
Db        196 ACTAGCAGGTTGCGTTTTGACCGTATATTGTATAGAGCTTCTATGGAAGGAAGTCAAGTC 137

Qy        210 IleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              ||||||:::   |||:::|||:::|||   ||||||||||||||||||||||
Db        136 ATCCCCCAGTTTCTTAACCTTGTTGGCACAGAAAAGCTGGACTGTGGCAGATTC 83
```

RESULT 59
BG695251
LOCUS BG695251 614 bp mRNA linear EST 04-MAY-2001
DEFINITION NISC_iv14c02.w2 Soares NMBP2 pituitary Mus musculus cDNA clone IMAGE:4317675 5', mRNA sequence.
ACCESSION BG695251
VERSION BG695251.1 GI:13954981
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 614)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
MGI:1598443
Plate: LLAM9922 row: E column: 4
Seq primer: T7 primer.
FEATURES Location/Qualifiers
source 1. .614
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4317675"
/tissue_type="pituitary gland"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares NMBP2 pituitary"
/note="Organ: brain; Vector: pT7T3D-Pac; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with a NotI - oligo(dT) primer
5'-AACTGGAAGAATTCGCGGCCGCGGCGCTTTTTTTTTTTTTTTTT-3'; double-stranded cDNA was ligated to EcoRI adaptors 5'-AATTCGGCACGAGG-3' AND 5'-CCTCGTGCCG-3'(Pharmacia), digested with NotI and cloned into the NotI and EcoRI sites of the pT7T3D-PacI vector. Library went through one round of normalization, and was constructed in the laboratory of M. Bento Soares (University of Iowa)."
ORIGIN

Alignment Scores:

Pred. No.:	8.6e-87	Length:	614
Score:	807.00	Matches:	155
Percent Similarity:	88.9%	Conservative:	22
Best Local Similarity:	77.9%	Mismatches:	20
Query Match:	67.7%	Indels:	3
DB:	2	Gaps:	1

US-10-757-745-2_COPY_121_347 (1-227) x BG695251 (1-614)

```
Qy         29 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 48
              ||||||||||||||||   |||||||||   :::||||||||||||||:::::::::|||
Db         12 TTTCTCCAGGAAGTTAT-CCCCCCTACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTAC 70

Qy         49 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 68
              ||||||||||||::||||||||||||||||||||||:::||||||||||   |||
Db         71 ACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGA 130

Qy         69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 88
              ||||||   ||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        131 GTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTG 190

Qy         89 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
              |||||||||:::|||::::::   |||||||||   ||||||||||||||||||||||
Db        191 CTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAG 250

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              |||||||||   |||:::||||||||:::   |||||||||   ||||||   ||||||
Db        251 AGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATG 310

Qy        129 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 148
              ||||||||||||:::|||   ||||||||||||||||||||||||||||||||||:::
Db        311 CAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAA 370

Qy        149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
              ||||||   :::||||||||||||:::|||:::   |||   ||||||||||||||||
Db        371 GAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGC 430

Qy        169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 188
              ||||||||||||||||||||||||||||||:::   |||:::||||||   |||
Db        431 AAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCT 490

Qy        189 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 208
              ||||||   |||   ||||||||||||||||||||||||   ||||||||||||||||
Db        491 GCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCAC 544

Qy        209 IleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              :::||||||:::||||||||||:::||||||||||||||||||||||||||||||
Db        545 CTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTT 601
```

RESULT 60
BM503555
LOCUS BM503555 580 bp mRNA linear EST 12-MAR-2002
DEFINITION ih33e07.y1 Melton Mouse E16 5 Pancreas Library 2 M16B2 Mus musculus cDNA clone IMAGE:5679061 5' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION BM503555
VERSION BM503555.1 GI:18665499
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

```
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Juliana Brown
            (brown@fas.harvard.edu)
            MGI:1956939 This sequence now available from the IMAGE consortium,
            for clone orders contact: info@image.llnl.gov
            Seq primer: -40RP from Gibco
            High quality sequence stop: 432.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="ICR"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5679061"
                     /sex="Both"
                     /tissue_type="Total pancreas"
                     /dev_stage="Embryonic day 16.5"
                     /lab_host="TOP10"
                     /clone_lib="Melton Mouse E16 5 Pancreas Library 2 M16B2"
                     /note="Organ: Pancreas; Vector: pBluescript II SK; Site_1:
                     NotI; Site_2: SalI; Library constructed using SuperScript
                     Plasmid Library kit (Life  Technologies). cDNA made by
                     oligo-dT priming.Size-selected by  column fractionation;
                     average insert size 1.06kb. Primary  library,
                     unamplified."
ORIGIN

Alignment Scores:
Pred. No.:                1.04e-86        Length:          580
Score:                    806.00          Matches:         151
Percent Similarity:       89.2%           Conservative:    23
Best Local Similarity:    77.4%           Mismatches:      19
Query Match:              67.6%           Indels:          2
DB:                       3               Gaps:            1

US-10-757-745-2_COPY_121_347 (1-227) x BM503555 (1-580)

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
              ||||||:::|||||||||||||||||||||||||||||   :::|||||||||||||||:::
Db          1 GATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCA 60

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::|||   |||||||||||:::|||||||||||||||||||||:::||||||
Db         61 GCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAG 120

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||   ||||||||   ||||||||||||||||||||||||||:::||||||||||||
Db        121 AAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATG 180

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSer 105
              ||||||||||||||||:::|||:::::   |||||||||   ||||||||||||||||
Db        181 AGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 240

Qy        106 HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeu 125
              ||||||||||||||||||   |||:::|||||||||:::   |||||||||   ||||||

Db        241 CATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTT 300

Qy        126 LysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu 145
              ||||||||||||||||||:::|||   ||||||||||||||||||||||||||||||||
Db        301 GGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTA 360

Qy        146 ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGlu 165
              ||||||:::||||||   :::|||||||||||||||:::|||:::   |||   ||||||
Db        361 AGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAA 420

Qy        166 PheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeu 185
              ||||||||||||||||||||||||||||||||||||||||:::   |||:::||||||
Db        421 TTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTC 480

Qy        186 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 205
              |||   ||||||   |||   |||||||||||||||||||||||       ||||||
Db        481 AGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAA 534

Qy        206 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||||||||:::||||||:::||||||||||||:::|||||||||
Db        535 GAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 579

RESULT 61
DN097147
LOCUS       DN097147                827 bp    mRNA    linear   EST 14-FEB-2005
DEFINITION  JGI_CABE7246.fwd NIH_XGC_tropOval Xenopus tropicalis cDNA clone
            IMAGE:7825694 5', mRNA sequence.
ACCESSION   DN097147
VERSION     DN097147.1  GI:59767495
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 827)
  AUTHORS   Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C.,
            Brokstein,P. and Lindquist,E.A.
  TITLE     DOE Joint Genome Institute Xenopus tropicalis EST project
  JOURNAL   Unpublished (2004)
COMMENT     Other_ESTs: JGI_CABE7246.rev
            Contact: Lindquist,E.A., Richardson,P.
            DOE Joint Genome Institute
            2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            Tel: 925 296 5600
            Fax: 925 296 5710
            Email: cdna@jgi-psf.org
            Tissue Procurement: Robert M. Grainger
            cDNA Library Preparation: Bruce Blumberg Laboratory, University of
            California, Irvine
            DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
            Clone Distribution: I.M.A.G.E. Consortium/LLNL:
            http://image.llnl.gov
            Naming Conventions: EST name is generated by the concatenation of
            the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
            indicates a forward sequencing read of the insert. It does not
            necessarily reflect the orientation of the insert.
            Plate: CABE 0073  row: 1  column: 12
            High quality sequence stop: 820.
FEATURES             Location/Qualifiers
     source          1. .827
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /strain="N6 (Nigerian 6th generation inbred)"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:7825694"
                     /sex="Female"
                     /tissue_type="Ovary"
                     /dev_stage="Adult"
                     /lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
                     /clone_lib="NIH_XGC_tropOval"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: XhoI; The
                     library was prepared from 5 ug of poly A+ RNA by oligo-dT
```

priming
(5'-ACTAGTGCGGCCGCCTAGGCCTCGAGTTTTTTTTTTTTTTTTTV-3') and Stratascript reverse transcriptase. After ligation of EcoRI adapters (5'-AATTCGGCACGAGG-3') followed by kinasing adapters and by XhoI digestion, the cDNA was size selected by chromatography on Sepharose CL-2B columns and fractions containing cDNAs larger than 1000 bp were ligated into EcoRI/XhoI-digested pCS107. Reference for library construction: Current Genomics 4, 635-644. Library constructed by Michelle Tabb and Bruce Blumberg (Dept of Developmental and Cell Biology, University of California, Irvine)."

ORIGIN

Alignment Scores:

Pred. No.:	1.78e-86	Length:	827
Score:	806.00	Matches:	147
Percent Similarity:	87.1%	Conservative:	29
Best Local Similarity:	72.8%	Mismatches:	26
Query Match:	67.6%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x DN097147 (1-827)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||||||||||||||   :::|||:::::::|||||||||||||||||||||||||||||
Db        221 ATAGATGGCCTTGATGAATCAAATGTCGCAGAAAGGGCTCGTGGTGTGTTCCTATTTG 280

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||:::|||||||||||||||||||||||||||||    |||
Db        281 GCCCTGTACAGTCCAGATGTAGTATTCTTACAGGAGGTTATTCCACCATACTATGAGTAC 340

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::   :::|||   ||||||||||||:::|||:::   |||||||||
Db        341 CTGAAGAAGAGAGCTGTCAGCTACACAATTATTACAGGTAATGAGGATGAATACTTCACT 400

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||:::|||||||||||||||||||||||||||   ||||||||||||:::|||:::|||
Db        401 GCTATGATGCTGAAAAAGAGCAGGGTAAAGTTAATAAGCCAAGAAATTGTACCATACCCA 460

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||   ||||||||||||||||||      :::||||||:::|||||||||   :::
Db        461 TCCACGGTGATGATGAGAAATTTGCTTGTAGCAAATGTGAATATCTCTGGCAACAGTATA 520

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||:::   |||:::   ||||||:::   |||
Db        521 TGCTTAATGACTTCTCATCTAGAGAGCACAAAAGATCATTCAAAGGAGCGATTAAAGCAG 580

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||   :::|||||||||||||||   :::|||||||     |||||||||||||||
Db        581 CTGGACATTGTGCTAAAAAAGATGATGGATGCTCCCCCTTTAGCAACTGTAATATTTGGA 640

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||:::||||||   :::   ||||||:::||||||||||||||
Db        641 GGGGATACAAACCTGCGAGATCAGGAGGTTGCTAAAATTGGAGGCATGCCAAATAATATA 700

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              :::|||||||||:::|||||||||||||||||:::|||||:::|||||||||||||||:::
Db        701 TTGGATGTTTGGGACTTCTTGGGTAAGCCTGAGCACTGTCGCTATACCTGGGACACAAAA 760

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              :::|||   ||||||              |||:::|||||||||||||||||||||||||||
Db        761 GTGAACAAGAACCTGCGTGCTCCCTACATTTGCAGGTTGCGCTTTGACCGCATATTTTTT 820

Qy        201 ArgAla 202
              ||||||
Db        821 AGAGCT 826

RESULT 62
BX846715
LOCUS       BX846715                984 bp    mRNA    linear   EST 11-DEC-2003
DEFINITION  BX846715 Wellcome CRC pSK egg Xenopus laevis cDNA clone
            IMAGp998O198238 ; IMAGE:3377514 5', mRNA sequence.
ACCESSION   BX846715
VERSION     BX846715.1  GI:39736212
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 984)
  AUTHORS   Heil,O., Neubert,P., Peters,M., Radelof,U., Schneider,D.,
            Schroth,A., Korn,B. and Landgrebe,J.
  TITLE     Xenopus laevis UniGene Set 1 (RZPDLIB No.988)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998O198238.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection (amp- resistant) (RZPDLIB
            No.998) http://www.rzpd.de/cgi-
            bin/products/showLib.pl.cgi/response?libNo=998 RZPDLIB; Xenopus
            laevis UniGene Set 1 (RZPDLIB No.988) http://www.rzpd.de/cgi-
            bin/products/showLib.pl.cgi/response?libNo=988 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            T7, 5' TAATACGACTCACTATAGGG 3'.
FEATURES              Location/Qualifiers
     source           1. .984
                      /organism="Xenopus laevis"
                      /mol_type="mRNA"
                      /db_xref="taxon:8355"
                      /clone="IMAGp998O198238 ; IMAGE:3377514"
                      /tissue_type="egg"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="Wellcome CRC pSK egg"
                      /note="Vector: pBluescript SK-; Site_1: NotI; Site_2:
                      EcoRI; cDNAs were oligo-dT primed and directionally
                      cloned. Library was constructed by N. Garrett, P. LeMaire,
                      A.M. Zorn, and  J.B. Gurdon (Wellcome/CRC Institute).
                      Note: This is a Xenopus Gene Collection (XGC) library."
ORIGIN

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 2.3e-86 | Length: | 984 |
| Score: | 806.00 | Matches: | 156 |
| Percent Similarity: | 83.0% | Conservative: | 35 |
| Best Local Similarity: | 67.8% | Mismatches: | 36 |
| Query Match: | 67.6% | Indels: | 3 |
| DB: | 5 | Gaps: | 0 |

US-10-757-745-2_COPY_121_347 (1-227) x BX846715 (1-984)

```
Qy   1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||| :::|||:::::::||||||||||||| ||||||||| |||
Db   252 ATAGATGGCCTTGATGAATCAAACGTTGCAGAGAGGGCTCGTGCTGTATGTTCCTGTTTG 311

Qy   21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||:::|||||||||:::||||||||||||||||||||||||||||| |||
Db   312 GCCCTGTACACTCCAGATGTAGTATTTTTGCAGGAAGTTATTCCACCATACTACTGTGAGTAC 371

Qy   41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::   :::|||:::|||||||||||||:::|||::: |||||||||
Db   372 CTGAAGAAGAGGGCCGTCAGCTACAAAATTATTACAGGTAATGAGGATGAATACTTTACT 431

Qy   61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||:::||||||||||||||||||||||||||| ||||||||||||:::|||:::|||
Db   432 GCTATGATGCTGAAAAAAAGCAGGGTGAAGTTAATTAGTCAAGAAATTGTACCATACCCA 491

Qy   81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
```
```

```
Db        492 TCCACATTAATGATGAGAAATTTGCTTGTAGCAAATGTGAACATCTCTGGCAATAGTATA 551

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120

Db        552 TGCTTAATGACTTCACATCTAGAAAGCACCAAAGATCATTCAAAGGAGCGATTAAAGCAG 611

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140

Db        612 CTGGACACTGTGCTTAAAAAGATGATGGATGCTCCCCCTTCAGCTACTGTAATATTTGGA 671

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db        672 GGAGATACAAATCTGAGAGATCAGGAGGTGGCTAAAATTGGAGGTCTGCCAAATACCATA 731

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp-AspThrGl 180

Db        732 TTGGATGTTTGGGAGTTTTTGGGTAAGCCTGAGCATTGTCGCTATACCTGGGGACACAAA 791

Qy        180 nMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePh 200

Db        792 ACTGAACAACAATCTGAGAGCTTGCTATACTAGCAGGTTGCGTTTTGACCGTATATTGTA 851

Qy        200 eArgAlaAlaAlaGluGluGlyHisIleIle-ProArgSerLeuAspLeuLeuGlyLeuG 220

Db        852 TAGAGCTTCTATGGAAGGAAGTCAAGTCATCCCCCCAGTTTCTTAACCTTGTTGGCACAG 911

Qy        220 luLysLeu-AspCysGlyArgPhe 227

Db        912 AAAAGCTGGGACTGTGGCAGATTT 935
```

```
RESULT 63
BX753763/c
LOCUS       BX753763                749 bp    mRNA    linear   EST 10-DEC-2003
DEFINITION  BX753763 XGC-gastrula Xenopus tropicalis cDNA clone TGas102g02 3',
            mRNA sequence.
ACCESSION   BX753763
VERSION     BX753763.1  GI:39660971
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (11_2003)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Croning MDR
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TGas102g02.q1kT7
            Sequencing primer: T7
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            cDNA was oligo dT primed from 5ug of poly A+ RNA from stages 10-13
            gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
            EcoRI at the 5' end and NotI at the 3' end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1- blue.
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TGas102g02"
                     /dev_stage="gastrula (stages 10.5-12 mixed)"
                     /lab_host="Escherichia coli XL1-blue"
                     /clone_lib="XGC-gastrula"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from stages
                     10-13 gastrulae.  EcoRI-NotI cut cDNA was then ligated
                     into pCS107 with EcoRI at the 5' end and NotI at the 3'
                     end."
ORIGIN
```

```
Alignment Scores:
Pred. No.:                4.67e-86      Length:          749
Score:                    802.00        Matches:         147
Percent Similarity:       84.6%         Conservative:    29
Best Local Similarity:    70.7%         Mismatches:      32
Query Match:              67.3%         Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x BX753763 (1-749)

Qy         20 LeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSer 39

Db        749 TTGGCCCTGTACAGTCCAGATGTAGTATTCTTACAGGAGGTTATTCCACCATACTATGAG 690

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 59

Db        689 TACCTGAAGAAGAGAGCTGTCAGCTACACAATTATTACAGGTAATGAGGATGAATACTTC 630

Qy         60 ThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 79

Db        629 ACTGCTATGATGCTGAAAAAGAGCAGGGTAAAGTTAATAAGCCAAGAAATTGTACCATAC 570

Qy         80 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 99

Db        569 CCATCCACGGTGATGATGAGAAATTTGCTTGTAGCAAATGTGAATATCTCTGGCAACAGT 510

Qy        100 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 119

Db        509 ATATGCTTAATGACTTCTCATCTAGAGAGCACAAAAGATCATTCAAAGGAGCGATTAAAG 450

Qy        120 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 139

Db        449 CAGCTGGACATTGTGCTAAAAAAGATGATGGATGCTCCCCCTTTAGCAACTGTAATATTT 390

Qy        140 AlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsn 159

Db        389 GGAGGGGATACAAACCTGCGAGATCAGGAGGTTGCTAAAATTGGAGGCATGCCAAATAAT 330

Qy        160 IleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThr 179

Db        329 ATATTGGATGTTTGGGACTTCTTGGGTAAGCCTGAGCACTGTCGCTATACCTGGGACACA 270

Qy        180 GlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe 199

Db        269 AAAGTGAACAAGAACCTGCGTGCTCCCTACATTTGCAGGTTGCGCTTTGACCGCATATTT 210

Qy        200 PheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeu 219

Db        209 TTTAGAGCTTCGCAGGAGGGAAGTCAAGTGATCCCCCAGTCTCTTTATCTTGCTGGCACT 150

Qy        220 GluLysLeuAspCysGlyArgPhe 227

Db        149 GAAAAGCTGGACTGTGGCAGATTT 126
```

```
RESULT 64
CA973696
LOCUS       CA973696                919 bp    mRNA    linear   EST 27-FEB-2003
DEFINITION  AGENCOURT_10802116 Wellcome CRC pSK egg Xenopus laevis cDNA clone
            IMAGE:6326013 5', mRNA sequence.
ACCESSION   CA973696
VERSION     CA973696.1  GI:27506350
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 919)
```

```
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: N. Garrett, P. LeMaire, A.M. Zorn, and J.B.
            Gurdon (Wellcome/CRC Institute)
             cDNA Library Preparation: N. Garrett, P. LeMaire, A.M. Zorn, and
            J.B. Gurdon (Wellcome/CRC Institute)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: XGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13773  row: e  column: 22
            High quality sequence stop: 698.
FEATURES             Location/Qualifiers
     source          1. .919
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:6326013"
                     /tissue_type="egg"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Wellcome CRC pSK egg"
                     /note="Vector: pBluescript SK-; Site_1: NotI; Site_2:
                     EcoRI; cDNAs were oligo-dT primed and directionally
                     cloned. Library was constructed by N. Garrett, P. LeMaire,
                     A.M. Zorn, and  J.B. Gurdon (Wellcome/CRC Institute).
                     Note: This is a Xenopus Gene Collection (XGC) library."
ORIGIN

Alignment Scores:
Pred. No.:                1.03e-84        Length:          919
Score:                    792.00          Matches:         150
Percent Similarity:       82.4%           Conservative:    33
Best Local Similarity:    67.6%           Mismatches:      38
Query Match:              66.4%           Indels:          1
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x CA973696 (1-919)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||||||||||||||   :::|||:::::|||||||||||||   |||||||||   |||
Db        254 ATAGATGGCCTTGATGAATCAAACGTTGCAGAGAGGGCTCGTGCTGTATGTTCCTGTTTG 313

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              |||||||||:::|||||||||:::|||||||||||||||||||||||||||||     |||
Db        314 GCCCTGTACACTCCAGATGTAGTATTTTTGCAGGAAGTTATTCCACCATACTGTGAGTAC 373

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::   :::|||:::|||||||||||||:::|||:::   |||||||||
Db        374 CTGAAGAAGAGGGCCGTCAGCTACAAAATTATTACAGGTAATGAGGATGAATACTTTACT 433

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||:::|||||||||||||||||||||||||||||   ||||||||||||:::|||:::|||
Db        434 GCTATGATGCTGAAAAAAAGCAGGGTGAAGTTAATTAGTCAAGAAATTGTACCATACCCA 493

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||   ||||||||||||||||||||      :::||||||:::|||||||||   :::
Db        494 TCCACATTAATGATGAGAAATTTGCTTGTAGCAAATGTGAACATCTCTGGCAATAGTATA 553

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              |||||||||||||||||||||||||||||:::   |||:::   ||||||:::   |||
Db        554 TGCTTAATGACTTCACATCTAGAAAGCACCAAAGATCATTCAAAGGAGCGATTAAAGCAG 613

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||       ||||||||||||||||   :::|||||||   ||||||||||||||||||
Db        614 CTGGACACTGTGCTTAAAAAGATGATGGATGCTCCCCCTTCAGCTACTGTAATATTTGGA 673

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||   ||||||   :::   ||||||||||||||   |||
```

```
Db        674 GGAGATACAAATCTGAGAGATCANGAGGTGGCTAAAATTGGAGGTCTGCCAAATACCATA 733

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              :::||||||||||||||||||||||||||||:::|||||:::|||||||||||||||:::
Db        734 TTGGATGTTTGGGAGTTTTTGGGTAAGCCTGAGCATTGTCGCTATACCTGGGACACAAAA 793

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              :::|||:::||||||                 :::|||||||||||||||||   :::
Db        794 CTGAACAACAATCTGAGAGCTTGCTATACTAGCAGGTTGCGTTTTGACCGTATATTGTAT 853

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArg-SerLeuAspLeuLeuGlyLeuGl 220
              ||||||:::   |||         :::|||||||:::    |||:::|||   |||   ||
Db        854 AGAGCTTCTATGGAAGGAAGTCAAGTCATCCCCCAGTTTTCTTAACCTTGNTGGCACAGA 913

Qy        220 uLys 221
              ||||
Db        914 AAAG 917

RESULT 65
BI760756
LOCUS       BI760756                757 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603044763F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5185334 5',
            mRNA sequence.
ACCESSION   BI760756
VERSION     BI760756.1  GI:15752334
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 757)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11462  row: m  column: 15
            High quality sequence start: 6
            High quality sequence stop: 755.
FEATURES             Location/Qualifiers
     source          1. .757
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5185334"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_116"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female.  Library is
                     oligo-dT primed and  directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.4 kb,
                     insert size range 1-3 kb.  Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber  (Invitrogen).  Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                1.35e-84        Length:          757
Score:                    790.00          Matches:         161
Percent Similarity:       98.8%           Conservative:    0
Best Local Similarity:    98.8%           Mismatches:      2
```

```
Query Match:          66.3%      Indels:        2
DB:                   3          Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x BI760756 (1-757)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 328

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 GC-TTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 387

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 447

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 507

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 567

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 627

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 687

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        688 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAA-ATT 746

Qy        161 ValAspVal 163
              |||||||||
Db        747 GTGGATGTC 755

RESULT 66
BM555041
LOCUS       BM555041               1103 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6545705 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5737150
            5', mRNA sequence.
ACCESSION   BM555041
VERSION     BM555041.1  GI:18795166
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1103)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12746  row: m  column: 23
            High quality sequence stop: 678.
FEATURES             Location/Qualifiers
     source          1. .1103
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5737150"
                     /tissue_type="duodenal adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_88"
                     /note="Organ: small intestine; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                     oligo-dT primed. Average insert size 1.767 kb. Library
                     enriched for full-length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:            2.37e-84      Length:        1103
Score:                790.00        Matches:       169
Percent Similarity:   88.6%         Conservative:  2
Best Local Similarity: 87.6%        Mismatches:    14
Query Match:          66.3%         Indels:        8
DB:                   3             Gaps:          3

US-10-757-745-2_COPY_121_347 (1-227) x BM555041 (1-1103)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 449

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 509

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 569

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 629

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 689

Qy        101 CysLeuMetThrSer-HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        690 TGCCTTATGACATCCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCA 749

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |  ||||||||||||||||||||||||||||||||||||||||       ||||||||||
Db        750 GNTAAAAATGGTTTTAAAGAAAATGCAAGAAGCTCCAGAGTCACCTTACGTTATATTTGC 809

Qy        140 aGlyAspThrAsnLeuArg-AspArgGluValThrArg-CysGly-GlyLeuProAsnAs 159
              |||||||||||||||||| |||||||||   |||||| |||||| |||||||||
Db        810 AGGAGATACAAATCTAAGGGGATCGAGAGGGTACCCGAATGTGGTGGGTTTACCCACAAC 869

Qy        159 nIleValAspValTrp------GluPheLeuGlyLysProLysHisCys---GlnTyrTh 176
              ||||||:::   |||          ||||||||||||||||||   |||   :::|||
Db        870 CATTGTGGAA---TGGCTGGGGAATTTTTTGGGCAAACCTAAACTTTGCCCAGAATACAT 926

Qy        176 rTrpAspThrGlnMetAsnSerAsnLeuGly 186
                  |||   |||   |||||||||
Db        927 GGGGAAACCCCAATGGAACCCTAATCTTGGG 957

RESULT 67
CA980929
LOCUS       CA980929               900 bp    mRNA    linear   EST 27-FEB-2003
DEFINITION  AGENCOURT_11283153 Wellcome CRC pSK egg Xenopus laevis cDNA clone
            IMAGE:6869619 5', mRNA sequence.
ACCESSION   CA980929
VERSION     CA980929.1  GI:27513583
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 900)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: N. Garrett, P. LeMaire, A.M. Zorn, and J.B.
            Gurdon (Wellcome/CRC Institute)
             cDNA Library Preparation: N. Garrett, P. LeMaire, A.M. Zorn, and
            J.B. Gurdon (Wellcome/CRC Institute)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: XGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14498  row: l  column: 02
            High quality sequence start: 10
            High quality sequence stop: 786.
FEATURES             Location/Qualifiers
     source          1. .900
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:6869619"
                     /tissue_type="egg"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Wellcome CRC pSK egg"
                     /note="Vector: pBluescript SK-; Site_1: NotI; Site_2:
                     EcoRI; cDNAs were oligo-dT primed and directionally
                     cloned. Library was constructed by N. Garrett, P. LeMaire,
                     A.M. Zorn, and  J.B. Gurdon (Wellcome/CRC Institute).
                     Note: This is a Xenopus Gene Collection (XGC) library."
ORIGIN

Alignment Scores:
Pred. No.:                3.05e-84        Length:          900
Score:                    788.00          Matches:         146
Percent Similarity:       83.6%           Conservative:    33
Best Local Similarity:    68.2%           Mismatches:      35
Query Match:              66.1%           Indels:          0
DB:                       6               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x CA980929 (1-900)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||   :::|||:::::|||||||||||||   |||||||||   |||
Db        254 ATAGATGGCCTTGATGAATCAAACGTTGCAGAGAGGGCTCGTGCTGTATGTTCCTGTTTG 313

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              |||||||||:::||||||||||:::|||||||||||||||||||||||||||||   |||
Db        314 GCCCTGTACACTCCAGATGTAGTATTTTTGCAGGAAGTTATTCCACCATACTGTGAGTAC 373

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||:::   :::|||:::|||||||||||||:::|||:::   |||||||||
Db        374 CTGAAGAAGAGGGCCGTCAGCTACAAAATTATTACAGGTAATGAGGATGAATACTTTACT 433

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||:::|||||||||||||||||||||||||||   |||||||||||:::|||:::|||
Db        434 GCTATGATGCTGAAAAAAAGCAGGGTGAAGTTAATTAGTCAAGAAATTGTACCATACCCA 493

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||  ||||||||||||||||||      :::|||||:::|||||||||   :::
Db        494 TCCACATTAATGATGAGAAATTTGCTTGTAGCAAATGTGAACATCTCTGGCAATAGTATA 553

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              |||||||||||||||||||||||||||||::   |||:::   ||||||:::   |||
Db        554 TGCTTAATGACTTCACATCTAGAAAGCACCAAAGATCATTCAAAGGAGCGATTAAAGCAG 613

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||      ||||||||||||||   :::|||||||   ||||||||||||||||||
```

```
Db        614 CTGGACACTGTGCTTAAAAAGATGATGGATGCTCCCCCTTCAGCTACTGTAATATTTGGA 673

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||:::||||||   :::   ||||||||||||||   |||
Db        674 GGAGATACAAATCTGAGAGATCAGGAGGTGGCTAAAATTGGAGGTCTGCCAAATACCATA 733

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              :::||||||||||||||||||||||||||:::||||||:::|||||||||||||||:::
Db        734 TTGGATGTTTGGGAGTTTTTGGGTAAGCCTGAGCATTGTCGCTATACCTGGGACACAAAA 793

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              :::|||:::||||||                 :::|||||||||||||||||   :::
Db        794 CTGAACAACAATCTGAGAGCTTGCTATACTAGCAGGTTGCGTTTTGACCGTATATTGTAT 853

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeu 214
              ||||||:::   |||         :::|||||||:::   |||
Db        854 AGAGCTTCTATGGAAGGAAGTCAAGTCATCCCCCAGTTTCTT 895

RESULT 68
BG119064
LOCUS       BG119064                927 bp    mRNA    linear   EST 30-JAN-2001
DEFINITION  602347589F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4442399 5',
            mRNA sequence.
ACCESSION   BG119064
VERSION     BG119064.1  GI:12612570
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 927)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10215  row: a  column: 24
            High quality sequence stop: 676.
FEATURES             Location/Qualifiers
     source          1. .927
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4442399"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_90"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.7 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                2.58e-83        Length:          927
Score:                    780.50          Matches:         177
Percent Similarity:       85.8%           Conservative:    5
Best Local Similarity:    83.5%           Mismatches:      26
Query Match:              65.5%           Indels:          8
DB:                       2               Gaps:            4

US-10-757-745-2_COPY_121_347 (1-227) x BG119064 (1-927)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 252

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 312

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 372

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 432

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
               ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db        433 AGTACCAAAATGATGAGAAACTT-TTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 491

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 551

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
               ||||||||||||||||||||||||    ||||||||||||||||    |||||||||   |||
Db        552 TTAAAAATGGTTTTAAAGAAAATGCA-GAGGCTCCGGAGTCAGT-ACAGTTATATCTGCA 609

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
               ||||||||||||    ||||||||||||||||||||||||||||||    ||||||
Db        610 GGAGATACAAATCT-AGGGATCGAGAGGTTACCAGATGTGGTGGTTAC---CAAAACATT 665

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
               ||||||||||||||    ||||||                  :::   |||:::||||||
Db        666 GTGGATGTCTGGGAGGTTTTGGGGAACTTAACATTGGCAGCATTC---TGGGAGACACAA 722

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
               |||:::   |||   |||||||||          ||||||                |||   |||
Db        723 ATGGACCCTAATTCTGGGATAACTGTGGGTGTGAAACTCGGTTGATCGAATATT---TTC 779

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
               |||||||||   :::   |||   :::   ||||||
Db        780 AGAGCAGCGCGAAAAAGGGGA---CTTTTCCCGAGG 812

RESULT 69
CA971409
LOCUS       CA971409                 896 bp    mRNA    linear   EST 27-FEB-2003
DEFINITION  AGENCOURT_10764928 Wellcome CRC pSK egg Xenopus laevis cDNA clone
            IMAGE:6322803 5', mRNA sequence.
ACCESSION   CA971409
VERSION     CA971409.1  GI:27504063
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 896)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: N. Garrett, P. LeMaire, A.M. Zorn, and J.B.
            Gurdon (Wellcome/CRC Institute)
             cDNA Library Preparation: N. Garrett, P. LeMaire, A.M. Zorn, and
            J.B. Gurdon (Wellcome/CRC Institute)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: XGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13764  row: p  column: 04
            High quality sequence stop: 611.
FEATURES    Location/Qualifiers
     source 1. .896
            /organism="Xenopus laevis"
            /mol_type="mRNA"
            /db_xref="taxon:8355"
            /clone="IMAGE:6322803"
            /tissue_type="egg"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="Wellcome CRC pSK egg"
            /note="Vector: pBluescript SK-; Site_1: NotI; Site_2:
            EcoRI; cDNAs were oligo-dT primed and directionally
            cloned. Library was constructed by N. Garrett, P. LeMaire,
            A.M. Zorn, and  J.B. Gurdon (Wellcome/CRC Institute).
            Note: This is a Xenopus Gene Collection (XGC) library."
ORIGIN

Alignment Scores:
Pred. No.:              6.51e-83        Length:         896
Score:                  777.00          Matches:        144
Percent Similarity:     83.0%           Conservative:   32
Best Local Similarity:  67.9%           Mismatches:     36
Query Match:            65.2%           Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CA971409 (1-896)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||    :::|||:::::|||||||||||    |||||||||   |||
Db        255 ATAGATGGCCTTGATGAATCAAACGTTGCAGAGAGGGCTCGTGCTGTATGTTCCTGTTTG 314

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
               |||||||||:::||||||||||::|||||||||||||||||||||||||||      |||
Db        315 GCCCTGTACACTCCAGATGTAGTATTTTTGCAGGAAGTTATTCCACCATACTGTGAGTAC 374

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
               ||||||||||||:::   :::|||:::||||||||||||:::|||:::   |||||||||
Db        375 CTGAAGAAGAGGGCCGTCAGCTACAAAATTATTACAGGTAATGAGGATGAATACTTTACT 434

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               |||:::|||||||||||||||||||||||||||||||    ||||||||||||:::|||:::|||
Db        435 GCTATGATGCTGAAAAAAAGCAGGGTGAAGTTAATTAGTCAAGAAATTGTACCATACCCA 494

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
               ||||||    ||||||||||||||||||||        :::||||||:::|||||||||   :::
Db        495 TCCACATTAATGATGAGAAATTTGCTTGTAGCAAATGTGAACATCTCTGGCAATAGTATA 554

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
               |||||||||||||||||||||||||||||:::   |||:::   ||||||:::   |||
Db        555 TGCTTAATGACTTCACATCTAGAAAGCACCAAAGATCATTCANAGGAGCGATTANAGCAG 614

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
               |||       ||||||   ||||||   :::||||||   ||||||||||||||||||
Db        615 CTGGACACTGTGCTTNAAAAGATGATGGATGCTCCCCCTTCAGCTACTGTAATATTTGGA 674

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
               ||||||||||||||||||||:::||||||          ||||||||||||||   |||
Db        675 GGAGATACAAATCTGAGAGATCAGGAGGTGGCTANAATTGGAGGTCTGCCAAATACCATA 734

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
               :::|||||||||||||||||||||||||||:::||||||:::|||||||||||||||:::
Db        735 TTGGATGTTTGGGAGTTTTTGGGTAAGCCTGAGCATTGTCGCTATACCTGGGACACAAAA 794

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
               :::|||:::||||||                  :::||||||||||||||||||   :::
Db        795 CTGAACAACAATCTGAGAGCTTGCTATACTAGCAGGTTGCGTTTTGACCGTATATTGTAT 854

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
               ||||||:::   |||         :::|||||||:::
Db        855 AGAGCTTCTATGGAAGGGAGTCAAGTCATCCCCCAG 890

RESULT 70
DN100293
```

```
LOCUS       DN100293                 802 bp    mRNA    linear   EST 02-AUG-2005
DEFINITION  JGI_CABE8921.fwd NIH_XGC_tropOval Xenopus tropicalis cDNA clone
            IMAGE:7827361 5', mRNA sequence.
ACCESSION   DN100293
VERSION     DN100293.1  GI:59772177
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 802)
  AUTHORS   Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C.,
            Brokstein,P. and Lindquist,E.A.
  TITLE     DOE Joint Genome Institute Xenopus tropicalis EST project
  JOURNAL   Unpublished (2004)
COMMENT     Other_ESTs: JGI_CABE8921.rev
            Contact: Lindquist,E.A., Richardson,P.
            DOE Joint Genome Institute
            2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            Tel: 925 296 5600
            Fax: 925 296 5710
            Email: cdna@jgi-psf.org
            Tissue Procurement: Robert M. Grainger
            cDNA Library Preparation: Bruce Blumberg Laboratory, University of
            California, Irvine
            DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
            Clone Distribution: I.M.A.G.E. Consortium/LLNL:
            http://image.llnl.gov
            Naming Conventions: EST name is generated by the concatenation of
            the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
            indicates a forward sequencing read of the insert. It does not
            necessarily reflect the orientation of the insert.
            Plate: CABE 0093  row: a  column: 23
            High quality sequence stop: 741.
FEATURES             Location/Qualifiers
     source          1. .802
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /strain="N6 (Nigerian 6th generation inbred)"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:7827361"
                     /sex="Female"
                     /tissue_type="Ovary"
                     /dev_stage="Adult"
                     /lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
                     /clone_lib="NIH_XGC_tropOval"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: XhoI; The
                     library was prepared from 5 ug of poly A+ RNA by oligo-dT
                     priming
                     (5'-ACTAGTGCGGCCGCCTAGGCCTCGAGTTTTTTTTTTTTTTTTV-3') and
                     Stratascript reverse transcriptase. After ligation of
                     EcoRI adapters (5'-AATTCGGCACGAGG-3') followed by kinasing
                     adapters and by XhoI digestion, the cDNA was size selected
                     by chromatography on Sepharose CL-2B columns and fractions
                     containing cDNAs larger than 1000 bp were ligated into
                     EcoRI/XhoI-digested pCS107. Reference for library
                     construction: Current Genomics 4, 635-644. Library
                     constructed by Michelle Tabb and Bruce Blumberg (Dept of
                     Developmental and Cell Biology, University of California,
                     Irvine)."
ORIGIN
```

Alignment Scores:

Pred. No.:	2.22e-82	Length:	802
Score:	772.00	Matches:	143
Percent Similarity:	86.4%	Conservative:	29
Best Local Similarity:	71.9%	Mismatches:	26
Query Match:	64.8%	Indels:	1
DB:	8	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x DN100293 (1-802)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||   :::|||:::::::||||||||||||||||||||||||||||||
Db        208 ATAGATGGCCTTGATGAATCAAATGTCGCAGAAAGGGCTCGTGGTGTGTGTTCCTATTTG 267

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||:::|||||||||||||||||||||||||||||||   |||
Db        268 GCCCTGTACAGTCCAGATGTAGTATTCTTACAGGAGGTTATTCCACCATACTATGAGTAC 327

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||||||||||:::   :::|||   |||||||||||:::|||:::   |||||||||
Db        328 CTGAAGAAGAGAGCTGTCAGCTACACAATTATTACAGGTAATGAGGATGAATACTTCACT 387

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              |||:::|||||||||||||||||||||||||   ||||||||||||:::|||:::|||
Db        388 GCTATGATGCTGAAAAAGAGCAGGGTAAAGTTAATAAGCCAAGAAATTGTACCATACCCA 447

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||   ||||||||||||||||||   :::||||||:::|||||||   :::
Db        448 TCCACGGTGATGATGAGAAATTTGCTTGTAGCAAATGTGAATATCTCTGGCAACAGTATA 507

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||:::   |||:::   ||||||:::   |||
Db        508 TGCTTAATGACTTCTCATCTAGAGAGCACAAAAGATCATTCAAAGGAGCGATTAAAGCAG 567

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              |||   :::|||||||||||||||   :::|||||||       ||||||||||||||
Db        568 CTGGACATTGTGCTAAAAAAGATGATGGATGCTCCCCCTTTAGCAACTGTAATATTTGGA 627

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||:::||||||   :::   ||||||:::||||||||||||||
Db        628 GGGGATACAAACCTGCGAGATCAGGAGGTTGCTAAAATTGGAGGCATGCCAAATAATATA 687

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              :::|||||||||:::|||||||||||||||:::|||||||:::|||||||||||||||:::
Db        688 TTGGATGTTTGGGACTTCTTGGGTAAGCCTGAGCACTGTCGCTATACCTGGGACACAAAA 747

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg-PheAspArgIle 198
              :::|||   ||||||              |||:::|||||| ||||||||||||
Db        748 GTGAACAAGAACCTGCGTGCTCCCTACATTTGCAGGTTGCGCCTTTGACCGCATA 802
```

```
RESULT 71
CK966606
LOCUS       CK966606                 608 bp    mRNA    linear   EST 16-MAR-2004
DEFINITION  4081935 BARC 9BOV Bos taurus cDNA clone 9BOV1_O16 5', mRNA
            sequence.
ACCESSION   CK966606
VERSION     CK966606.1  GI:45484580
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 608)
  AUTHORS   Sonstegard, T.S., Van Tassell, C.P., Matukumalli, L.K., Harhay,
            G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
  TITLE     Production of EST from cDNA libraries derived from immunologically
            activated bovine gut
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Tad S. Sonstegard
            Bovine Functional Genomics Laboratory
            Animal and Natural Resources Institute
            Bdlg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
            Tel: 3015048416
            Fax: 3015048414
            Email: tads@anri.barc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            0.000925 using options -trim_alt " -trim_fasta. Vector identified
            by cross_match using options -minmatch 12 -minscore 12
            Plate: 1  row: O  column: 16
            Seq primer: CCCAGTCACGACGTTGTAAAACG
            High quality sequence stop: 608.
FEATURES             Location/Qualifiers
```

```
     source          1. .608
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /strain="Holstein"
                     /db_xref="taxon:9913"
                     /clone="9BOV1_O16"
                     /sex="Male"
                     /tissue_type="Pooled"
                     /dev_stage="Multiple"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="BARC 9BOV"
                     /note="Organ: Abomasum; Vector: pAgen-1; Site_1: EcoRV;
                     Site_2: NotI; Equimolar amounts of mRNA extracted from
                     fundic and pyloric abomasums of 18 and 21 week old steers.
                     Exposure to Ostertagia ostertagi was initiated at 15 weeks
                     of age. fundic and pyloric abomasum"
ORIGIN

Alignment Scores:
Pred. No.:                3.15e-81        Length:         608
Score:                    761.00          Matches:        147
Percent Similarity:       90.7%           Conservative:   9
Best Local Similarity:    85.5%           Mismatches:     15
Query Match:              63.8%           Indels:         1
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x CK966606 (1-608)

Qy          57 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 76
               ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db           3 GGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAGTGTAAAAGCCAATAAATT 62

Qy          77 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 96
               |||||||||||:::|||:::|||||||||||||||||||||||||||||||:::||||||
Db          63 ATTCCTTTTCCAAATACGCAAATGATGAGAAACCTTTTGTGTGTTCATGTGAGTGTGTCA 122

Qy          97 GlyAsnGluLeuCysLeuMetThrSerHis-LeuGluSerThrArgGlyHisAlaAlaGl 116
                  |||||||||||||||||||||||||| ||||||||||||||||||||||||||   ||
Db         123 CGAAATGAACTTTGCCTTATGACTTCTCACTTTGGAAAGCACCAGAGGGCACGCTAACGA 182

Qy         116 uArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaTh 136
               |||||||||||||   |||||||||||:::|||||||||||||||||||   |||   ||
Db         183 ACGAATGAATCAGTTCAAAATGGTTTTAGAGAAAATGCAAGAGGCTCCAGGATCACCTAC 242

Qy         136 rValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLe 156
               |||||||||||||||||||||||||||:::|||:::|||||||||||   ||||||||||
Db         243 TGTTATATTTGCTGGAGATACAAATCTCAAGGATCAAGAAGTTACCACATGTGGTGGTTT 302

Qy         156 uProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrTh 176
               |||||||||||:::|||||||||||||||||||||||||||||   ||||||||||||||
Db         303 ACCCAACAACATTTTGGATGTCTGGGAGTTTTTGGGCAAACCTAACCATTGCCAGTATAC 362

Qy         176 rTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAs 196
               |||||||||   |||||||||||||||||||||||   |||   |||||||||||||||||
Db         363 ATGGGATACACTAATGAACTCTAATCTTGGAATAGCTGCTAATTGTAAACTTCGTTTTGA 422

Qy         196 pArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLe 216
               |||||||||||||   ||||||||||:::   |||||||||||||||||:::||||||||||
Db         423 TCGAATATTTTTTACAGCTGCAGCAAAAGGTGGCCACATTATTCCCCAAAGTTTGGACCT 482

Qy         216 uLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
               |||||||||||||||||||||||||||   |||
Db         483 GCTTGGGCTGGAAAAACTGGACTGTGGTACATTT 516

RESULT 72
BP349936
LOCUS       BP349936                 581 bp    mRNA    linear   EST 17-SEP-2004
DEFINITION  BP349936 Sugano cDNA library, brain Homo sapiens cDNA clone
            SZR06923, mRNA sequence.
ACCESSION   BP349936
VERSION     BP349936.1  GI:52279921
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
            Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  TITLE     Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
  JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
   PUBMED   15342556
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SZR06923"
                     /tissue_type="brain"
                     /clone_lib="Sugano cDNA library, brain"
ORIGIN

Alignment Scores:
Pred. No.:                6.8e-81         Length:         581
Score:                    758.00          Matches:        148
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              63.6%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BP349936 (1-581)

Qy           1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         138 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 197

Qy          21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         198 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 257

Qy          41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         258 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 317

Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         318 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 377

Qy          81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         378 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 437

Qy         101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         438 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 497

Qy         121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         498 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 557

Qy         141 GlyAspThrAsnLeuArgAspArg 148
               |||||||||||||||||||||||
Db         558 GGAGATACAAATCTAAGGGATCGA 581

RESULT 73
BU943127
LOCUS       BU943127                 826 bp    mRNA    linear   EST 18-OCT-2002
DEFINITION  AGENCOURT_10558906 NIH_MGC_127 Homo sapiens cDNA clone
            IMAGE:6716518 5', mRNA sequence.
```

```
ACCESSION   BU943127
VERSION     BU943127.1  GI:24131946
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 826)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: NCI
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3031  row: l  column: 22
            High quality sequence stop: 490.
FEATURES             Location/Qualifiers
     source          1. .826
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6716518"
                     /tissue_type="mixed (pool of 40 RNAs)"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_127"
                     /note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
                     Site_2: SfiI (ggccgcctcggcc); Double-stranded cDNA was
                     prepared from a pool of 40 cell line polyA+ RNAs (bladder
                     - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
                     4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
                     kidnney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
                     5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
                     salivary gland - 1.3%, and skin - 2.3%).  5' and 3'
                     adaptors were used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'.  Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 1-2 kb
                     size fraction (other fractions present in NIH_MGC_126 and
                     NIH_MGC_128).  Library created in the laboratory of T.
                     Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC
                     Library."
ORIGIN

Alignment Scores:
Pred. No.:              2.01e-80        Length:         826
Score:                  756.00          Matches:        142
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            63.4%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x BU943127 (1-826)

Qy          86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSer 105
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           3 AGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCC 62

Qy         106 HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeu 125
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 CATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTA 122

Qy         126 LysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu 145
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 AAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTA 182

Qy         146 ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGlu 165
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         183 AGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAG 242

Qy         166 PheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeu 185
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         243 TTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTT 302

Qy         186 GlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGlu 205
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         303 GGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAA 362

Qy         206 GluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGly 225
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         363 GAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGT 422

Qy         226 ArgPhe 227
               ||||||
Db         423 AGATTT 428

RESULT 74
BG070243/c
LOCUS       BG070243                714 bp    mRNA    linear   EST 18-DEC-2003
DEFINITION  H3085C02-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
            H3085C02 3', mRNA sequence.
ACCESSION   BG070243
VERSION     BG070243.2  GI:40069891
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 714)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
            Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,
            Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
   PUBMED   10922068
COMMENT     On Jan 26, 2001 this sequence version replaced gi:12552812.
            Other_ESTs: H3085C02-5
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3085  row: C  column: 02
            Seq primer: -21M13 Forward
            High quality sequence stop: 714
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .714
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:H3085C02-3"
                     /db_xref="taxon:10090"
                     /clone="H3085C02"
                     /sex="Clones arrayed from a variety of cDNA libraries"
                     /dev_stage="Clones arrayed from a variety of cDNA
                     libraries"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse 15K cDNA Clone Set"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 15,247 clones from 11
                     embryo  cDNA libraries (including preimplantation stage
                     embryos from  unfertilized egg to blastocyst, embryonic
                     part of E7.5 embryos,  extraembryonic part of E7.5
                     embryos, and E12.5 female mesonephros/gonad)  and one
```

newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse develolpment, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Alignment Scores:

Pred. No.:	4.93e-80	Length:	714
Score:	752.00	Matches:	141
Percent Similarity:	88.9%	Conservative:	19
Best Local Similarity:	78.3%	Mismatches:	18
Query Match:	63.1%	Indels:	2
DB:	2	Gaps:	1

US-10-757-745-2_COPY_121_347 (1-227) x BG070243 (1-714)

```
Qy         48 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 67
              |||   |||||||||||:::|||||||||||||||||||||||||:::|||||||||
Db        711 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 652

Qy         68 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 87
              |||||||||   |||||||||||||||||||||||||:::|||||||||||||||:::
Db        651 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAGC 592

Qy         88 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 107
              |||||||||||:::|||::::::   |||||||||   |||||||||||||||||||||
Db        591 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 532

Qy        108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
              ||||||||||||   |||:::|||||||||:::   |||||||||   ||||||   |||
Db        531 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 472

Qy        128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 147
              ||||||||||||||:::|||   ||||||||||||||||||||||||||||||||||||
Db        471 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 412

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
              :::||||||   :::|||||||||||||||:::|||:::   |||   ||||||||||||
Db        411 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 352

Qy        168 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 187
              ||||||||||||||||||||||||||||||||:::   |||:::||||||   |||
Db        351 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 292

Qy        188 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 207
              ||||||   |||   |||||||||||||||||||||||||||||||||||||||||||
Db        291 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGG 238

Qy        208 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||:::||||||:::|||||||||||||:::|||||||||||||||||||||||||||||
Db        237 CACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTT 178
```

RESULT 75
BP226464
LOCUS BP226464 565 bp mRNA linear EST 15-SEP-2004
DEFINITION BP226464 Sugano cDNA library, dermoid cancer Homo sapiens cDNA clone DMC02648, mRNA sequence.
ACCESSION BP226464
VERSION BP226464.1 GI:52099369
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 565)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES Location/Qualifiers
source 1. .565
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DMC02648"
/clone_lib="Sugano cDNA library, dermoid cancer"
/note="dermoid cancer"
ORIGIN

Alignment Scores:

Pred. No.:	4.59e-80	Length:	565
Score:	751.00	Matches:	141
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	63.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x BP226464 (1-565)

```
Qy         87 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 AACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCAT 63

Qy        107 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 TTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAG 123

Qy        127 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 AAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGG 183

Qy        147 AspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPhe 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTT 243

Qy        167 LeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGly 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGA 303

Qy        187 IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAG 363

Qy        207 GlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArg 226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGA 423

Qy        227 Phe 227
              |||
Db        424 TTT 426
```

Search completed: March 11, 2006, 03:40:11
Job time : 6111.65 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 03:44:14 ; Search time 656.943 Seconds
(without alignments)
797.263 Million cell updates/sec

Title: US-10-757-745-2_COPY_121_347
Perfect score: 1192
Sequence: 1 IDGLDLNNLSERARGVCSYL.....HIIPRSLDLLGLEKLDCGRF 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7673375 seqs, 1153648444 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170339_4330/app_query.fasta_1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1676_@runat_09032006_170339_4330 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	196	16.4	952	8	US-10-750-185-25698	Sequence 25698, A
c	2	196	16.4	952	8	US-10-750-623-25698	Sequence 25698, A
	3	114.5	9.6	634	6	US-09-925-065A-783536	Sequence 783536,
	4	111	9.3	606	6	US-09-925-065A-783537	Sequence 783537,
	5	97.5	8.2	2928	7	US-10-330-773-785	Sequence 785, App
	6	97.5	8.2	47040	7	US-10-330-773-784	Sequence 784, App
	7	94.5	7.9	1769	7	US-10-330-773-788	Sequence 788, App
	8	94.5	7.9	49355	7	US-10-330-773-787	Sequence 787, App
c	9	93.5	7.8	228835	7	US-10-330-773-398	Sequence 398, App
	10	88.5	7.4	70708	7	US-10-330-773-582	Sequence 582, App
	11	88	7.4	2080	12	US-11-140-625-2	Sequence 2, Appli
c	12	87.5	7.3	190276	8	US-10-661-966-1	Sequence 1, Appli
	13	87	7.3	852	9	US-11-096-568A-29667	Sequence 29667, A
	14	87	7.3	981	9	US-11-096-568A-29182	Sequence 29182, A
	15	87	7.3	1820	12	US-11-024-959-213	Sequence 213, App
	16	86.5	7.3	3075	12	US-11-194-246-145	Sequence 145, App
	17	86	7.2	645	12	US-11-098-686-9487	Sequence 9487, Ap
c	18	86	7.2	1457619	12	US-11-098-686-8739	Sequence 8739, Ap
	19	85	7.1	600	6	US-09-925-065A-279087	Sequence 279087,
	20	85	7.1	1830	12	US-11-078-951-1	Sequence 1, Appli
	21	85	7.1	1939	12	US-11-078-951-10	Sequence 10, Appl
c	22	85	7.1	196716	9	US-11-114-798-53	Sequence 53, Appl
c	23	84.5	7.1	24861	7	US-10-330-773-248	Sequence 248, App
c	24	84	7.0	154548	12	US-11-121-086-33	Sequence 33, Appl
	25	83.5	7.0	81210	8	US-10-995-561-13295	Sequence 13295, A
	26	83.5	7.0	171732	12	US-11-121-086-98	Sequence 98, Appl
	27	83.5	7.0	645179	8	US-10-995-561-13293	Sequence 13293, A
	28	83.5	7.0	1691140	12	US-11-091-018-1	Sequence 1, Appli
	29	83	7.0	628	6	US-09-925-065A-882093	Sequence 882093,
	30	83	7.0	2368	9	US-11-096-568A-2579	Sequence 2579, Ap
	31	83	7.0	440684	7	US-10-330-773-114	Sequence 114, App
c	32	82.5	6.9	623	6	US-09-925-065A-858142	Sequence 858142,
c	33	82.5	6.9	623	6	US-09-925-065A-858143	Sequence 858143,
c	34	82.5	6.9	623	6	US-09-925-065A-858144	Sequence 858144,
	35	82.5	6.9	659	6	US-09-925-065A-509852	Sequence 509852,
c	36	82.5	6.9	110469	7	US-10-330-773-314	Sequence 314, App
c	37	82.5	6.9	172147	12	US-11-112-908-22	Sequence 22, Appl
c	38	82.5	6.9	188682	12	US-11-112-908-23	Sequence 23, Appl
	39	82	6.9	600	6	US-09-925-065A-279086	Sequence 279086,
c	40	82	6.9	118544	7	US-10-330-773-76	Sequence 76, Appl
	41	81.5	6.8	1199	8	US-10-750-185-54729	Sequence 54729, A
	42	81.5	6.8	1199	8	US-10-750-623-54729	Sequence 54729, A
c	43	81.5	6.8	1814	6	US-09-925-065A-702911	Sequence 702911,
c	44	81.5	6.8	194534	7	US-10-330-773-458	Sequence 458, App
c	45	81.5	6.8	440684	7	US-10-330-773-114	Sequence 114, App
	46	81	6.8	1466	9	US-11-096-568A-22794	Sequence 22794, A
	47	81	6.8	2532	8	US-10-645-441-11	Sequence 11, Appl
	48	81	6.8	3828	8	US-10-821-234-746	Sequence 746, App
	49	81	6.8	98862	12	US-11-121-086-76	Sequence 76, Appl
c	50	81	6.8	98862	12	US-11-121-086-76	Sequence 76, Appl
	51	81	6.8	215308	12	US-11-121-086-77	Sequence 77, Appl
c	52	81	6.8	215308	12	US-11-121-086-77	Sequence 77, Appl
c	53	80.5	6.8	593	6	US-09-925-065A-871423	Sequence 871423,
	54	80.5	6.8	791	6	US-09-925-065A-918150	Sequence 918150,
	55	80.5	6.8	791	6	US-09-925-065A-944994	Sequence 944994,
	56	80.5	6.8	2580	9	US-11-072-512-1854	Sequence 1854, Ap
	57	80.5	6.8	69648	7	US-10-330-773-911	Sequence 911, App
	58	80.5	6.8	1080000	8	US-10-928-446A-1	Sequence 1, Appli
	59	80.5	6.8	1080000	8	US-10-928-446A-181	Sequence 181, App
	60	80.5	6.8	1080000	8	US-10-928-446A-183	Sequence 183, App
	61	80.5	6.8	1080000	8	US-10-928-446A-185	Sequence 185, App
	62	80.5	6.8	1080000	8	US-10-928-446A-187	Sequence 187, App
	63	80.5	6.8	1080000	8	US-10-928-446A-189	Sequence 189, App
	64	80.5	6.8	1080000	8	US-10-928-446A-191	Sequence 191, App
	65	80.5	6.8	1080000	8	US-10-928-446A-193	Sequence 193, App
	66	80.5	6.8	1080000	8	US-10-928-446A-195	Sequence 195, App
	67	80.5	6.8	1080000	8	US-10-928-446A-197	Sequence 197, App
	68	80.5	6.8	1080000	8	US-10-928-446A-199	Sequence 199, App
	69	80.5	6.8	1080000	8	US-10-928-446A-201	Sequence 201, App
	70	80	6.7	664	6	US-09-925-065A-861077	Sequence 861077,
	71	80	6.7	1111	12	US-11-177-987-9	Sequence 9, Appli
	72	79.5	6.7	100387	12	US-11-121-086-94	Sequence 94, Appl
	73	79.5	6.7	176760	12	US-11-121-086-51	Sequence 51, Appl
c	74	79.5	6.7	187745	12	US-11-121-086-83	Sequence 83, Appl
c	75	79	6.6	1267	6	US-09-925-065A-81444	Sequence 81444, A

ALIGNMENTS

```
RESULT 1
US-10-750-185-25698/c
; Sequence 25698, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25698
;   LENGTH: 952
;   TYPE: DNA
;   ORGANISM: Bovine  19866881510436
US-10-750-185-25698

Alignment Scores:
Pred. No.:              1.05e-12        Length:         952
Score:                  196.00          Matches:        38
Percent Similarity:     91.3%           Conservative:   4
Best Local Similarity:  82.6%           Mismatches:     4
Query Match:            16.4%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-750-185-25698 (1-952)

Qy          49 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 68
               :::|||       |||   ||||||||||||||||||||||||||||||||||||||||||
Db         800 CAAATACCTGAAGGTCGTGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 741

Qy          69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 88
               ||||||  ||||||||||||||||||||||||||:::|||:::|||||||||||||||
Db         740 GTGAAGTTTAAAAGCCAAGAAATTATTCCTTTTCCAAATACGCAAATGATGAGAAACCTT 681

Qy          89 LeuCysValHisValAsn 94
               ||||||||||||||:::
Db         680 TTGTGTGTTCATGTAAGT 663

RESULT 2
US-10-750-623-25698/c
; Sequence 25698, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25698
;   LENGTH: 952
;   TYPE: DNA
;   ORGANISM: Bovine  19866881510436
US-10-750-623-25698

Alignment Scores:
Pred. No.:              1.05e-12        Length:         952
Score:                  196.00          Matches:        38
Percent Similarity:     91.3%           Conservative:   4
Best Local Similarity:  82.6%           Mismatches:     4
Query Match:            16.4%           Indels:         0
DB:                     8               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-750-623-25698 (1-952)

Qy          49 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 68
               :::|||       |||   ||||||||||||||||||||||||||||||||||||||||||
Db         800 CAAATACCTGAAGGTCGTGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 741

Qy          69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 88
               ||||||  ||||||||||||||||||||||||||:::|||:::|||||||||||||||
Db         740 GTGAAGTTTAAAAGCCAAGAAATTATTCCTTTTCCAAATACGCAAATGATGAGAAACCTT 681

Qy          89 LeuCysValHisValAsn 94
               ||||||||||||||:::
Db         680 TTGTGTGTTCATGTAAGT 663

RESULT 3
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
;   LENGTH: 634
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783536

Alignment Scores:
Pred. No.:              0.00216         Length:         634
Score:                  114.50          Matches:        27
Percent Similarity:     69.0%           Conservative:   2
Best Local Similarity:  64.3%           Mismatches:     10
Query Match:            9.6%            Indels:         3
DB:                     6               Gaps:           1

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-783536 (1-634)

Qy           1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         373 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 432

Qy          21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
               ||||||             ::::::|||         |||           |||   |||
Db         433 GCTTTGTAAGTATTATCACTTCTATTTAATAGTAATGTG---------TCTTATGTATAT 483

Qy          41 LeuLys 42
```

```
                    |||
Db        484 ACAAAG 489

RESULT 4
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-783537

Alignment Scores:
Pred. No.:                0.00524        Length:          606
Score:                    111.00         Matches:         22
Percent Similarity:       100.0%         Conservative:    0
Best Local Similarity:    100.0%         Mismatches:      0
Query Match:              9.3%           Indels:          0
DB:                       6              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-783537 (1-606)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 581

Qy         21 AlaLeu 22
              ||||||
Db        582 GCTTTG 587

RESULT 5
US-10-330-773-785
; Sequence 785, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785
;   LENGTH: 2928
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(2928)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-785

Alignment Scores:
Pred. No.:                1.59           Length:          2928
Score:                    97.50          Matches:         51
Percent Similarity:       37.1%          Conservative:    40
Best Local Similarity:    20.8%          Mismatches:      111
Query Match:              8.2%           Indels:          43
DB:                       7              Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-785 (1-2928)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              ||||||          :::         :::      |||   ||||||::::::   |||
Db       1041 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 1100

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||          |||:::          :::          |||
Db       1101 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 1148

Qy         51 IleThrGlyHisGluGluGlyTyrPhe-------------------------------- 59
                 |||::::::       :::|||
Db       1149 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 1205

Qy         60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                   |||:::         ::::::|||   ||||||
Db       1206 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 1265

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
              |||          |||          :::      :::          |||  :::   :::
Db       1266 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 1325

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                                   ||||||   :::   |||:::      :::||||||:::
Db       1326 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 1385

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              :::      |||          |||      |||          :::|||:::   :::
Db       1386 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 1445

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
              |||          :::   |||   :::      |||      ||||||   |||
Db       1446 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 1505

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                 |||                                    |||:::          :::
Db       1506 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 1565

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
              :::   |||          |||:::   :::              ||||||
Db       1566 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 1613

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                    :::      :::      |||:::      |||   |||::::::
Db       1614 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 1667

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                 |||                 :::      :::||||||||||||      |||:::
Db       1668 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 1718

Qy        222 LeuAspCysGlyArg 226
              :::         |||
Db       1719 ATTGGGCCCAACCGG 1733

RESULT 6
US-10-330-773-784
; Sequence 784, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
```

```
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 784
;   LENGTH: 47040
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(47040)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-784

Alignment Scores:
Pred. No.:                61.4          Length:          47040
Score:                    97.50         Matches:         51
Percent Similarity:       37.1%         Conservative:    40
Best Local Similarity:    20.8%         Mismatches:      111
Query Match:              8.2%          Indels:          43
DB:                       7             Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-784 (1-47040)

Qy          11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
               ||||||         :::         :::      |||   ||||||::::::   |||
Db       35153 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 35212

Qy          31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
               |||||||||          |||:::         :::          |||
Db       35213 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 35260

Qy          51 IleThrGlyHisGluGluGlyTyrPhe-------------------------------- 59
                  |||::::::      :::|||
Db       35261 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 35317

Qy          60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                   |||:::         ::::::|||   ||||||
Db       35318 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 35377

Qy          73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
               |||      |||      :::     :::      |||   :::   :::
Db       35378 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 35437

Qy          92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                        |||:::   :::   |||:::      :::|||||:::
Db       35438 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 35497

Qy         109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
               :::      |||          |||      |||          :::|||:::   :::
Db       35498 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 35557

Qy         129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
               |||          :::   |||   :::      |||      ||||||   |||
Db       35558 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 35617

Qy         147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                  |||                                ||:::          :::
Db       35618 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 35677

Qy         165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
               :::   |||         ||:::   :::          ||||||
Db       35678 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 35725

Qy         182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                     :::      :::      |||:::         |||   |||::::::
Db       35726 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 35779

Qy         202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                  |||          :::      :::|||||||||||||||      |||:::
Db       35780 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 35830

Qy         222 LeuAspCysGlyArg 226
               :::      |||
Db       35831 ATTGGGCCCAACCGG 35845
```

```
RESULT 7
US-10-330-773-788
; Sequence 788, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 788
;   LENGTH: 1769
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-788

Alignment Scores:
Pred. No.:                1.84          Length:          1769
Score:                    94.50         Matches:         50
Percent Similarity:       37.1%         Conservative:    41
Best Local Similarity:    20.4%         Mismatches:      111
Query Match:              7.9%          Indels:          43
DB:                       7             Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-788 (1-1769)

Qy          11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
               ||||||         :::         :::      |||   ||||||::::::   |||
Db         713 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 772

Qy          31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
               |||||||||          |||:::         :::          |||
Db         773 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 820

Qy          51 IleThrGlyHisGluGluGlyTyrPhe-------------------------------- 59
                  |||::::::      :::|||
Db         821 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 877

Qy          60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                   |||:::         :::::::|||   ||||||
Db         878 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 937

Qy          73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
               :::      |||          :::      :::      |||   :::   :::
Db         938 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 997

Qy          92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                        |||:::   :::   |||:::      :::|||||:::
Db         998 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 1057

Qy         109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
               :::      |||          |||      |||          :::|||:::   :::
Db        1058 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 1117

Qy         129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
               |||          :::   |||   :::      |||      ||||||   |||
Db        1118 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 1177

Qy         147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                  |||                                ||:::          :::
Db        1178 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 1237

Qy         165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
               :::   |||         ||:::   :::          ||||||
Db        1238 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 1285

Qy         182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                     :::      :::      |||:::         |||   |||::::::
Db        1286 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 1339
```

```
Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
Db       1340 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 1390

Qy        222 LeuAspCysGlyArg 226
Db       1391 ATTGGACCCAACAGG 1405

RESULT 8
US-10-330-773-787
; Sequence 787, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 787
;   LENGTH: 49355
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(49355)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-787

Alignment Scores:
Pred. No.:              147          Length:         49355
Score:                  94.50        Matches:        50
Percent Similarity:     37.1%        Conservative:   41
Best Local Similarity:  20.4%        Mismatches:     111
Query Match:            7.9%         Indels:         43
DB:                     7            Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-787 (1-49355)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
Db      38299 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 38358

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
Db      38359 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 38406

Qy         51 IleThrGlyHisGluGluGlyTyrPhe---------------------------------- 59
Db      38407 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 38463

Qy         60 ----------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
Db      38464 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 38523

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
Db      38524 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 38583

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
Db      38584 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 38643

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
Db      38644 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 38703

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
Db      38704 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 38763

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
Db      38764 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 38823

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
Db      38824 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 38871

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
Db      38872 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 38925

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
Db      38926 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 38976

Qy        222 LeuAspCysGlyArg 226
Db      38977 ATTGGACCCAACAGG 38991

RESULT 9
US-10-330-773-398/c
; Sequence 398, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 398
;   LENGTH: 228835
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-398

Alignment Scores:
Pred. No.:              1.45e+03     Length:         228835
Score:                  93.50        Matches:        45
Percent Similarity:     41.5%        Conservative:   43
Best Local Similarity:  21.2%        Mismatches:     85
Query Match:            7.8%         Indels:         39
DB:                     7            Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-398 (1-228835)

Qy          8 AsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspVal 27
Db      91221 AATGCCCCAGTTAAAAGTCACGGACTGGCAAATTGGATAAAGAGTCAAGACCCATGGGTG 91162

Qy         28 IlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsn 47
Db      91161 TGCTGTATTCAGGAG-----------------ACCCATTTCATGTGCAAAGACGCACAT 91120

Qy         48 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 67
Db      91119 AGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAACAAATGGAAAGCAAAAAAAAGCA 91060

Qy         68 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn 87
Db      91059 GGGGTTACAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGAC 91000

Qy         88 LeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCysLeuMet 103
Db      90999 AAAGAAGGCCACCACATAATGGTAAGGGGATCAATTCACCAAGAAGAGCTAACTATCCTA 90940

Qy        104 ThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMet 123
Db      90939 AATATATATGCACCCAATACA---------GGAGCAACCAGATTC------ATAAAGCAA 90895
```

```
Qy        124 ValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly------ 141
              ||||||:::   :::|||         :::|||   |||:::|||
Db      90894 GTCCTTAGAGATCTACAAAGAGACTTAGACTCCCACACAATAATAATGGGAGACTTTAAA 90835

Qy        142 ------------------AspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
                               |||   :::::::|||   :::                  |||
Db      90834 ACCCCACTGTCAGTATTAGACAGATCAATGAGACAGAAGGTTAACAAGGATATCGAGGGC 90775

Qy        156 LeuProAsn--------------AsnIleValAspValTrpGluPheLeuGlyLysPro 170
              |||   :::              :::::::::|||::::::      |||      |||
Db      90774 TTGAACTCAGCTCTGCACCAAGTAGACCTAATAGACATCTACAGAACTCTC---CACCCC 90718

Qy        171 Lys-----------------------------HisCysGlnTyrThrTrpAspThrGln 180
              |||                           ||||||  |||:::
Db      90717 AAATCAACAGAATATACATTCTTCTCAGCACCACATTGCACTTATTCTAAAATTGACCAC 90658

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              :::      :::      :::  :::  ||||||
Db      90657 TTAGTTGGAAGTAAAGCACTCCTCAGCAAATGTAAA 90622

RESULT 10
US-10-330-773-582
; Sequence 582, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 582
;   LENGTH: 70708
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-582

Alignment Scores:
Pred. No.:              1.19e+03        Length:         70708
Score:                  88.50           Matches:        44
Percent Similarity:     43.3%           Conservative:   40
Best Local Similarity:  22.7%           Mismatches:     77
Query Match:            7.4%            Indels:         33
DB:                     7               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-582 (1-70708)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::         :::         :::   :::::::::
Db      17945 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 18004

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                   |||   |||          :::||||||                :::::::|||
Db      18005 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACACATCTC 18046

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||:::::::::          |||   |||       :::   |||
Db      18047 ACGTGCAGAAACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 18106

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::|||||||   :::         :::         |||
Db      18107 GGAAAACAAAAAAAGGCAAGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 18166

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              |||||:::   |||:::             ::::::   |||   |||:::
Db      18167 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 18226

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||                |||  |||
Db      18227 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 18277

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   ||||||      :::|||          :::|||   |||:::
Db      18278 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 18331

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db      18332 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 18388

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                                 :::::::::|||:::
Db      18389 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGATCTAATAGACATC 18448

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::|||||:::
Db      18449 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 18487

RESULT 11
US-11-140-625-2
; Sequence 2, Application US/11140625
; Publication No. US20060026706A1
; GENERAL INFORMATION:
;  APPLICANT: Van Wezel, Gilles P.
;  APPLICANT:  Vijgenboom, Erik
;  TITLE OF INVENTION: A method for marker-less integration of a sequence of interest
;  TITLE OF INVENTION:  into the genome of a cell
;  FILE REFERENCE: P62377SUS00
;  CURRENT APPLICATION NUMBER: US/11/140,625
;  CURRENT FILING DATE:  2005-05-27
;  PRIOR APPLICATION NUMBER: EP 02080000.9
;  PRIOR FILING DATE: 2002-11-28
;  PRIOR APPLICATION NUMBER: PCT/NL2003/000835
;  PRIOR FILING DATE: 2003-11-27
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
;   LENGTH: 2080
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Sequence of the glkA region amplified from the S. coelicolor M145
;   OTHER INFORMATION: genome
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(2080)
;   OTHER INFORMATION: Sequence starts at nucleotide position 37 and ends at nucleotide
;   OTHER INFORMATION: position 2116
US-11-140-625-2

Alignment Scores:
Pred. No.:              13.1            Length:         2080
Score:                  88.00           Matches:        46
Percent Similarity:     37.6%           Conservative:   40
Best Local Similarity:  20.1%           Mismatches:     101
Query Match:            7.4%            Indels:         42
DB:                     12              Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-11-140-625-2 (1-2080)

Qy          5 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 24
              ::::::   :::|||   :::         :::         :::         :::
Db       1141 AACATCCGCTCACTGCGCGACGACACCGACGCCCTCGCCCGCGTCATCAAGGCCTGCGCG 1200

Qy         25 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys--- 43
              |||||:::::   |||||||||       |||   ::::::   :::   ||||||
Db       1201 CCCGACCTGGTCCTCCTCCAGGAGGCC---CCCCGCTTCTTCCGCTGGCGCAAGAAGATC 1257

Qy         44 ------------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 59
                         :::   :::         ::::::::|||
Db       1258 ACCCGGCTCGCCGCCGCCGGCGATCTGGTCCTGCTCTCCGGCGGCGGAACCGCCGCCGGC 1317

Qy         60 ThrAlaIleMetLeuLys---LysSerArgValLysLeuLysSerGlnGluIleIlePro 78
                 |||::::::         ::::::   |||:::                ::::::|||
```

```
Db       1318 CCCGCGCTGCTCTGCTCCCTGCGGGCCACCGTGGAGCGCACCGAGGACGTGCTGCTCCCG 1377

Qy         79 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 98

Db       1378 CTCACCCCCGGCCGGCACCGCCGGGGCATCGCCGCCGCCGTGGTCCGGATCGGCGGCGCC 1437

Qy         99 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 118

Db       1438 CGCATCGGCGTACTCAGCTCCCACCTGTCCCTG---------GACGCGGACGAGCGCCAC 1488

Qy        119 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 138

Db       1489 GAACAGGCCGGGCTGCTCCTCGACCACCTG---GCCGCCCTCGGCGTGAAGCACGCCGTG 1545

Qy        139 PheAlaGlyAspThrAsnLeuArg------AspArgGluValThrArgCysGlyGlyLeu 156

Db       1546 GCGGGCGGCGACCTGAACGAGCGCCCCGGCGGCCGCACCTTCCGGCGGCTCGGC------ 1599

Qy        157 ProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 176

Db       1600 ---GAAGGGCTGCGCGACTGCTGG------ACCGCCGCGCCCTGGGGCGGCGAGTACACC 1650

Qy        177 TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAsp 196

Db       1651 TTCCCC-------------------------GCGACCGCCCCGGACCGCCGGATCGAC 1683

Qy        197 ArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeu 216

Db       1684 GCGGTCTTCGTCACCGAG---------------------------------------- 1701

Qy        217 LeuGlyLeuGluLysLeuAspCysGly 225

Db       1702 ---GGCATCGAGGTGCTGGGCTGCGGA 1725
```

RESULT 12
US-10-661-966-1/c
; Sequence 1, Application US/10661966
; Publication No. US20050277118A1
; GENERAL INFORMATION:
; APPLICANT: Roth, Richard B.
; APPLICANT: Nelson, Matthew Roberts
; APPLICANT: Braun, Andreas
; TITLE OF INVENTION: METHODS FOR IDENTIFYING SUBJECTS AT RISK
; TITLE OF INVENTION: OF MELANOMA AND TREATMENTS THEREOF
; FILE REFERENCE: 524592003800
; CURRENT APPLICATION NUMBER: US/10/661,966
; CURRENT FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: 60/410,595
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: 60/422,344
; PRIOR FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 190276
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-661-966-1

Alignment Scores:

Pred. No.:	5.73e+03	Length:	190276
Score:	87.50	Matches:	43
Percent Similarity:	42.8%	Conservative:	40
Best Local Similarity:	22.2%	Mismatches:	78
Query Match:	7.3%	Indels:	33
DB:	8	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-661-966-1 (1-190276)

```
Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21

Db     163218 TTAAACGTAAATGGACTGAATGCTCCAATTAAAAGACACAGACTGGCAAACTGGATAAAG 163159
```

```
Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41

Db     163158 AGTCAAGACCCATCAGTGTGCTGCATACAGGAA----------------ACCCATCTC 163117

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61

Db     163116 ACATGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAGT 163057

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81

Db     163056 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 162997

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly---------- 97

Db     162996 ACAAAGATCAAAAGAGACCAAGAAGGCCATTACATAATGGTAAAGGGACCAATTCAACAA 162937

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117

Db     162936 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 162886

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137

Db     162885 TTC------ATAAAGCAAGTCCTTAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 162832

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145

Db     162831 ATAATG---GGAGACTTTAACATCCCACTGTCAACATTAGACAGATCAACGCGACAGAAA 162775

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163

Db     162774 GTTAATAAGGACACCCAGGAATTGAACTCAGCTCTGCACCAAGGGGACCTAATAGACATC 162715

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177

Db     162714 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 162676
```

RESULT 13
US-11-096-568A-29667
; Sequence 29667, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29667
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(852)
; OTHER INFORMATION: Ceres Seq. ID no. 4921051
US-11-096-568A-29667

Alignment Scores:

Pred. No.:	5.31	Length:	852
Score:	87.00	Matches:	48
Percent Similarity:	41.7%	Conservative:	40
Best Local Similarity:	22.7%	Mismatches:	67
Query Match:	7.3%	Indels:	56
DB:	9	Gaps:	12

US-10-757-745-2_COPY_121_347 (1-227) x US-11-096-568A-29667 (1-852)

```
Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30

Db         55 AAACGTTCGAAGGCGATTCTTGATAATATAAAGAATTTTGAAGCTGATTTCATCTGTCTA 114

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLys------ArgSerSerAsnTyr 48
```

```
Db        115 CAAGAAGTA---GACGAGTACCATAGTTTTTTCGATAGAAACATGGAAGCTCAAGGATAC 171

Qy         49 Glu---IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 67

Db        172 TCTGGAATTCCTATTGAAAATAAAGAAGGCTATGAATGCGCAATCTTTTTCAAGCCC--- 228

Qy         68 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet------ 85

Db        229 ------------AAATTTGCAGAGTTCATTACATACCAAACAACCAGAATCCAAGGATAC 276

Qy         86 ------ArgAsnLeuLeuCysValHisValAsn---------------------------- 94

Db        277 ACCAAATATGAAAACCTTTGCGTGGCGCCGAGTAGTAGTACCGTAAGTAGTGAGTCTAGT 336

Qy         95 ------ValSerGlyAsnGluLeuCysLeuMet---------------------------- 103

Db        337 GACGTTGTTAACGCTGAAGAGCTTTCTGTGGTAATGGTAGCGTTCAAGATTTTAAAACCA 396

Qy        104 ------------------------ThrSerHisLeuGluSerThrArgGlyHisAlaAla 115

Db        397 TTTAATCATGTCGTTATTATTGCAAGCTCTCATTTAAAATCCGGGAAACCTGATCGATGG 456

Qy        116 Glu------ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 133

Db        457 GATGATTTAAAGCTAGCTCAAGTCAAGACTCTAATGACTGAATTAGCCTCGTTCAAGGAG 516

Qy        134 ------------------------SerAlaThrValIlePheAlaGlyAspThrAsnLeu 145

Db        517 ATTATATCAGCGCTAACAAATTGCAGTCCATCGGTGATCCTTGCTGGTGACTTCAATTCG 576

Qy        146 ArgAsp------ArgGluValThrArgCysGlyGlyLeuProAsnAsnIle------Val 161

Db        577 AAACCCTACGTGCATAAATACATAAATTCAGACAATATACCGTCTGATATTGATCTACGA 636

Qy        162 AspValTrpGluPheLeu---GlyLysProLys 171

Db        637 AGTGTGTATGAATTTACGAAAGGAGAGCCAAGG 669
```

RESULT 14
US-11-096-568A-29182
; Sequence 29182, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29182
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(981)
; OTHER INFORMATION: Ceres Seq. ID no. 4362003
US-11-096-568A-29182

Alignment Scores:

Pred. No.:	6.4	Length:	981
Score:	87.00	Matches:	48
Percent Similarity:	41.7%	Conservative:	40
Best Local Similarity:	22.7%	Mismatches:	67
Query Match:	7.3%	Indels:	56
DB:	9	Gaps:	12

US-10-757-745-2_COPY_121_347 (1-227) x US-11-096-568A-29182 (1-981)

```
Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30

Db        184 AAACGTTCGAAGGCGATTCTTGATAATATAAAGAATTTTGAAGCTGATTTCATCTGTCTA 243

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLys------ArgSerSerAsnTyr 48

Db        244 CAAGAAGTA---GACGAGTACCATAGTTTTTTCGATAGAAACATGGAAGCTCAAGGATAC 300

Qy         49 Glu---IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 67

Db        301 TCTGGAATTCCTATTGAAAATAAAGAAGGCTATGAATGCGCAATCTTTTTCAAGCCC--- 357

Qy         68 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet------ 85

Db        358 ------------AAATTTGCAGAGTTCATTACATACCAAACAACCAGAATCCAAGGATAC 405

Qy         86 ------ArgAsnLeuLeuCysValHisValAsn---------------------------- 94

Db        406 ACCAAATATGAAAACCTTTGCGTGGCGCCGAGTAGTAGTACCGTAAGTAGTGAGTCTAGT 465

Qy         95 ------ValSerGlyAsnGluLeuCysLeuMet---------------------------- 103

Db        466 GACGTTGTTAACGCTGAAGAGCTTTCTGTGGTAATGGTAGCGTTCAAGATTTTAAAACCA 525

Qy        104 ------------------------ThrSerHisLeuGluSerThrArgGlyHisAlaAla 115

Db        526 TTTAATCATGTCGTTATTATTGCAAGCTCTCATTTAAAATCCGGGAAACCTGATCGATGG 585

Qy        116 Glu------ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 133

Db        586 GATGATTTAAAGCTAGCTCAAGTCAAGACTCTAATGACTGAATTAGCCTCGTTCAAGGAG 645

Qy        134 ------------------------SerAlaThrValIlePheAlaGlyAspThrAsnLeu 145

Db        646 ATTATATCAGCGCTAACAAATTGCAGTCCATCGGTGATCCTTGCTGGTGACTTCAATTCG 705

Qy        146 ArgAsp------ArgGluValThrArgCysGlyGlyLeuProAsnAsnIle------Val 161

Db        706 AAACCCTACGTGCATAAATACATAAATTCAGACAATATACCGTCTGATATTGATCTACGA 765

Qy        162 AspValTrpGluPheLeu---GlyLysProLys 171

Db        766 AGTGTGTATGAATTTACGAAAGGAGAGCCAAGG 798
```

RESULT 15
US-11-024-959-213
; Sequence 213, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 213
; LENGTH: 1820
; TYPE: DNA
; ORGANISM: Pinus radiata
US-11-024-959-213

Alignment Scores:

Pred. No.:	14.4	Length:	1820
Score:	87.00	Matches:	57
Percent Similarity:	36.4%	Conservative:	33

```
Best Local Similarity:  23.1%          Mismatches:     80
Query Match:            7.3%           Indels:         77
DB:                     12             Gaps:           13

US-10-757-745-2_COPY_121_347 (1-227) x US-11-024-959-213 (1-1820)

Qy         38 TyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr-GlyHisGluGluGl 57
              :::   |||::::::|||||              |||           |||:::|||||
Db        302 TTTGAATACGTAGAGAAGAGGAAGGGAGGGATAGAGACAGAGGTGCTCGCACCAGGAAGG 361

Qy         57 yTyrPheThrAlaIleMetLeuLysLysSerArg------------------------- 68
              |                    :::   |||
Db        362 G----------------------AGGGATAGAGACAGGGAGGGATAGAGACAGATACA 397

Qy         69 -------ValLysLeuLysSerGlnGluIleIleProPhe---------------ProSe 81
                  ::::::|||:::   |||         ||||||                 |||||
Db        398 GATGGATATTGAACTGGAGGATCAA---------CCCTTTGATTTGGATTTCCATCCTTC 448

Qy         81 rThrLysMetMet--------------------ArgAsnLeuLeuCysValHisValAs 94
              |      ::::::                    |||      |||       :::|||::
Db        449 TGCTCCTATTGTTGCAGTGGCACTTATTACAGGCCGCCTCCAGTTGTTTCGTTATGTAGA 508

Qy         94 nValSerGly------AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 112
              ::::|||                |||      :::|||:::|||   ||||||   |||
Db        509 CATTTCATCAGAGCCCGAAAGATTATGGACTGTGACTGCTCATACTGAGTCCTGTCGA-- 566

Qy        112 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 132
              ||||||      :::|||          ||||||              :::
Db        567 ----GCTGCTCGATTTATTAATGCTGGTAGTTCTGTGTTAACGGCGTCTCCAGACTGCTC 622

Qy        132 oGluSerAlaThrValIlePheAlaGlyAsp---------------------------- 142
              ||||||   :::      |||
Db        623 AATTCTTGCCACAAATGTTGAAACTGGACAGCCAGTAGCCAGACTTGACAATGCTCACGG 682

Qy        143 -------------------ThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLe 156
                                |||||||||        |||   |||        |||
Db        683 GGCTGCAATAAATTGCTTAACCAATTTA------ACAGAGTCAACTATTGCATCTGGAGA 736

Qy        156 uProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCys-------- 173
              |||   |||:::   ||||||:::         :::          |||
Db        737 TGAGAATGGAATCATTAAGGTTTGGGAT------ACACGACAGAATTCCTGTTGCAACAA 790

Qy        174 --------------------------------GlnTyrThrTrpAspThrGlnMet-- 181
                                              ::::::      ||||||
Db        791 GTTTAAGGCTCATGAAGACTATATTTCGGATATGGAGTTTGTTCCTGATACTATGCAGCT 850

Qy        182 -------AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePh 199
                    |||      |||       :::   |||   ||||||   :::::::::
Db        851 ACTTGGAACAAGTGGTGATGGGACCCTTTCTGTTTGCAATCTTCGTAAAAATAAGGTCCA 910

Qy        199 ePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLe 219
              |||:::          |||      ::::::      :::|||          |||
Db        911 TGCCCGATCTGAATTTTCAGAAGATGAACTGCTTTCAGTGGCCTTGATGAAGAATGGG-- 968

Qy        219 uGluLysLeuAspCysGly 225
               :::|||:::   ||||||
Db        969 -AAGAAGGTAGTATGTGGT 986

RESULT 16
US-11-194-246-145
; Sequence 145, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
;  APPLICANT: Mott, John
;  APPLICANT:  Trepod, Catherine
;  APPLICANT:  Arvidson, Staffan
;  TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND MET
;  TITLE OF INVENTION:  USE
;  FILE REFERENCE: 00592.US1 (M&R 268.05920101)
;  CURRENT APPLICATION NUMBER: US/11/194,246
;  CURRENT FILING DATE:  2005-08-01
;  PRIOR APPLICATION NUMBER: US/10/274,586
;  PRIOR FILING DATE: 2002-10-21
;  PRIOR APPLICATION NUMBER: US 60/345,438
;  PRIOR FILING DATE: 2001-10-19
;  NUMBER OF SEQ ID NOS: 621
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 145
;    LENGTH: 3075
;    TYPE: DNA
;    ORGANISM: ARTIFICIAL
;    FEATURE:
;    OTHER INFORMATION: Sequence of the alr coding sequence and flanking regions.
US-11-194-246-145

Alignment Scores:
Pred. No.:              32.9           Length:         3075
Score:                  86.50          Matches:        45
Percent Similarity:     41.5%          Conservative:   40
Best Local Similarity:  22.0%          Mismatches:     87
Query Match:            7.3%           Indels:         33
DB:                     12             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-11-194-246-145 (1-3075)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||   ::::::      ::::::|||:::|||          :::   |||   |||
Db       2444 ACGCCAGTGCTTGTTGATGGCAAAGATGTTATGCCTGAAGTCAATGCTGTGCTAGCTAAA 2503

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
              :::                |||||:::|||      :::||||||
Db       2504 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGTAAAGCC 2563

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProProSerThr 82
              :::          :::   :::         ::::::   |||:::      |||
Db       2564 ATTACGGATGTTGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 2623

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
              :::   :::|||         |||   |||   :::|||         ||||||   |||
Db       2624 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 2683

Qy         98 AsnGlu---------------------------------------LeuCysLeuMetThr 104
                                                         |||   |||:::
Db       2684 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 2743

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              |||         :::|||:::                :::                :::
Db       2744 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCTCAAAGTGCGCGTGATTGGTTA 2803

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||      ::::::            |||             |||      ||||||
Db       2804 CTAAAAGCAGCGAAAGATGAAAGTGCAGTTGCAAAACATTTTGCAGCATTATCAACCAAT 2863

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              :::|||       |||   :::   |||          ||||||:::   :::   |||
Db       2864 GCTAAAGAT------GTAGAAAAATTTGGCATTGATACCAATAATATGTTTGAATTTTGG 2917

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
              ::::::::::|||               :::|||:::   |||               |||
Db       2918 GATTGGGTTGGCGGT-----------CGTTACTCTTTATGG--------------TCA 2950

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAla 203
              :::|||::::::   |||       :::   |||:::          |||   |||
Db       2951 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAAC------TTTGAAGCGTTA 3004

Qy        204 AlaGluGluGlyHis 208
                     |||
Db       3005 TTAAATGGTGCGCAT 3019

RESULT 17
US-11-098-686-9487
; Sequence 9487, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
;  APPLICANT: Kapur, Vivek and Gebhart, Connie J.
;  TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
```

```
;  TITLE OF INVENTION:  FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
;  FILE REFERENCE: 09531-128001
;  CURRENT APPLICATION NUMBER: US/11/098,686
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: PCT/US03/31318
;  PRIOR FILING DATE: 2003-10-01
;  PRIOR APPLICATION NUMBER: US 60/416,395
;  PRIOR FILING DATE: 2002-10-04
;  NUMBER OF SEQ ID NOS: 11433
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9487
;   LENGTH: 645
;   TYPE: DNA
;   ORGANISM: Lawsonia intracellularis
US-11-098-686-9487

Alignment Scores:
Pred. No.:                4.82          Length:         645
Score:                    86.00         Matches:        35
Percent Similarity:       41.2%         Conservative:   33
Best Local Similarity:    21.2%         Mismatches:     77
Query Match:              7.2%          Indels:         20
DB:                       12            Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x US-11-098-686-9487 (1-645)

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              :::|||:::     |||   :::          |||:::::::::      |||
Db         31 CTAATGATAACAAGTAGTTTACTTGGGGCTTGTGGGCAACAACTTGTAGAAATTCCTGAA 90

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly------------ 97
                           |||      |||   :::   |||   |||
Db         91 GTA--------------TTACCTAAGGTAACAATTGGTGTTGTAGGACCTATACAGCCT 135

Qy         98 ---AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla--- 115
                        |||:::     ::::::       ||||||   |||
Db        136 ATGGGAACAACAGATCTTTTAGCTGGTTATATACCAGAAGATCGTGGTTTTGCTACAGAG 195

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              |||   :::   :::|||:::  ::::::          |||
Db        196 GAAAACCTTATAGAACTTCGCTCTTTGATTATAGATACGCTAAGAGCACAAACTAAACGT 255

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
                     |||          :::      |||      :::|||||      |||
Db        256 GAATATGAATTTATTCCAGCAGCTGATGGTGCAGATCCTACAATTACACGTATGGCTGGA 315

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              :::   |||:::   ::::::   |||      :::|||||         ::::::
Db        316 GTT---AATAGTGCACTAGAATATTGGGTTGCTGTAGGAAAACAACTAGGTGTAGAGTTT 372

Qy        176 --------------------ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 188
                               :::|||     :::      |||:::   |||:::|||
Db        373 TTGATTATACCTTATGTTTTAAGTTGGCATCCTAGAGAGGGAAGTAGTGCAGGTGTGACA 432

Qy        189 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 208
              ::::::   :::|||      |||      |||                ||||||
Db        433 TCTTCTGCTGAGTTGACTATAGAT---TTTTTTCTTATAGATACAAAAGGAGAAGGAGCA 489

Qy        209 IleIleProArgSer 213
              ::::::   ||||||
Db        490 TTACTAGGAAGATCT 504

RESULT 18
US-11-098-686-8739/c
; Sequence 8739, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
;  APPLICANT: Kapur, Vivek and Gebhart, Connie J.
;  TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
;  FILE REFERENCE: 09531-128001
;  CURRENT APPLICATION NUMBER: US/11/098,686
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: PCT/US03/31318
;  PRIOR FILING DATE: 2003-10-01
;  PRIOR APPLICATION NUMBER: US 60/416,395
;  PRIOR FILING DATE: 2002-10-04
;  NUMBER OF SEQ ID NOS: 11433
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8739
;   LENGTH: 1457619
;   TYPE: DNA
;   ORGANISM: Lawsonia intracellularis
US-11-098-686-8739

Alignment Scores:
Pred. No.:                1.25e+05      Length:         1457619
Score:                    86.00         Matches:        35
Percent Similarity:       41.2%         Conservative:   33
Best Local Similarity:    21.2%         Mismatches:     77
Query Match:              7.2%          Indels:         20
DB:                       12            Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x US-11-098-686-8739 (1-1457619)

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              :::|||:::     |||   :::          |||:::::::::      |||
Db     721928 CTAATGATAACAAGTAGTTTACTTGGGGCTTGTGGGCAACAACTTGTAGAAATTCCTGAA 721869

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly------------ 97
                           |||      |||   :::   |||   |||
Db     721868 GTA--------------TTACCTAAGGTAACAATTGGTGTTGTAGGACCTATACAGCCT 721824

Qy         98 ---AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla--- 115
                        |||:::     ::::::       ||||||   |||
Db     721823 ATGGGAACAACAGATCTTTTAGCTGGTTATATACCAGAAGATCGTGGTTTTGCTACAGAG 721764

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              |||   :::   :::|||:::  ::::::          |||
Db     721763 GAAAACCTTATAGAACTTCGCTCTTTGATTATAGATACGCTAAGAGCACAAACTAAACGT 721704

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
                     |||          :::      |||      :::|||||      |||
Db     721703 GAATATGAATTTATTCCAGCAGCTGATGGTGCAGATCCTACAATTACACGTATGGCTGGA 721644

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              :::   |||:::   ::::::   |||      :::|||||         ::::::
Db     721643 GTT---AATAGTGCACTAGAATATTGGGTTGCTGTAGGAAAACAACTAGGTGTAGAGTTT 721587

Qy        176 --------------------ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr 188
                               :::|||     :::      |||:::   |||:::|||
Db     721586 TTGATTATACCTTATGTTTTAAGTTGGCATCCTAGAGAGGGAAGTAGTGCAGGTGTGACA 721527

Qy        189 AlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHis 208
              ::::::   :::|||      |||      |||                ||||||
Db     721526 TCTTCTGCTGAGTTGACTATAGAT---TTTTTTCTTATAGATACAAAAGGAGAAGGAGCA 721470

Qy        209 IleIleProArgSer 213
              ::::::   ||||||
Db     721469 TTACTAGGAAGATCT 721455

RESULT 19
US-09-925-065A-279087
; Sequence 279087, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
```

```
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 279087
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-279087

Alignment Scores:
Pred. No.:                5.74          Length:          600
Score:                    85.00         Matches:         44
Percent Similarity:       44.3%         Conservative:    41
Best Local Similarity:    22.9%         Mismatches:      86
Query Match:              7.1%          Indels:          22
DB:                       6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-279087 (1-600)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyr---TyrSerTyrLeuLys 42
              ||||||   ::::::      :::   :::   |||   :::   :::   |||:::::: 
Db          6 TCACCAAAGCTCCTTCTGAGAAAGATTCTTTGCCCCCAGTTCCCATTTCACTATATGAGA 65

Qy         43 --------------------------LysArgSerSerAsnTyrGluIleIle------ 51
                                        |||   :::   |||:::   :::::: 
Db         66 ATGAACCCTGGAAATATCTGGAATCAGAAGAATACCAAGAACCATAGGGTTCTCATCCCA 125

Qy         52 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys--- 70
              ||||||                  |||   ::::::        ||||||:::|||    
Db        126 ACTGGGTGGACTATTGCCACTACTACAAGGGTGATATACCCCCACAGCAGACTCAAAAGA 185

Qy         71 ---------LeuLysSerGlnGluIle---IleProPheProSerThrLysMetMetArg 86
                    :::      |||   :::   ||||||   ||||||   |||:::
Db        186 CATGTATGCATAGGAAAACAGTCGCTGGGAATCCCTGCCCCATCTGCCAAGGTCACAAGT 245

Qy         87 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 106
                ::|||   :::   :::      |||   |||   |||   :::   |||
Db        246 TGCATGTTGACTTTAGGAATGTTAAGCTCTTGGAACAGTTTGTCAGTGCCCACACGGGTA 305

Qy        107 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
              |||   ||||||      :::             |||   |||   ||||||:::
Db        306 TTATCTTCCACGCT-CCATACACAGGAGTCTGTATGAAGCAGCACAAGATGTTGACCCAT 364

Qy        127 LysMetGlnGluAlaProGluSerAlaThrValIlePhe---AlaGlyAspThrAsnLeu 145
                ::|||:::|||   :::       :::   :::
Db        365 GCCATCCAGAAAGCCAGGGATCATGGTCTCCTCAGTTACCACATCTCCCAAGTTGAACCC 424

Qy        146 ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGlu 165
              ||||||   :::   :::         |||       :::       :::   |||
Db        425 CGGGACCTTGATTTTAGTGCCTCTCATGGAGCTATGAGTACTTCTCTGCCAGTCCCCACC 484

Qy        166 PheLeu---GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
                ::   |||:::|||   :::       |||:::|||       :::       |||
Db        485 CTSGTCTCAGGTGAACCCTGGTACCTATGGTACAGCTGG------AAACAGCCATCAGAG 538

Qy        185 LeuGlyIleThrAlaAlaCysLysLeuArgPheAsp 196
                  ::::::      |||:::|||      |||
Db        539 AAAGAACTGTCTCACCTTTGCCAGCTCTATCAGGAT 574

RESULT 20
US-11-078-951-1
; Sequence 1, Application US/11078951
; Publication No. US20050272119A1
; GENERAL INFORMATION:
;  APPLICANT: Kubota, Kazuishi; Nakahara, Kaori
;  APPLICANT:  Hara, Aayako; Ozeki, Yohei
;  APPLICANT:  Iijima, Yasuteru
;  TITLE OF INVENTION: 2',5'-Oligoadenylate phosphodiesterase
;  FILE REFERENCE: 05090C/HG
;  CURRENT APPLICATION NUMBER: US/11/078,951
;  CURRENT FILING DATE:  2005-03-11
;  PRIOR APPLICATION NUMBER: JP 2002-267797
;  PRIOR FILING DATE: 2002-09-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1830
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1827)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
US-11-078-951-1

Alignment Scores:
Pred. No.:                24.9          Length:          1830
Score:                    85.00         Matches:         43
Percent Similarity:       39.2%         Conservative:    40
Best Local Similarity:    20.3%         Mismatches:      77
Query Match:              7.1%          Indels:          52
DB:                       12            Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-11-078-951-1 (1-1830)

Qy         23 TyrSerProAspValIlePheLeuGlnGluVal--------------------------- 33
              |||:::   |||||||||   ||||||||||
Db       1024 TACAACGCCGATGTCATCTGTTTGCAGGAGGTTGACCGCGCAGTGTTTTCTGACAGCTTG 1083

Qy         34 IleProProTyrTyrSerTyr--------------------LeuLysLys--------- 43
              :::|||          ::::::                  :::|||:::
Db       1084 GTACCCGCCCTAGAGGCCTTCGGGCTCGAGGGGGTGTTTCGAATCAAGCAGCACGAAGGC 1143

Qy         44 ---------------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 58
                           |||   |||   :::   :::::::::   |||:::      :::
Db       1144 CTGGCCACTTTCTACCGAAAGTCTAAGTTCAGCCTTCTTAGCCAGCATGACATTTCATTC 1203

Qy         59 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 78
              :::   |||:::                  :::|||      :::::::::::::
Db       1204 TACGAAGCCCTCGAGTCCGACCCACTTCACAAAGAACTG---CTGGAGAAACTAGTTTTG 1260

Qy         79 PheProSerThr------LysMetMetArgAsnLeuLeuCysValHisValAsnVal--- 95
              :::||||||          |||:::::::::              :::   |||:::|||
Db       1261 TACCCATCAGCGCAGGAGAAGGTGCTCCAGAGATCTTCTGTTCTTCAGGTTTCAGTTCTT 1320

Qy         96 ---------------SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 110
                           |||         :::|||:::      :::|||||||
Db       1321 CAGTCTACAAAGGACTCTTCTAAAAGGATATGTGTTGCTAATACCCATCTTTACTGGCAT 1380

Qy        111 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 130
                          |||:::   |||:::   :::   |||      ::::::
Db       1381 CCTAAAGGTGGGTATATTCGCCTCATTCAAATGGCAGTAGCCTTGGCTCACATAAGACAT 1440

Qy        131 AlaPro------------GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 146
                              |||||||||   ||||||   |||
Db       1441 GTTTCATGTGATCTGTATCCTGGCATACCAGTTATATTTTGTGGGGACTTTAATAGTACA 1500

Qy        147 AspArg------------GluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
                                        |||   :::|||   :::      |||
Db       1501 CCATCAACAGGAATGTATCATTTTGTCATCAATGGCAGCATTCCAGAGGATCATGAAGAC 1560

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 182
              |||         |||:::   :::   |||
Db       1561 ---TGGGCTTCCAATGGGGAGGAGGAAAGATGC------------------------- 1590

Qy        183 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 194
                 |||:::   :::|||         ||||||:::
Db       1591 ---AATATGTCTCTTACACATTTCTTCAAGCTGAAA 1623
```

```
RESULT 21
US-11-078-951-10
; Sequence 10, Application US/11078951
; Publication No. US20050272119A1
; GENERAL INFORMATION:
;  APPLICANT: Kubota, Kazuishi; Nakahara, Kaori
;  APPLICANT:  Hara, Aayako; Ozeki, Yohei
;  APPLICANT:  Iijima, Yasuteru
;  TITLE OF INVENTION: 2',5'-Oligoadenylate phosphodiesterase
;  FILE REFERENCE: 05090C/HG
;  CURRENT APPLICATION NUMBER: US/11/078,951
;  CURRENT FILING DATE:  2005-03-11
;  PRIOR APPLICATION NUMBER: JP 2002-267797
;  PRIOR FILING DATE: 2002-09-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 1939
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-078-951-10

Alignment Scores:
Pred. No.:              26.9          Length:          1939
Score:                  85.00         Matches:         43
Percent Similarity:     39.2%         Conservative:    40
Best Local Similarity:  20.3%         Mismatches:      77
Query Match:            7.1%          Indels:          52
DB:                     12            Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-11-078-951-10 (1-1939)

Qy         23 TyrSerProAspValIlePheLeuGlnGluVal--------------------------- 33
              |||:::   |||||||||   ||||||||||||
Db       1095 TACAACGCCGATGTCATCTGTTTGCAGGAGGTTGACCGCGCAGTGTTTTCTGACAGCTTG 1154

Qy         34 IleProProTyrTyrSerTyr---------------------LeuLysLys--------- 43
              :::|||         ::::::                    :::|||:::
Db       1155 GTACCCGCCCTAGAGGCCTTCGGGCTCGAGGGGGTGTTTCGAATCAAGCAGCACGAAGGC 1214

Qy         44 ---------------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 58
                           |||   |||   :::   :::::::::   |||:::      :::
Db       1215 CTGGCCACTTTCTACCGAAAGTCTAAGTTCAGCCTTCTTAGCCAGCATGACATTTCATTC 1274

Qy         59 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 78
              :::   |||:::              :::|||      ::::::::::::
Db       1275 TACGAAGCCCTCGAGTCCGACCCACTTCACAAAGAACTG---CTGGAGAAACTAGTTTTG 1331

Qy         79 PheProSerThr------LysMetMetArgAsnLeuLeuCysValHisValAsnVal--- 95
              :::||||||         |||::::::::::              :::   |||:::|||
Db       1332 TACCCATCAGCGCAGGAGAAGGTGCTCCAGAGATCTTCTGTTCTTCAGGTTTCAGTTCTT 1391

Qy         96 ---------------SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 110
                           |||       :::|||:::     :::||||||
Db       1392 CAGTCTACAAAGGACTCTTCTAAAAGGATATGTGTTGCTAATACCCATCTTTACTGGCAT 1451

Qy        111 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 130
                        |||:::   |||:::  :::  |||      ::::::
Db       1452 CCTAAAGGTGGGTATATTCGCCTCATTCAAATGGCAGTAGCCTTGGCTCACATAAGACAT 1511

Qy        131 AlaPro------------GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArg 146
                              |||||||||   ||||||   |||
Db       1512 GTTTCATGTGATCTGTATCCTGGCATACCAGTTATATTTTGTGGGGACTTTAATAGTACA 1571

Qy        147 AspArg------------GluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
                              |||   :::|||  :::      |||
Db       1572 CCATCAACAGGAATGTATCATTTTGTCATCAATGGCAGCATTCCAGAGGATCATGAAGAC 1631

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 182
              |||       |||:::   :::   |||
Db       1632 ---TGGGCTTCCAATGGGGAGGAGGAAAGATGC--------------------------- 1661

Qy        183 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 194
                 |||:::   :::|||          ||||||:::
Db       1662 ---AATATGTCTCTTACACATTTCTTCAAGCTGAAA 1694

RESULT 22
US-11-114-798-53/c
; Sequence 53, Application US/11114798
; Publication No. US20060035246A1
; GENERAL INFORMATION:
;  APPLICANT: WU, RINA
;  APPLICANT:  MARQUEZ, ABBEY
;  TITLE OF INVENTION: CHROMOGENIC IN SITU HYBRIDIZATION METHODS, KITS, AND
;  TITLE OF INVENTION:  COMPOSITIONS
;  FILE REFERENCE: 0618.011.0004
;  CURRENT APPLICATION NUMBER: US/11/114,798
;  CURRENT FILING DATE:  2005-04-26
;  PRIOR APPLICATION NUMBER: 10/173,525
;  PRIOR FILING DATE: 2002-06-17
;  PRIOR APPLICATION NUMBER: 09/952,851
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 53
;   LENGTH: 196716
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: nucleotide sequence
US-11-114-798-53

Alignment Scores:
Pred. No.:              1.18e+04      Length:          196716
Score:                  85.00         Matches:         44
Percent Similarity:     42.6%         Conservative:    37
Best Local Similarity:  23.2%         Mismatches:      77
Query Match:            7.1%          Indels:          32
DB:                     9             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-11-114-798-53 (1-196716)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   :::::::::            |||
Db     194346 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCG 194290

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                 |||      :::||||||              :::::::|||      |||
Db     194289 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCATGTGCAGAGAC 194248

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||      :::   |||                 |||
Db     194247 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGAGCTACCAAGCCAATGGAAAACAAAAA 194188

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::       :::         |||      ||||||:::
Db     194187 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 194128

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------------GluLeuCys 101
              |||:::             ::::::   |||   |||:::             ||||||
Db     194127 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 194068

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||           |||   |||       :::
Db     194067 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 194023

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||          :::|||   |||:::|||       |||
Db     194022 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 193966

Qy        142 AspThrAsnLeu-------------------------------------------ArgAsp 147
              |||   |||                                             :::|||
Db     193965 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 193906

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
```

```
Db     193905 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATCTACAGAACTCTC 193846

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
Db     193845 ---CACCCCAAATCAACAGAATATACATTT 193819
```

RESULT 23
US-10-330-773-248/c
; Sequence 248, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 248
; LENGTH: 24861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-248

Alignment Scores:

Pred. No.:	884	Length:	24861
Score:	84.50	Matches:	44
Percent Similarity:	43.5%	Conservative:	40
Best Local Similarity:	22.8%	Mismatches:	76
Query Match:	7.1%	Indels:	33
DB:	7	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-248 (1-24861)

```
Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
Db       6489 TTAAATGTAAATGGGCTAAACGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 6430

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
Db       6429 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAG----------------ACCCATCTC 6388

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
Db       6387 ACGTGCAGAGACACACATAGGCTCAAAATAAAGGGATAGAGGAAGATCTACCAAACAAAT 6328

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
Db       6327 GGAAAACAAAAAAAAGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAATCA 6268

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
Db       6267 ACGAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 6208

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
Db       6207 GAAGAGCTAACTATCCTAAATATATATGCACCCAATATA---------GGAGCACCCAGA 6157

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
Db       6156 TTC------ATAAAGCAAGTCCTTAGAGACCTACAGAGAGACTTAGACTCCCACACAATA 6103

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
Db       6102 ATAATG---GGAGACTTTAACATCCCACTGTCAACTTTAGACAGATTAATGAGACAGAAA 6046

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
Db       6045 GTTAACAAGGATATCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACTTAATAGACATC 5986

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 176
```

```
Db       5985 TACAGAACTCTC---CACCCCAAATCAACAGAATACACT 5950
```

RESULT 24
US-11-121-086-33/c
; Sequence 33, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33
; LENGTH: 154548
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-33

Alignment Scores:

Pred. No.:	1.12e+04	Length:	154548
Score:	84.00	Matches:	45
Percent Similarity:	42.6%	Conservative:	36
Best Local Similarity:	23.7%	Mismatches:	77
Query Match:	7.0%	Indels:	32
DB:	12	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-11-121-086-33 (1-154548)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
Db      34938 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 34882

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
Db      34881 TCAGTGTGCTGTATTCAGGAA----------------ACCCATCTCACGTGCAGAGAC 34840

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db      34839 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 34780

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db      34779 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 34720

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn----------GluLeuCys 101
Db      34719 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATYCAACAAGAGGAGCTAACT 34660

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db      34659 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 34615

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db      34614 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 34558

Qy        142 AspThrAsnLeu----------------------------------------ArgAsp 147
Db      34557 GACTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGAT 34498

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
Db      34497 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 34438

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
Db      34437 ---CACCCCAAATCAACAGAATATACATTT 34411
```

```
RESULT 25
US-10-995-561-13295
; Sequence 13295, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13295
;   LENGTH: 81210
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-13295

Alignment Scores:
Pred. No.:                5.5e+03       Length:          81210
Score:                    83.50         Matches:         44
Percent Similarity:       41.8%         Conservative:    37
Best Local Similarity:    22.7%         Mismatches:      80
Query Match:              7.0%          Indels:          33
DB:                       8             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-995-561-13295 (1-81210)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              :::::||||||          |||   |||   :::       :::   :::::::::
Db      22693 GTAAATGGACTA---------AATGCTTCAATCAAAAGACACAGACTGGCAAATTGGATA 22743

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                        |||   |||       :::||||||                  ::::::
Db      22744 AAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCAT 22785

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||       |||   ::::::         |||   |||       :::   |||
Db      22786 CTCACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCA 22845

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                    ||||||:::   |||   :::       :::          |||
Db      22846 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAA 22905

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------ 98
                ||||||:::   |||:::         ::::::   |||   |||:::
Db      22906 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAA 22965

Qy         99 ------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                   ||||||   ::::::     :::   :::       |||       |||
Db      22966 CAGGAAGAGCTAACTATCCTAAATATATATGCACCCAACAATACAGGA------GCACCC 23019

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||       :::|||   ||||||     :::|||         :::|||   |||
Db      23020 AGATTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACA 23073

Qy        137 ValIlePheAlaGlyAspThrAsnLeu-------------------------------- 145
              :::|||     ||||||   |||
Db      23074 ATAATAATT---GGAGACTTTAACACCCCACTGTCAATATTAGACAGATCAACGAGACAA 23130

Qy        146 ---------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
                      :::|||   :::           ::::::::::|||
Db      23131 AAAGTTAACAAAGATAGCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGAC 23190

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 176
              ::::::      |||       |||||||      :::|||||||
Db      23191 ATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACA 23229

RESULT 26
US-11-121-086-98
```

```
; Sequence 98, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 98
;   LENGTH: 171732
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-98

Alignment Scores:
Pred. No.:                1.47e+04      Length:          171732
Score:                    83.50         Matches:         44
Percent Similarity:       41.8%         Conservative:    37
Best Local Similarity:    22.7%         Mismatches:      80
Query Match:              7.0%          Indels:          33
DB:                       12            Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-11-121-086-98 (1-171732)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              :::::||||||          |||   |||   :::       :::   :::::::::
Db     153282 GTAAATGGACTA---------AATGCTTCAATCAAAAGACACAGACTGGCAAATTGGATA 153332

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                        |||   |||       :::||||||                  ::::::
Db     153333 AAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCAT 153374

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||       |||   ::::::         |||   |||       :::   |||
Db     153375 CTCACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCA 153434

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                    ||||||:::   |||   :::       :::          |||
Db     153435 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAA 153494

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------ 98
                ||||||:::   |||:::         ::::::   |||   |||:::
Db     153495 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAA 153554

Qy         99 ------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                   ||||||   ::::::     :::   :::       |||       |||
Db     153555 CAGGAAGAGCTAACTATCCTAAATATATATGCACCCAACAATACAGGA------GCACCC 153608

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||       :::|||   ||||||     :::|||         :::|||   |||
Db     153609 AGATTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACA 153662

Qy        137 ValIlePheAlaGlyAspThrAsnLeu-------------------------------- 145
              :::|||     ||||||   |||
Db     153663 ATAATAATT---GGAGACTTTAACACCCCACTGTCAATATTAGACAGATCAACGAGACAA 153719

Qy        146 ---------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
                      :::|||   :::           ::::::::::|||
Db     153720 AAAGTTAACAAAGATAGCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGAC 153779

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 176
              ::::::      |||       |||||||      :::|||||||
Db     153780 ATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACA 153818

RESULT 27
US-10-995-561-13293
; Sequence 13293, Application US/10995561
; Publication No. US20050272054A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13293
;   LENGTH: 645179
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-13293

Alignment Scores:
Pred. No.:                8.41e+04        Length:         645179
Score:                    83.50           Matches:        44
Percent Similarity:       42.3%           Conservative:   38
Best Local Similarity:    22.7%           Mismatches:     79
Query Match:              7.0%            Indels:         33
DB:                       8               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-995-561-13293 (1-645179)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::         :::         :::   |||::::::
Db     123212 TTAAATATAAATGGACTAAATTCTGCAATTAAAAGACACAAACTGGCAAGTTGGATAAAG 123271

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                   |||   |||       :::||||||                  ::::::|||
Db     123272 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 123313

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::         |||   |||       :::   |||
Db     123314 ACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAAT 123373

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                 ||||||:::   |||   :::         :::         |||
Db     123374 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGACAAAACAGACTTTAAACCA 123433

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::         ::::::   |||   |||:::
Db     123434 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 123493

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                |||||   ::::::       :::       :::|||           |||   |||
Db     123494 GAGGAGCTAACTATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGA 123544

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                :::|||   ||||||       :::|||         :::|||   |||:::
Db     123545 TTC------ATAAAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 123598

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||       ||||||   |||
Db     123599 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 123655

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                  :::|||   :::                         :::::::::|||:::
Db     123656 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATC 123715

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::       |||       ||||||       :::|||||:::
Db     123716 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 123754

RESULT 28
US-11-091-018-1
; Sequence 1, Application US/11091018
; Publication No. US20050287551A1
; GENERAL INFORMATION:
;  APPLICANT: Gretarsdottir, Solveig
;  APPLICANT:  Thorleifsson, Gudmar
;  APPLICANT:  Gulcher, Jeffrey R.
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR HUMAN STROKE;
;  TITLE OF INVENTION:  METHODS OF TREATMENT
;  FILE REFERENCE: 2345.2010-016
;  CURRENT APPLICATION NUMBER: US/11/091,018
;  CURRENT FILING DATE:  2005-03-25
;  PRIOR APPLICATION NUMBER: PCT/US03/29906
;  PRIOR FILING DATE: 2003-09-25
;  PRIOR APPLICATION NUMBER: 10/255,120
;  PRIOR FILING DATE: 2002-09-25
;  PRIOR APPLICATION NUMBER: 10/419,723
;  PRIOR FILING DATE: 2003-04-18
;  PRIOR APPLICATION NUMBER: 10/650,120
;  PRIOR FILING DATE: 2003-08-27
;  PRIOR APPLICATION NUMBER: 10/067,514
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 09/811,352
;  PRIOR FILING DATE: 2001-03-19
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1691140
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1691140)
;   OTHER INFORMATION: n=A,T,C or G
US-11-091-018-1

Alignment Scores:
Pred. No.:                2.96e+05        Length:         1691140
Score:                    83.50           Matches:        47
Percent Similarity:       47.2%           Conservative:   45
Best Local Similarity:    24.1%           Mismatches:     67
Query Match:              7.0%            Indels:         36
DB:                       12              Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-11-091-018-1 (1-1691140)

Qy          4 LeuAspLeuAsnAsnLeu------SerGluArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||       |||   :::         :::   :::::::::
Db     311845 TTAAATGTAAATGGGCTAACTGCTTCAATTAAAAGACACAGACTGGCAAACTGGATAAAG 311904

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                   |||   |||   ::::::||||||                  ::::::|||
Db     311905 AGTCAAGACCCATCAGTGTGCTATATTCAGGAA----------------ACCCATCTC 311946

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::         |||   |||       :::   |||
Db     311947 ACGTGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 312006

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                 ||||||:::   |||   :::         :::         |||
Db     312007 GGAAAGCAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 312066

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::         ::::::   |||   |||:::
Db     312067 ACAAAGATCAAAAGAGACAAAGAAGCCCATTACATAATGGTAAAGGGATCAATTCAACAG 312126

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                |||||   ::::::       :::       :::|||                 |||
Db     312127 GAAGAGCTAACTATCCTAAATATATATGCACCTAATACA---------GGAGTACCCAGA 312177

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                :::|||   ||||||:::   ::::::         ::::::   |||:::
Db     312178 TTC------ATAAAGCAAGTCCTTAGAGACCTAGAAAGAGACTTAGACGCCCATACAATA 312231

Qy        138 IlePheAlaGlyAsp------------------------ThrAsnLeuArgAspArgGl 149
              :::       ||||||                     |||:::   |||||||||::
Db     312232 ATG------GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 312285

Qy        149 uValThrArg-CysGlyGlyLeu--------------ProAsn-AsnIleValAspVal 163
```

```
Db     312286 GTTAACAAGGATGTCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATC 312345

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 176
Db     312346 TACAGAACTCTC---CACCCCAAATCTACAGAATATACA 312381
```

RESULT 29
US-09-925-065A-882093
; Sequence 882093, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 882093
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-882093

Alignment Scores:

Pred. No.:	10.5	Length:	628
Score:	83.00	Matches:	36
Percent Similarity:	48.5%	Conservative:	28
Best Local Similarity:	27.3%	Mismatches:	42
Query Match:	7.0%	Indels:	26
DB:	6	Gaps:	7

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-882093 (1-628)

```
Qy         65 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet  84
Db          7 AAAAAAGCAAGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATC  66

Qy         85 MetArgAsn--------------LeuLeuCysValHisValAsnValSerGlyAsnGlu  99
Db         67 AAAAGACACAAAGAAGGTAATTACATAATGGTAAAGGAGGGATCAATTCAACAAGAAGAA 126

Qy        100 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 119
Db        127 CTAACTATCCTAAATATGTATGCACCCAATACA---------GGAGCACCCAGATTC--- 174

Qy        120 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePhe 139
Db        175 ---ATAAAGCAGGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACAATAATAATG 231

Qy        140 AlaGlyAspThrAsnLeu--------------------ArgAspArgGluValThrArg 152
Db        232 ---GGAGACTTTAACACCCCACTGTCAATATTAGATCAACAAGACAGAAGGTTAACAAGG 288

Qy        153 CysGly-GlyLeu--------------ProAsn-AsnIleValAspValTrpGluPheL 167
Db        289 ATATCCAGGACTTGAACGCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTC 348

Qy        167 euGlyLysProLysHisCysGlnTyrThrTrp 177
Db        349 TC---CACCCCAAATCAACAGAATATACATTC 377
```

RESULT 30
US-11-096-568A-2579
; Sequence 2579, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 2579
; LENGTH: 2368
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2368)
; OTHER INFORMATION: Ceres Seq. ID no. 15176027
US-11-096-568A-2579

Alignment Scores:

Pred. No.:	60	Length:	2368
Score:	83.00	Matches:	63
Percent Similarity:	35.3%	Conservative:	39
Best Local Similarity:	21.8%	Mismatches:	85
Query Match:	7.0%	Indels:	102
DB:	9	Gaps:	16

US-10-757-745-2_COPY_121_347 (1-227) x US-11-096-568A-2579 (1-2368)

```
Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg---GlyValCysSerTyrLeu  20
Db        796 GAAGGGTTTAACATAAATGTTCCATGGGATCAAGGACGATGTGGAGATGCAGATTATCTT 855

Qy         21 AlaLeu--------------------------------TyrSerProAspValIlePhe  29
Db        856 GCTGCATGGGACCATATCTTGATTCCTGTGGCGAGGGAATTTAACCCGGATGTTATTTTT 915

Qy         30 LeuGlnGlu-----------------------------------------ValIleProPro  36
Db        916 CTTTCAGCTGGATTTGATGCAGCTATTAATGATCCGCTAGGGGGTTGTTGTGTGACCCCA 975

Qy         37 Tyr---TyrSerTyrLeuLysLysArgSerSerAsnTyr---------GluIleIleThr  52
Db        976 TATGGATATTCAGTTATGTTGAAGAAGCTGATGGAGTTTGCACAAGGAAAGATAGTGCTG 1035

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer------------Arg  68
Db       1036 GCGTTAGAGGGAGGGTACAATCTTGACTCAATTGCAAAGTCTTCACTTGCATGTGTCCAA 1095

Qy         69 ValLysLeuLysSerGlnGluIle------------------IleProPheProSerThr  82
Db       1096 GTTTTGCTTGAAGACAAACAAATTCAAGGACCTCCGGAAGCATACCCATTCGAGTCGACA 1155

Qy         83 ------------LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn  98
Db       1156 TGGCGTGTCATACAAGCGGTGCGCAAGAGGCTATGCACATATTGGCCTTCACTTGCAGAT 1215

Qy         99 GluLeu-----------------------------------------CysLeuMetThrSer 105
Db       1216 GAATTATCGTGGAAGCTAATTAATCAGAAAACTCCTACTCCAATCATCCTTATCTCAAGC 1275

Qy        106 HisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeu 125
Db       1276 TCCGACTCTGAAACTGAAGACAACGCCCAAGGACTCCTGGATCAAATGTCG--------- 1326

Qy        126 LysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu 145
Db       1327 AAACTCAGCATTGAAAACCCTCAAGGT----------------------ACACTACTT 1362
```

```
Qy        146 ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle---------- 160
Db       1363 GAAAACCACCAAGTAGAA----------CCTGCATCTACATCTTGGAGAGCTGAT 1407

Qy        161 ---------ValAspValTrp----------GluPheLeuGlyLysProLysHis 172
Db       1408 CTCGCAAAGGTTGACGTTTGGTATGCCTCTTTTGGATCCAACATGTGGAAACCAAGGTTT 1467

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
Db       1468 CTTTGCTATATTCAAGGGGGACAGGTCGAT---------GGATTGAAAAAGGTATGTGTT 1518

Qy        193 LeuArgPheAsp----------ArgIle 198
Db       1519 GGCTCAATGGATAAAAGTCCACCAAAGGAAACAGTGTGGGAAACATTTCCACATAGATTG 1578

Qy        199 PhePheArgAlaAlaAlaGluGluGly 207
Db       1579 TTCTTTGGCCGTGAATCTTCAGTTGGT 1605
```

RESULT 31
US-10-330-773-114
; Sequence 114, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 440684
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(440684)
; OTHER INFORMATION: n = A,T,C or G
US-10-330-773-114

Alignment Scores:

Pred. No.:	5.83e+04	Length:	440684
Score:	83.00	Matches:	43
Percent Similarity:	44.3%	Conservative:	43
Best Local Similarity:	22.2%	Mismatches:	74
Query Match:	7.0%	Indels:	34
DB:	7	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-114 (1-440684)

```
Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
Db     366847 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 366906

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
Db     366907 AGTCAAGACCCATCAGTGTGCTGTGTTCAGGAA----------------ACCCATCTC 366948

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
Db     366949 ACGTGCAGAGACACAAATAGGCTCAGAATAAAGGGATGGAGGAAGATCTAC---CAAGCA 367005

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
Db     367006 AATGGAAAACAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 367065

Qy         82 ThrLysMetMetArg----------AsnLeuLeuCysValHisValAsnValSerGly 97
Db     367066 ACAAAGGTCAAAAGAGAGAAAGAAGGCCATTACATAATGGTAAAGGGTTCAATTCAACAA 367125

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
Db     367126 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 367176

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
Db     367177 TTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACAATA 367230

Qy        138 IlePheAlaGlyAspThrAsnLeuArg---------- 146
Db     367231 ATAATG---GGAGACTCTAACACCCCACTGTCAACATTAGACAGATCAACAAGACAGAAA 367287

Qy        147 ---AspArgGluValThrArgCysGlyGlyLeuProAsn------AsnIleValAspVal 163
Db     367288 GTTAACAAGGATATCCAGGAATTGAGCTCAGCTCTTAACCAAGCAGACCTAATTGACATC 367347

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
Db     367348 TACAGAACTCTC---CACCCCAAATCAGCAGAATATACATTC 367386
```

RESULT 32
US-09-925-065A-858142/c
; Sequence 858142, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 858142
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-858142

Alignment Scores:

Pred. No.:	11.8	Length:	623
Score:	82.50	Matches:	36
Percent Similarity:	41.5%	Conservative:	23
Best Local Similarity:	25.4%	Mismatches:	46
Query Match:	6.9%	Indels:	38
DB:	6	Gaps:	6

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-858142 (1-623)

```
Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
Db        491 GGTCACCAAAAAGGAAACTGCACTGCAGCTCCTCTTGCCAGA---------CAGACCCAG 441

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
Db        440 GAGCAGTCAAGGATGCCAGCTCACAAAACACCTCTAATTAGATCATGAATCTGTATAGGC 381

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
Db        380 TGTAAT----------CCCAGCACTTTGGGA 360

Qy        113 HisAlaAlaGluArgMet---------AsnGlnLeuLysMetValLeuLysLysMetGln 129
```

```
Db        359 GGCCGAGGCGGGCAGATGACTTGAGGCCAGGAGTTCAAGACCAGCCTGGCCAACATGGT- 301

Qy        130 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 149
              |||   |||   |||                                     ::::::   :::
Db        300 GAAATCCCCGTCTCTACT-----------------------------AAAAACACAAAA 271

Qy        150 ValThrArgCysGlyGlyLeu-----------------ProAsnAsnIleValAspVal 163
              :::|||:::|||||||||:::                 |||:::|||
Db        270 ATTACCCAGTGTGGTGGCGTGTGGGGGAGGCTTTAGCACGAGAATCACATAGACCCAGGA 211

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThr------TrpAspThrGlnMet 181
                  :::||||||  ||||||  |||      |||  |||:::
Db        210 GGCAGAGGTTGCAGTGAGCCAAAATTGTGCCAATGCACTCCAGCCTGGGCGACAGAGCAA 151

Qy        182 AsnSer 183
              :::|||
Db        150 GACTCT 145

RESULT 33
US-09-925-065A-858143/c
; Sequence 858143, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 858143
;   LENGTH: 623
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-858143

Alignment Scores:
Pred. No.:              11.8          Length:         623
Score:                  82.50         Matches:        36
Percent Similarity:     41.5%         Conservative:   23
Best Local Similarity:  25.4%         Mismatches:     46
Query Match:            6.9%          Indels:         38
DB:                     6             Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-858143 (1-623)

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||:::::::|||      ||||||      |||   :::         :::   :::
Db        491 GGTCACCAAAAAGGAAACTGCACTGCAGCTCCTCTTGCCAGA---------CAGACCCAG 441

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
                 |||      :::|||         |||   ::::::|||:::   :::|||:::
Db        440 GAGCAGTCAAGGATGCCAGCTCACAAAACACCTCTAATTAGATCATGAATCTGTATAGGC 381

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
                |||                                         ||||||   |||
Db        380 TGTAAT-------------------------------------CCCAGCACTTTGGGA 360

Qy        113 HisAlaAlaGluArgMet---------AsnGlnLeuLysMetValLeuLysLysMetGln 129
                  :::|||           :::   |||      |||      |||
Db        359 GGCCGAGGCGGGCAGATGACTTGAGGCCAGGAGTTCAAGACCAGCCTGGCCAACATGGT- 301

Qy        130 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 149
              |||   |||   |||                                     ::::::   :::
Db        300 GAAATCCCCGTCTCTACT-----------------------------AAAAACACAAAA 271

Qy        150 ValThrArgCysGlyGlyLeu-----------------ProAsnAsnIleValAspVal 163
              :::|||:::|||||||||:::                 |||:::|||
Db        270 ATTACCCAGTGTGGTGGCGTGTGGGGGAGGCTTTAGCACGAGAATCACATAGACCCAGGA 211

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThr------TrpAspThrGlnMet 181
                  :::||||||  ||||||  |||      |||  |||:::
Db        210 GGCAGAGGTTGCAGTGAGCCAAAATTGTGCCAATGCACTCCAGCCTGGGCGACAGAGCAA 151

Qy        182 AsnSer 183
              :::|||
Db        150 GACTCT 145

RESULT 34
US-09-925-065A-858144/c
; Sequence 858144, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 858144
;   LENGTH: 623
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-858144

Alignment Scores:
Pred. No.:              11.8          Length:         623
Score:                  82.50         Matches:        36
Percent Similarity:     41.5%         Conservative:   23
Best Local Similarity:  25.4%         Mismatches:     46
Query Match:            6.9%          Indels:         38
DB:                     6             Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-858144 (1-623)

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||:::::::|||      ||||||      |||   :::         :::   :::
Db        491 GGTCACCAAAAAGGAAACTGCACTGCAGCTCCTCTTGCCAGA---------CAGACCCAG 441

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
                 |||      :::|||         |||   ::::::|||:::   :::|||:::
Db        440 GAGCAGTCAAGGATGCCAGCTCACAAAACACCTCTAATTAGATCATGAATCTGTATAGGC 381

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
                |||                                         ||||||   |||
Db        380 TGTAAT-------------------------------------CCCAGCACTTTGGGA 360

Qy        113 HisAlaAlaGluArgMet---------AsnGlnLeuLysMetValLeuLysLysMetGln 129
                  :::|||           :::   |||      |||      |||
Db        359 GGCCGAGGCGGGCAGATGACTTGAGGCCAGGAGTTCAAGACCAGCCTGGCCAACATGGT- 301

Qy        130 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 149
              |||   |||   |||                                     ::::::   :::
```

```
Db        300 GAAATCCCCGTCTCTACT----------------------------AAAAACACAAAA 271

Qy        150 ValThrArgCysGlyGlyLeu-----------------ProAsnAsnIleValAspVal 163
              :::|||:::|||||||||:::                 |||:::|||
Db        270 ATTACCCAGTGTGGTGGCGTGTGGGGGAGGCTTTAGCACGAGAATCACATAGACCCAGGA 211

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThr------TrpAspThrGlnMet 181
                          :::||||||   ||||||  |||     |||   |||:::
Db        210 GGCAGAGGTTGCAGTGAGCCAAAATTGTGCCAATGCACTCCAGCCTGGGCGACAGAGCAA 151

Qy        182 AsnSer 183
              :::|||
Db        150 GACTCT 145

RESULT 35
US-09-925-065A-509852
; Sequence 509852, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 509852
;   LENGTH: 659
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-509852

Alignment Scores:
Pred. No.:              12.8          Length:         659
Score:                  82.50         Matches:        32
Percent Similarity:     42.2%         Conservative:   17
Best Local Similarity:  27.6%         Mismatches:     42
Query Match:            6.9%          Indels:         25
DB:                     6             Gaps:           5

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-509852 (1-659)

Qy         17 CysSerTyrLeuAlaLeu-----------------------------TyrSerProAsp 26
              |||||||:::|||   |||                            ||||||
Db        211 TGTTCATTTTTGTGTCTGTACTTCAACCTGAACACCACCGGGATCAACTACTCCCTAAAA 270

Qy         27 ValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSer 46
              :::   ||||||   |||:::   |||   :::|||   ::::::          :::|||
Db        271 ATCTGCTTCCTT---GAACTGTACCCCTTCTGGTATTTCTGGATGCATATGGAAAACTCC 327

Qy         47 --------------------------AsnTyrGluIleIleThrGlyHisGluGluGly 57
                                       :::|||      ::::::      ::::::
Db        328 ATCATGTGTGTTGTCCAGGCCCTTCCCCACTACAGGACAGTTTCTTGTTGTAAAAAAATA 387

Qy         58 TyrPheThrAlaIle------MetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              |||   |||   |||      :::|||:::   |||   ||||||         |||
Db        388 TATATAACTACAATCTAAGTCATTCTGGAGACTGTTAGATTCAAGCTG---------GAA 438

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
                 |||   |||          :::   |||   |||               |||
Db        439 TTCATACACTTTAATGGTGGCCTTTTGAAAAGACCGTTAACCAAAAGCAGTGCTAATGCC 498

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArg 111
                          ||||||::::::      |||   ||||||   |||
Db        499 TTCAATATGCATTTGTGTATTCTTTGTAYGCATAAAGAATCTCCCAGA 546

RESULT 36
US-10-330-773-314/c
; Sequence 314, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 314
;   LENGTH: 110469
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-314

Alignment Scores:
Pred. No.:              1.08e+04      Length:         110469
Score:                  82.50         Matches:        44
Percent Similarity:     42.3%         Conservative:   38
Best Local Similarity:  22.7%         Mismatches:     79
Query Match:            6.9%          Indels:         33
DB:                     7             Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-314 (1-110469)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   |||::::::
Db      32170 TTAAATGTAAACGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG 32111

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                    |||   |||       :::||||||                 :::::::|||
Db      32110 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 32069

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::         |||   |||       :::   |||
Db      32068 ACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 32009

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                    ||||||:::   |||   :::         :::          |||
Db      32008 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 31949

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::             ::::::   |||   |||:::
Db      31948 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 31889

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||               |||   |||
Db      31888 GAGGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 31838

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                    :::|||   ||||||       :::|||         :::|||   |||:::
Db      31837 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 31784

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db      31783 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 31727

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                                 :::::::::|||:::
Db      31726 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATC 31667

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||       ||||||      :::|||||:::
Db      31666 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 31628
```

```
RESULT 37
US-11-112-908-22/c
; Sequence 22, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
;  APPLICANT: Harris, Cole
;  APPLICANT:  Davis, Lisa M.
;  TITLE OF INVENTION: Breast Cancer Biomarkers
;  FILE REFERENCE: 04-164-US
;  CURRENT APPLICATION NUMBER: US/11/112,908
;  CURRENT FILING DATE:  2005-04-22
;  PRIOR APPLICATION NUMBER: US 60/564,758
;  PRIOR FILING DATE: 2004-04-23
;  PRIOR APPLICATION NUMBER: US 60/575,978
;  PRIOR FILING DATE: 2004-06-01
;  PRIOR APPLICATION NUMBER: US 60/631,702
;  PRIOR FILING DATE: 2004-11-30
;  PRIOR APPLICATION NUMBER: US 60/633,826
;  PRIOR FILING DATE: 2004-12-07
;  NUMBER OF SEQ ID NOS: 511
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
;   LENGTH: 172147
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-112-908-22

Alignment Scores:
Pred. No.:              1.93e+04        Length:         172147
Score:                  82.50           Matches:        31
Percent Similarity:     45.0%           Conservative:   28
Best Local Similarity:  23.7%           Mismatches:     47
Query Match:            6.9%            Indels:         25
DB:                     12              Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x US-11-112-908-22 (1-172147)

Qy         65 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 84
              ||||||:::   |||  :::       :::         |||      ||||||:::
Db     136463 AAAAAAGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATT 136404

Qy         85 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn----------GluLeu 100
                ||:::         ::::::   |||   |||:::             ||||||
Db     136403 AAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTA 136344

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              :::::   ::::::      :::|||     |||
Db     136343 ACTATCCTAAATACATATGCACCCAATACAGGAGCACAC--------------AGATTC 136299

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              :::|||   ||||||:::   :::|||         :::|||   |||:::|||
Db     136298 ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACAATAATAATG--- 136242

Qy        141 GlyAspThrAsnLeuArg----------------------------------AspArg 148
              ||||||   |||                                       |||:::
Db     136241 GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACAAGACAGAAAGTTGACAAG 136182

Qy        149 GluValThrArgCysGlyGlyLeuProAsn------AsnIleValAspValTrpGluPhe 166
              ::::::                  |||      :::::::::|||::::::
Db     136181 GATATCCAGGAATTGAACTCAACTCTGAACCAAGCGGACCTAATAGACATCTACAGAACT 136122

Qy        167 LeuGlyLysProLysHisCysGlnTyrThrTrp 177
              |||      ||||||      :::|||||:::
Db     136121 CTC---CACCCCAAATCAACAGAATACACATTC 136092

RESULT 38
US-11-112-908-23/c
; Sequence 23, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
;  APPLICANT: Harris, Cole
;  APPLICANT:  Davis, Lisa M.
;  TITLE OF INVENTION: Breast Cancer Biomarkers
;  FILE REFERENCE: 04-164-US
;  CURRENT APPLICATION NUMBER: US/11/112,908
;  CURRENT FILING DATE:  2005-04-22
;  PRIOR APPLICATION NUMBER: US 60/564,758
;  PRIOR FILING DATE: 2004-04-23
;  PRIOR APPLICATION NUMBER: US 60/575,978
;  PRIOR FILING DATE: 2004-06-01
;  PRIOR APPLICATION NUMBER: US 60/631,702
;  PRIOR FILING DATE: 2004-11-30
;  PRIOR APPLICATION NUMBER: US 60/633,826
;  PRIOR FILING DATE: 2004-12-07
;  NUMBER OF SEQ ID NOS: 511
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
;   LENGTH: 188682
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-112-908-23

Alignment Scores:
Pred. No.:              2.18e+04        Length:         188682
Score:                  82.50           Matches:        31
Percent Similarity:     45.0%           Conservative:   28
Best Local Similarity:  23.7%           Mismatches:     47
Query Match:            6.9%            Indels:         25
DB:                     12              Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x US-11-112-908-23 (1-188682)

Qy         65 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 84
              ||||||:::   |||  :::       :::         |||      ||||||:::
Db      86084 AAAAAAGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATT 86025

Qy         85 MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn----------GluLeu 100
                ||:::         ::::::   |||   |||:::             ||||||
Db      86024 AAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTA 85965

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              :::::   ::::::      :::|||     |||
Db      85964 ACTATCCTAAATACATATGCACCCAATACAGGAGCACAC--------------AGATTC 85920

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              :::|||   ||||||:::   :::|||         :::|||   |||:::|||
Db      85919 ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACAATAATAATG--- 85863

Qy        141 GlyAspThrAsnLeuArg----------------------------------AspArg 148
              ||||||   |||                                       |||:::
Db      85862 GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACAAGACAGAAAGTTGACAAG 85803

Qy        149 GluValThrArgCysGlyGlyLeuProAsn------AsnIleValAspValTrpGluPhe 166
              ::::::                  |||      :::::::::|||::::::
Db      85802 GATATCCAGGAATTGAACTCAACTCTGAACCAAGCGGACCTAATAGACATCTACAGAACT 85743

Qy        167 LeuGlyLysProLysHisCysGlnTyrThrTrp 177
              |||      ||||||      :::|||||:::
Db      85742 CTC---CACCCCAAATCAACAGAATACACATTC 85713

RESULT 39
US-09-925-065A-279086
; Sequence 279086, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
```

```
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 279086
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-279086

Alignment Scores:
Pred. No.:              12.9          Length:         600
Score:                  82.00         Matches:        44
Percent Similarity:     43.8%         Conservative:   40
Best Local Similarity:  22.9%         Mismatches:     87
Query Match:            6.9%          Indels:         22
DB:                     6             Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-279086 (1-600)

Qy          24 SerProAspValIlePheLeuGlnGluValIleProProTyr---TyrSerTyrLeuLys 42
               ||||||   ::::::      :::   :::   |||   :::   :::   |||:::::: 
Db           6 TCACCAAAGCTCCTTCTGAGAAAGATTCTTTGCCCCCAGTTCCCATTTCACTATATGAGA 65

Qy          43 --------------------------LysArgSerSerAsnTyrGluIleIle------ 51
                                         |||   :::   |||:::   :::::: 
Db          66 ATGAACCCTGGAAATATCTGGAATCAGAAGAATACCAAGAACCATAGGGTTCTCATCCCA 125

Qy          52 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys--- 70
               ||||||                |||   ::::::        ||||||:::|||
Db         126 ACTGGGTGGACTATTGCCACTACTACAAGGGTGATATACCCCCACAGCAGACTCAAAAGA 185

Qy          71 ---------LeuLysSerGlnGluIle---IleProPheProSerThrLysMetMetArg 86
                      :::      |||   :::   ||||||   ||||||   |||:::
Db         186 CATGTATGCATAGGAAAACAGTCGCTGGGAATCCCTGCCCCATCTGCCAAGGTCACAAGT 245

Qy          87 AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 106
               :::|||   :::   :::       |||   |||   |||   :::   |||
Db         246 TGCATGTTGACTTTAGGAATGTTAAGCTCTTGGAACAGTTTGTCAGTGCCCACACGGGTA 305

Qy         107 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
               |||   ||||||      :::           |||   |||   ||||||:::
Db         306 TTATCTTCCACGCT-CCATACACAGGAGTCTGTATGAAGCAGCACAAGATGTTGACCCAT 364

Qy         127 LysMetGlnGluAlaProGluSerAlaThrValIlePhe---AlaGlyAspThrAsnLeu 145
               :::|||:::|||   :::         :::   :::
Db         365 GCCATCCAGAAAGCCAGGGATCATGGTCTCCTCAGTTACCACATCTCCCAAGTTGAACCC 424

Qy         146 ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGlu 165
               ||||||          :::         |||   :::   :::   |||
Db         425 CGGGACCTTGRTTTTAGTGCCTCTCATGGAGCTATGAGTACTTCTCTGCCAGTCCCCACC 484

Qy         166 PheLeu---GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
               :::   |||:::|||   :::       |||:::|||       :::       |||
Db         485 CTCGTCTCAGGTGAACCCTGGTACCTATGGTACAGCTGG------AAACAGCCATCAGAG 538

Qy         185 LeuGlyIleThrAlaAlaCysLysLeuArgPheAsp 196
               ::::::      |||:::|||       |||
Db         539 AAAGAACTGTCTCACCTTTGCCAGCTCTATCAGGAT 574

RESULT 40
US-10-330-773-76/c
; Sequence 76, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76
;   LENGTH: 118544
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-76

Alignment Scores:
Pred. No.:              1.36e+04      Length:         118544
Score:                  82.00         Matches:        47
Percent Similarity:     45.6%         Conservative:   42
Best Local Similarity:  24.1%         Mismatches:     72
Query Match:            6.9%          Indels:         34
DB:                     7             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-76 (1-118544)

Qy           1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
               :::::::||||||          |||   |||   :::         :::   :::::::::
Db       98772 GTAAATGGGCTA---------AATGCTTCAATTAAAAGACACAGACTGGCAAACTGGATA 98722

Qy          21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                           |||   |||       :::   |||                 ::::::
Db       98721 AAGAGTCAAGACCCATCAGTGTGCTGTATTTAGGAA-----------------ACCCAT 98680

Qy          41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
               |||       |||   ::::::          |||   |||       :::   |||
Db       98679 CTCATATGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCA 98620

Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                        ||||||:::   |||   :::         :::         |||
Db       98619 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAA 98560

Qy          81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------ 98
               |||||:::   |||:::           ::::::   |||   |||:::
Db       98559 CCGACAAAGATCAAAAGACACAAAGAAGGCTATTACATAATGGTAAAGGGATCAATTCAA 98500

Qy          99 ------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                     ||||||   ::::::      :::      :::|||           |||
Db       98499 CAAGAAGAGCTAACTATCCTGAATATATATGCACCCAATACA---------GGAGCACCC 98449

Qy         117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
               |||         :::|||   ||||||:::   :::|||         :::      |||
Db       98448 AGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTACCACACA 98395

Qy         137 ValIlePheAlaGlyAspThrAsnLeu--------------------ArgAspArgGlu 149
               :::|||      ||||||  |||                     :::|||||:::
Db       98394 ATAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGATCAACAAGACAGAAAG 98338

Qy         150 ValThrArgCysGly-GlyLeu--------------ProAsn-AsnIleValAspValT 164
               :::||||||       |||:::              |||::: ::::::::::|||::::
Db       98337 ATAACAAGGATATCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCT 98278

Qy         164 rpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
               ::      |||      ||||||      :::|||||:::
Db       98277 ACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 98240

RESULT 41
US-10-750-185-54729
; Sequence 54729, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
```

```
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 54729
;    LENGTH: 1199
;    TYPE: DNA
;    ORGANISM: Bovine  19866881341043
US-10-750-185-54729

Alignment Scores:
Pred. No.:                36.7          Length:          1199
Score:                    81.50         Matches:         54
Percent Similarity:       36.3%         Conservative:    31
Best Local Similarity:    23.1%         Mismatches:      64
Query Match:              6.8%          Indels:          85
DB:                       8             Gaps:            14

US-10-757-745-2_COPY_121_347 (1-227) x US-10-750-185-54729 (1-1199)

Qy          21 AlaLeuTyrSerProAspValIle-----------------------PheLeuGlnGlu 32
               |||:::|||:::|||   :::|||                       |||||||||
Db          18 GCAATGTATAACCCAGGGTTAATCCAATGCTTTAACATGATTTCTTCTTTTTTGCAGACT 77

Qy          33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
                  :::    :::|||   |||     :::   |||   |||          :::
Db          78 AATGTTGGAAAGTTTTATTTGTATTTCGGAAGGCATTCTCAAAATGTAAATGCTTTGGAG 137

Qy          53 GlyHisGlu---------------------------------------------------- 55
                  |||
Db         138 GCTGTGGAAACAGACATATGTATATGAAAGTAAAATCAAGAGTGCTTAGAAGCTGGATAA 197

Qy          56 ---------------GluGlyTyrPheThrAlaIleMetLeuLysLysSerArg------ 68
                           :::       :::|||  ::|||  |||||||||
Db         198 GTAAAGAGAATGAGAAAGCCAGCAATAAGTGCCAAATTACTTTTTAAATCTAGAAACCGA 257

Qy          69 ------------ValLysLeuLysSer------------------------------- 73
                        :::   ||||||
Db         258 TCATGTCACTTTCTCAGCTTAAAATGTTTCTGTGGTTCTCTATCACACTGAGAATAAATT 317

Qy          74 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 93
               |||       :::|||:::|||          |||   ||||||   ::::::
Db         318 CAAAACTCCTTACCCTGGCCCTGTGTGACTTGG------AACTGCCTTTGC---AATCTC 368

Qy          94 AsnValSerGlyAsnGluLeuCysLeuMet---------------------------Thr 104
                  :::       |||   |||  ::::::                          |||
Db         369 ACCTTACTATCTAACTCTCTTATGATGATACGGTCAATCAGATGGCCTTCTCGCCACACT 428

Qy         105 SerHisLeuGlu--------------------SerThrArgGlyHisAlaAlaGluArg 117
               ||||||                        |||:::   |||              |||
Db         429 TCTCACCCCAGTAGGGTTTGTGCACTTGCTTTGTCCTCTGCTGGGTAATGCTTTTCCAGA 488

Qy         118 Met------AsnGlnLeuLysMet----------------------------ValLeu 125
               |||      :::  |||||||||                            |||
Db         489 ATGTAAACAAGTCTGCTCAAAATGGAGGACAAAACAGATGCAGACTATGCCCTAAGTTTA 548

Qy         126 LysLysMetGlnGluAlaProGluSerAlaThrValIle------PheAlaGlyAspThr 143
               |||          :::      :::      ||||||              |||:::
Db         549 AAAACACAGAGCCAAACCTAAAAACAAAAGACAGTGTTCCCATGGCCACCAGGGGAATTT 608

Qy         144 AsnLeuArgAspArgGluValThrArg---CysGly---GlyLeuProAsn--------- 158
                  :::      :::::::|||   ::::::  ||||||  |||:::  |||
Db         609 AGCCCGAGTAACAAGGAGAAGTCGAAACCGTGTGGCCTGGGAATGCAAAATGGGACAACC 668

Qy         159 AsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
                  :::   |||   |||:::||||||   |||   ||||||
Db         669 ACTTTGGAAGACAGTTGGCAGTTTCTTACAAAACTAAAACAC 710

RESULT 42
US-10-750-623-54729
; Sequence 54729, Application US/10750623
```

```
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 54729
;    LENGTH: 1199
;    TYPE: DNA
;    ORGANISM: Bovine  19866881341043
US-10-750-623-54729

Alignment Scores:
Pred. No.:                36.7          Length:          1199
Score:                    81.50         Matches:         54
Percent Similarity:       36.3%         Conservative:    31
Best Local Similarity:    23.1%         Mismatches:      64
Query Match:              6.8%          Indels:          85
DB:                       8             Gaps:            14

US-10-757-745-2_COPY_121_347 (1-227) x US-10-750-623-54729 (1-1199)

Qy          21 AlaLeuTyrSerProAspValIle-----------------------PheLeuGlnGlu 32
               |||:::|||:::|||   :::|||                       |||||||||
Db          18 GCAATGTATAACCCAGGGTTAATCCAATGCTTTAACATGATTTCTTCTTTTTTGCAGACT 77

Qy          33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
                  :::    :::|||   |||     :::   |||   |||          :::
Db          78 AATGTTGGAAAGTTTTATTTGTATTTCGGAAGGCATTCTCAAAATGTAAATGCTTTGGAG 137

Qy          53 GlyHisGlu---------------------------------------------------- 55
                  |||
Db         138 GCTGTGGAAACAGACATATGTATATGAAAGTAAAATCAAGAGTGCTTAGAAGCTGGATAA 197

Qy          56 ---------------GluGlyTyrPheThrAlaIleMetLeuLysLysSerArg------ 68
                           :::       :::|||  ::|||  |||||||||
Db         198 GTAAAGAGAATGAGAAAGCCAGCAATAAGTGCCAAATTACTTTTTAAATCTAGAAACCGA 257

Qy          69 ------------ValLysLeuLysSer------------------------------- 73
                        :::   ||||||
Db         258 TCATGTCACTTTCTCAGCTTAAAATGTTTCTGTGGTTCTCTATCACACTGAGAATAAATT 317

Qy          74 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 93
               |||       :::|||:::|||          |||   ||||||   ::::::
Db         318 CAAAACTCCTTACCCTGGCCCTGTGTGACTTGG------AACTGCCTTTGC---AATCTC 368

Qy          94 AsnValSerGlyAsnGluLeuCysLeuMet---------------------------Thr 104
                  :::       |||   |||  ::::::                          |||
Db         369 ACCTTACTATCTAACTCTCTTATGATGATACGGTCAATCAGATGGCCTTCTCGCCACACT 428

Qy         105 SerHisLeuGlu--------------------SerThrArgGlyHisAlaAlaGluArg 117
               ||||||                        |||:::   |||              |||
Db         429 TCTCACCCCAGTAGGGTTTGTGCACTTGCTTTGTCCTCTGCTGGGTAATGCTTTTCCAGA 488

Qy         118 Met------AsnGlnLeuLysMet----------------------------ValLeu 125
               |||      :::  |||||||||                            |||
Db         489 ATGTAAACAAGTCTGCTCAAAATGGAGGACAAAACAGATGCAGACTATGCCCTAAGTTTA 548

Qy         126 LysLysMetGlnGluAlaProGluSerAlaThrValIle------PheAlaGlyAspThr 143
               |||          :::      :::      ||||||              |||:::
Db         549 AAAACACAGAGCCAAACCTAAAAACAAAAGACAGTGTTCCCATGGCCACCAGGGGAATTT 608
```

```
Qy        144 AsnLeuArgAspArgGluValThrArg---CysGly---GlyLeuProAsn--------- 158
Db        609 AGCCCGAGTAACAAGGAGAAGTCGAAACCGTGTGGCCTGGGAATGCAAAATGGGACAACC 668

Qy        159 AsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
Db        669 ACTTTGGAAGACAGTTGGCAGTTTCTTACAAAACTAAAACAC 710

RESULT 43
US-09-925-065A-702911/c
; Sequence 702911, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 702911
;   LENGTH: 1814
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-702911

Alignment Scores:
Pred. No.:              63.3          Length:          1814
Score:                  81.50         Matches:         45
Percent Similarity:     38.8%         Conservative:    38
Best Local Similarity:  21.0%         Mismatches:      85
Query Match:            6.8%          Indels:          46
DB:                     6             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-702911 (1-1814)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
Db       1607 GAAGCGGTTGACATCAGTAAACTGTCTCCTTCTTCTAGCTTCATCTGCTAATTCCAGTGC 1548

Qy         22 LeuTyrSerProAspVal--------------------------------IlePheLeu 30
Db       1547 TTGTACAAGAACAATATCATCATTAGAGGAGAAAATGGTCAGAGGAGGTGTATCTGGATC 1488

Qy         31 GlnGluValIle----------------------------------------ProPro 36
Db       1487 AGGGAAGTTACGCTGAAGTGGATCATAGTGGATGCCATCATAAATAAGCAGAACCCTTTT 1428

Qy         37 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 56
Db       1427 GGTATATCCTGCATCTTCCCCAAAACGATCAATTCTTACTGTSTGTGTATCCACTACACA 1368

Qy         57 GlyTyrPheThrAlaIleMetLeuLysLysSerArg---ValLysLeuLysSerGlnGlu 75
Db       1367 ---TATTTCACATTGGTAAAACTTGGACAAAATCGATATCTCTATTGCTCCTCCCCAAGT 1311

Qy         76 IleIleProPhe---ProSerThrLysMetMet------------ArgAsnLeuLeuCys 90
Db       1310 GTCATCCCTTTTGATCCAGTCACAGTACTCTTGATTTGTTTTCCCAGTATTGCCTCACT 1251

Qy         91 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 110
```

```
Db       1250 ATAGAAGTCTGGATCGCTTGCTACAATTTGTGCTATGAGGCGTCTCATCTCAGGGGCACA 1191

Qy         111 -ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGl 130
Db       1190 AGCTGGATTCAAGACTCCTCCTTCGACGACATAGTACACACTAGTAAAGAGGCAAGAGTT 1131

Qy         130 uAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluVa 150
Db       1130 GTCTGCTGGGACCACGGTTCTGGTAAGCACAGGCAAAGTTTCCCTGACGTAACTAGAAGC 1071

Qy         150 lThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPr 170
Db       1070 ACCACGT--------------------TTAGTAAATGCAGGTGAACTTCTGGGC----- 1037

Qy         170 oLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
Db       1036 ------------CTGGTTTGGTCTTCTTCAATGATCAGC 1010

RESULT 44
US-10-330-773-458/c
; Sequence 458, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 458
;   LENGTH: 194534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(194534)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-458

Alignment Scores:
Pred. No.:              2.98e+04      Length:          194534
Score:                  81.50         Matches:         44
Percent Similarity:     41.8%         Conservative:    37
Best Local Similarity:  22.7%         Mismatches:      80
Query Match:            6.8%          Indels:          33
DB:                     7             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-458 (1-194534)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
Db      25781 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG 25722

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
Db      25721 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 25680

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
Db      25679 ACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 25620

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
Db      25619 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 25560

Qy         82 ThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnValSerGly 97
Db      25559 ACAAAGATCAAAAGAGACAAAGAAGGCCATTATTTAATGGTAAAGGGATCAATTCAACAA 25500

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
```

```
Db      25499 GAAGAGCTACCTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 25449

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137

Db      25448 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGAGTTAGACTCCCACACATTA 25395

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145

Db      25394 ATAATG---GGAGACTTTAACACCCCACCGTCAACATTAGACAGATCAACGAGACAGAAA 25338

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163

Db      25337 GTCAACAAGGATACCCAGGAATTGAACTCAGCTTTGCACCAAGCAGACCTAATAGACATC 25278

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177

Db      25277 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 25239

RESULT 45
US-10-330-773-114/c
; Sequence 114, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 114
;   LENGTH: 440684
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(440684)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-114

Alignment Scores:
Pred. No.:                8.73e+04        Length:          440684
Score:                    81.50           Matches:         43
Percent Similarity:       42.3%           Conservative:    39
Best Local Similarity:    22.2%           Mismatches:      79
Query Match:              6.8%            Indels:          33
DB:                       7               Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-114 (1-440684)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21

Db     253549 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAC 253490

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41

Db     253489 GGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 253448

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61

Db     253447 ACGTGCAGAGACACAAATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 253388

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81

Db     253387 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTGTCTGATAAAACAGACTTTAAACCA 253328

Qy         82 ThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnValSerGly 97

Db     253327 ACAAAGGTCAAAAGAGAGAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 253268

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117

Db     253267 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 253217

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137

Db     253216 TTC------ATAAAGCAAGTCCTGAGGGACCTACAAAGAGACTTAGACTCCCACACAATA 253163

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145

Db     253162 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAAACAGAAA 253106

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163

Db     253105 GTTAACAAGGATATCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 253046

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177

Db     253045 TACAGAACTCTG---CACCCCAAATCAACAGAATATACATTT 253007

RESULT 46
US-11-096-568A-22794
; Sequence 22794, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 22794
;   LENGTH: 1466
;   TYPE: DNA
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1466)
;   OTHER INFORMATION: Ceres Seq. ID no. 12410153
US-11-096-568A-22794

Alignment Scores:
Pred. No.:                54.7            Length:          1466
Score:                    81.00           Matches:         48
Percent Similarity:       37.6%           Conservative:    29
Best Local Similarity:    23.4%           Mismatches:      61
Query Match:              6.8%            Indels:          67
DB:                       9               Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-11-096-568A-22794 (1-1466)

Qy          9 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 28

Db        496 ATGTCAGAAAAATTTCGCGGTTTG----------------------CCTGGAGTAGTT 531

Qy         29 PheLeuGlnGluValIleProProTyrTyrSerTyr-LeuLysLysArgSerSer----- 46

Db        532 TTC---------ATTTTGCCTGATTCATATCTATATCCAGAAACAAAGGAGTACGGAGGA 582

Qy         47 ----------------------------------AsnTyrGluIleIleThrGlyHi 54

Db        583 GACAAATATGACAATGGTGTCATCACTCCAAGACCACCACCTGTTCATTATAGCAGACCA 642

Qy         54 sGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGl 74

Db        643 TCAAGAACTGACAGGAACCGTAACTACCGAGGAAACTACCAGGATGGCCCTCCACAGCAA 702

Qy         74 nGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAs 94

Db        703 GGAAATT--------------------------ACCAGAACAACCGTCCTCCACCAGAA 735

Qy         94 nValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAl 114

Db        736 GGTG------------GTTACCAGAACAACCCGCCGCAGCAAGGAAACTACCAGACATAC 783
```

```
Qy        114 aAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSe 134
Db        784 CGCTCGC-------------------------AGCAAGATGGAAGAGGCTATGCCCCA 816

Qy        134 rAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGl 154
Db        817 CAGCAGAATTATG-------------------------CACAAGGTGGTCAGG----- 844

Qy        154 yGlyLeuProAsnAsnIleValAspValTrp---GluPheLeuGlyLysProLysHisCy 173
Db        845 ---------------ATGGTAGAGGTTTGGAAGGAATGATTACACAGACCGTTCAGGCT 888

Qy        173 sGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLe 193
Db        889 ACAA------TGGACCACCTGATTTTCGAAGTCAAACTCAGTACCAAGGGCATGTAAATC 942

Qy        193 uArgPheAspArg 197
Db        943 CAGCTGGGCAAGG 955

RESULT 47
US-10-645-441-11
; Sequence 11, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Ryba, Nicholas J.P.
;  APPLICANT:  Nelson, Greg
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Chandrashekar, Jayaram
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by the Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: Mammalian Sweet Taste Receptors
;  FILE REFERENCE: 02307E-120110US
;  CURRENT APPLICATION NUMBER: US/10/645,441
;  CURRENT FILING DATE:  2003-08-20
;  PRIOR APPLICATION NUMBER: US/09/927,315
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: US 60/302,898
;  PRIOR FILING DATE: 2001-07-03
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 2532
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   OTHER INFORMATION: mouse T1R2 sweet taste receptor
US-10-645-441-11

Alignment Scores:
Pred. No.:              112          Length:         2532
Score:                  81.00        Matches:        50
Percent Similarity:     32.4%        Conservative:   27
Best Local Similarity:  21.0%        Mismatches:     77
Query Match:            6.8%         Indels:         84
DB:                     8            Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-645-441-11 (1-2532)

Qy          5 AspLeuAsnAsnLeuSerGluArgAlaArgGly--------------------------- 15
Db        250 GAAATCAACAACTGTAGCTCTTTGCTGCCCGGCGTGCTGCTCGGCTACGAGATGGTGGAT 309

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal------ 33
Db        310 GTCTGCTACCTCTCCAACAATATCCAGCCTGGGCTCTACTTCCTGTCACAGATAGATGAC 369

Qy         34 ---IleProProTyrTyrSerTyrLeuLysLysArg----------------------- 44

Db        370 TTCCTGCCCATCCTCAAAGACTACAGCCAGTACAGGCCCCAAGTGGTGGCTGTTATTGGC 429

Qy         45 ------SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
Db        430 CCAGACAACTCTGAGTCTGCCATCACCGTGTCCAACATTCTCTCCTACTTCCTCGTGCCA 489

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
Db        490 CAGGTCACATATAGCGCCATCACCGACAAGCTGCAAGACAAGCGGCGCTTCCCTGCC--- 546

Qy         83 LysMetMetArgAsnLeu-----------------------------------LeuCys 90
Db        547 ---ATGCTGCGCACTGTGCCCAGCGCCACCCACCACATCGAGGCCATGGTGCAACTGATG 603

Qy         91 ValHis--------------------ValAsnValSerGlyAsnGluLeuCysLeuMet 103
Db        604 GTTCACTTCCAGTGGAACTGGATCGTGGTGCTGGTGAGCGATGACGATTATGGCCGAGAG 663

Qy        104 ThrSerHisLeu------------------------------------------------ 107
Db        664 AACAGCCACCTGCTGAGCCAGCGTCTGACCAACACTGGCGACATCTGCATTGCCTTCCAG 723

Qy        108 ------------------GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db        724 GAGGTTCTGCCCGTACCAGAACCCAACCAGGCTGTGAGGCCTGAGGAGCAGGACCAACTG 783

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db        784 GACAACATCCTGGACAAGCTGCGGCGGACTTCGGCGCGTGTGGTGGTGATATTCTCGCCG 843

Qy        142 AspThrAsnLeuArgAsp------ArgGluValThrArgCysGlyGlyLeuProAsnAsn 159
Db        844 GAGCTGAGCCTGCACAACTTCTTCCGTGAGGTGCTGCGC---------------------- 882

Qy        160 IleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
Db        883 ------------TGGAACTTCACGGGC-----------------TTTGTGTGG 906

RESULT 48
US-10-821-234-746
; Sequence 746, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
;  APPLICANT: Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Andarmani, Susan
;  APPLICANT:  Tang, Y. Tom
;  TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;  FILE REFERENCE: 821A
;  CURRENT APPLICATION NUMBER: US/10/821,234
;  CURRENT FILING DATE:  2004-04-07
;  PRIOR APPLICATION NUMBER: US 60/462,047
;  PRIOR FILING DATE: 2003-04-07
;  NUMBER OF SEQ ID NOS: 1704
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 746
;   LENGTH: 3828
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-821-234-746

Alignment Scores:
Pred. No.:              194          Length:         3828
Score:                  81.00        Matches:        44
Percent Similarity:     42.6%        Conservative:   37
Best Local Similarity:  23.2%        Mismatches:     77
Query Match:            6.8%         Indels:         32
DB:                     8            Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-821-234-746 (1-3828)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
```

```
Db         52 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCA 108

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45

Db        109 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 150

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65

Db        151 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 210

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85

Db        211 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 270

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------------GluLeuCys 101

Db        271 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 330

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121

Db        331 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 375

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141

Db        376 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 432

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147

Db        433 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 492

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167

Db        493 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 552

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177

Db        553 ---CACCCCAAATCAACAGAATATACCTTT 579
```

RESULT 49
US-11-121-086-76
; Sequence 76, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76
; LENGTH: 98862
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-76

Alignment Scores:

Pred. No.:	1.4e+04	Length:	98862
Score:	81.00	Matches:	37
Percent Similarity:	44.3%	Conservative:	25
Best Local Similarity:	26.4%	Mismatches:	52
Query Match:	6.8%	Indels:	26
DB:	12	Gaps:	7

US-10-757-745-2_COPY_121_347 (1-227) x US-11-121-086-76 (1-98862)

```
Qy         57 GlyTyrPheThrAlaIleMet---LeuLysLysSerArgValLysLeuLysSerGlnGlu 75

Db      96270 GGAATATTTACCAAGCAAATGGAAAACAAAAAAGCAGGGGTTGCAATCCTAGTCTCTGAT 96329

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95

Db      96330 AAAAGAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAAGCCATTACATAGTGGTA 96389

Qy         96 SerGlyAsn------------GluLeuCysLeuMetThrSerHisLeuGluSerThrArg 111

Db      96390 AAGGGATCAGTGCAACAAGAAGAGCTAACTATCCTAAATATACATGCACCCAATACAAGA 96449

Qy        112 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 131

Db      96450 ---------GCATCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGAC 96494

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu------------------ 145

Db      96495 TTAGACTCCCACACAATAATAATG---GGAGACTTTAACACACCACTGTCAACATTAGAC 96551

Qy        146 ------------------------ArgAspArgGluValThrArgCysGlyGlyLeuPro 157

Db      96552 AGATCAATGAGACAGAAAGTTAACAAGGATATCCAGGAATTGAACTCAGCTCTGCACCAA 96611

Qy        158 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177

Db      96612 GCAGACCTAATAGATATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACCTTC 96668
```

RESULT 50
US-11-121-086-76/c
; Sequence 76, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76
; LENGTH: 98862
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-76

Alignment Scores:

Pred. No.:	1.4e+04	Length:	98862
Score:	81.00	Matches:	37
Percent Similarity:	44.3%	Conservative:	25
Best Local Similarity:	26.4%	Mismatches:	52
Query Match:	6.8%	Indels:	26
DB:	12	Gaps:	7

US-10-757-745-2_COPY_121_347 (1-227) x US-11-121-086-76 (1-98862)

```
Qy         57 GlyTyrPheThrAlaIleMet---LeuLysLysSerArgValLysLeuLysSerGlnGlu 75

Db      92049 GGAATATTTACCAAGCAAATGGAAAACAAAAAAGCAGGGGTTGCAATCCTAGTCTCTGAT 91990

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95

Db      91989 AAAAGAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAAGCCATTACATAGTGGTA 91930

Qy         96 SerGlyAsn------------GluLeuCysLeuMetThrSerHisLeuGluSerThrArg 111

Db      91929 AAGGGATCAGTGCAACAAGAAGAGCTAACTATCCTAAATATACATGCACCCAATACAAGA 91870

Qy        112 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 131

Db      91869 ---------GCATCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGAC 91825

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu------------------ 145
```

```
                ::::|||   |||:::|||      ||||||   |||
Db      91824 TTAGACTCCCACACAATAATAATG---GGAGACTTTAACACACCACTGTCAACATTAGAC 91768

Qy        146 ------------------------ArgAspArgGluValThrArgCysGlyGlyLeuPro 157
                                     :::|||   :::
Db      91767 AGATCAATGAGACAGAAAGTTAACAAGGATATCCAGGAATTGAACTCAGCTCTGCACCAA 91708

Qy        158 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::::::::|||::::::      |||      ||||||      :::|||||:::
Db      91707 GCAGACCTAATAGATATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACCTTC 91651

RESULT 51
US-11-121-086-77
; Sequence 77, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77
;   LENGTH: 215308
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-77

Alignment Scores:
Pred. No.:              3.89e+04        Length:         215308
Score:                  81.00           Matches:        37
Percent Similarity:     44.3%           Conservative:   25
Best Local Similarity:  26.4%           Mismatches:     52
Query Match:            6.8%            Indels:         26
DB:                     12              Gaps:           7

US-10-757-745-2_COPY_121_347 (1-227) x US-11-121-086-77 (1-215308)

Qy         57 GlyTyrPheThrAlaIleMet---LeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              |||   ||||||      |||      ||||||:::   |||   :::         :::
Db      10299 GGAATATTTACCAAGCAAATGGAAAACAAAAAAGCAGGGGTTGCAATCCTAGTCTCTGAT 10358

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
                 |||      ||||||:::   |||:::             ::::::   |||
Db      10359 AAAAGAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAAGCCATTACATAGTGGTA 10418

Qy         96 SerGlyAsn-----------GluLeuCysLeuMetThrSerHisLeuGluSerThrArg 111
                 |||:::           ||||||  ::::::      |||      :::||||||
Db      10419 AAGGGATCAGTGCAACAAGAAGAGCTAACTATCCTAAATATACATGCACCCAATACAAGA 10478

Qy        112 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 131
                 |||  |||         :::|||   ||||||:::   :::|||
Db      10479 ---------GCATCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGAC 10523

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu----------------- 145
              :::|||   |||:::|||      ||||||   |||
Db      10524 TTAGACTCCCACACAATAATAATG---GGAGACTTTAACACACCACTGTCAACATTAGAC 10580

Qy        146 ------------------------ArgAspArgGluValThrArgCysGlyGlyLeuPro 157
                                     :::|||   :::
Db      10581 AGATCAATGAGACAGAAAGTTAACAAGGATATCCAGGAATTGAACTCAGCTCTGCACCAA 10640

Qy        158 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::::::::|||::::::      |||      ||||||      :::|||||:::
Db      10641 GCAGACCTAATAGATATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACCTTC 10697

RESULT 52
US-11-121-086-77/c
; Sequence 77, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77
;   LENGTH: 215308
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-77

Alignment Scores:
Pred. No.:              3.89e+04        Length:         215308
Score:                  81.00           Matches:        37
Percent Similarity:     44.3%           Conservative:   25
Best Local Similarity:  26.4%           Mismatches:     52
Query Match:            6.8%            Indels:         26
DB:                     12              Gaps:           7

US-10-757-745-2_COPY_121_347 (1-227) x US-11-121-086-77 (1-215308)

Qy         57 GlyTyrPheThrAlaIleMet---LeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              |||   ||||||      |||      ||||||:::   |||   :::         :::
Db       6078 GGAATATTTACCAAGCAAATGGAAAACAAAAAAGCAGGGGTTGCAATCCTAGTCTCTGAT 6019

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
                 |||      ||||||:::   |||:::             ::::::   |||
Db       6018 AAAAGAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAAAGCCATTACATAGTGGTA 5959

Qy         96 SerGlyAsn-----------GluLeuCysLeuMetThrSerHisLeuGluSerThrArg 111
                 |||:::           ||||||  ::::::      |||      :::||||||
Db       5958 AAGGGATCAGTGCAACAAGAAGAGCTAACTATCCTAAATATACATGCACCCAATACAAGA 5899

Qy        112 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 131
                 |||  |||         :::|||   ||||||:::   :::|||
Db       5898 ---------GCATCCAGATTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGAC 5854

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu----------------- 145
              :::|||   |||:::|||      ||||||   |||
Db       5853 TTAGACTCCCACACAATAATAATG---GGAGACTTTAACACACCACTGTCAACATTAGAC 5797

Qy        146 ------------------------ArgAspArgGluValThrArgCysGlyGlyLeuPro 157
                                     :::|||   :::
Db       5796 AGATCAATGAGACAGAAAGTTAACAAGGATATCCAGGAATTGAACTCAGCTCTGCACCAA 5737

Qy        158 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::::::::|||::::::      |||      ||||||      :::|||||:::
Db       5736 GCAGACCTAATAGATATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACCTTC 5680

RESULT 53
US-09-925-065A-871423/c
; Sequence 871423, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
```

```
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 871423
;   LENGTH: 593
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-871423

Alignment Scores:
Pred. No.:                19              Length:         593
Score:                    80.50           Matches:        42
Percent Similarity:       41.8%           Conservative:   39
Best Local Similarity:    21.6%           Mismatches:     80
Query Match:              6.8%            Indels:         33
DB:                       6               Gaps:           7

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-871423 (1-593)

Qy           4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
Db         586 TTAAATTTAAATGGGCTAAATGCCCCAATTAAAAGACAAAAACCAGCAAATTGGATAAAG 527

Qy          22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
Db         526 GGTCAAGACCCATCAGTGTGCTATATTCAGGAG-----------------ACCCATCTC 485

Qy          42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
Db         484 ACGTGCAGAGACACACATAGGCACAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 425

Qy          62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
Db         424 GGAAAGCAAAAAAAAGGAGGKGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 365

Qy          82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
Db         364 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 305

Qy          99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
Db         304 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA--------------GGAGAA 260

Qy         118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
Db         259 CCCAAATTCATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACAATA 200

Qy         138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
Db         199 ATAATG---GGAGACTTTAACACCCCACTGTCAATATTAGACAGATCAATGAGACAGAAG 143

Qy         146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
Db         142 GTTAACAAGGATATCCAGGACTTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATC 83

Qy         164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
Db          82 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 44

RESULT 54
US-09-925-065A-918150
; Sequence 918150, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 918150
;   LENGTH: 791
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-918150

Alignment Scores:
Pred. No.:                27.8            Length:         791
Score:                    80.50           Matches:        42
Percent Similarity:       41.2%           Conservative:   38
Best Local Similarity:    21.6%           Mismatches:     81
Query Match:              6.8%            Indels:         33
DB:                       6               Gaps:           7

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-918150 (1-791)

Qy           4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
Db         210 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAGG 269

Qy          22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
Db         270 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 311

Qy          42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
Db         312 AAGTGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 371

Qy          62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
Db         372 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTACTCTCTGATAGAACAGATTTTAAACCA 431

Qy          82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
Db         432 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 491

Qy          99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
Db         492 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA--------------GGAGAA 536

Qy         118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
Db         537 CCCAGATTCATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 596

Qy         138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
Db         597 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGATAGATCAATGAGACAGAAA 653

Qy         146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
Db         654 GTTAATAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGAACTAATAGACATC 713

Qy         164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
Db         714 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 752

RESULT 55
US-09-925-065A-944994
; Sequence 944994, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 944994
;   LENGTH: 791
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-944994

Alignment Scores:
Pred. No.:                27.8          Length:          791
Score:                    80.50         Matches:         42
Percent Similarity:       41.2%         Conservative:    38
Best Local Similarity:    21.6%         Mismatches:      81
Query Match:              6.8%          Indels:          33
DB:                       6             Gaps:            7

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-944994 (1-791)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db        210 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAGG 269

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||      :::||||||                        ::::::|||
Db        270 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 311

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
              |||   |||   ::::::          |||   |||       :::   |||
Db        312 AAGTGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 371

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                 ||||||:::   |||   :::         :::          |||
Db        372 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTACTCTCTGATAGAACAGATTTTAAACCA 431

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::         ::::::   |||   |||:::
Db        432 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 491

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||
Db        492 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA--------------GGAGAA 536

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                 :::|||   ||||||      :::|||         :::|||   |||:::
Db        537 CCCAGATTCATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 596

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db        597 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGATAGATCAATGAGACAGAAA 653

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                                 :::::|||:::
Db        654 GTTAATAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGAACTAATAGACATC 713

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
                 :::      |||      ||||||      :::|||||:::
Db        714 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 752
```

```
RESULT 56
US-11-072-512-1854
; Sequence 1854, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1854
;   LENGTH: 2580
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-072-512-1854

Alignment Scores:
Pred. No.:                132           Length:          2580
Score:                    80.50         Matches:         48
Percent Similarity:       35.5%         Conservative:    35
Best Local Similarity:    20.5%         Mismatches:      84
Query Match:              6.8%          Indels:          67
DB:                       9             Gaps:            11

US-10-757-745-2_COPY_121_347 (1-227) x US-11-072-512-1854 (1-2580)

Qy         30 LeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGlu 49
              |||         ||||||            ::::::                ::::|||:::
Db        725 CTCCATGTAGCAATCCCAGCCACCATGCAGCATATT--------------GATTACCAG 769

Qy         50 IleIleThrGlyHisGluGluGlyTyr---PheThrAlaIleMetLeuLysLysSerArg 68
              :::   |||          :::   |||   :::||||||||||:::   |||   :::
Db        770 CTTGAGACTGCCCGGAGCCAGGTATACCGCTACACAGCCATCCTTGAGAAGGCAGCCAAC 829

Qy         69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn--- 87
                 |||         |||:::   |||:::         :::   :::   |||
Db        830 GACCCACTCATGAGGGCTGAACTTCAGCCATATATGTTAAAAGAAGCGCTATATAATATA 889

Qy         88 ----------------------LeuLeuCysValHisValAsnValSerGlyAsnGlu 99
                                    ::::::         :::                |||
Db        890 CGACAATATGAGGTCTTCCTTCAGATCATTTTGTCAGATATAATTACTGGTGCTCAAGAA 949

Qy        100 LeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsn 119
              :::   :::||||||      |||                              :::
Db        950 GTGGAAATGATGACAAAACAGTTA--------------------GGAGCCCATCTGGAA 988

Qy        120 GlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal---Ile 138
              ||||||||||||   :::|||         |||   |||                |||
Db        989 CAACTAAAAATGACCATAAAATCAAAGATAGCGGTCCCAACATCACAAGTCTTTCCTATC 1048
```

```
Qy        139 PheAlaGly------------AspThrAsnLeuArgAspArgGluValThrArgCysGly 154
              |||                  |||:::|||:::|||      :::         |||
Db       1049 TTCATTGCACTTTCTACTCTGTGGACCAGCTTGCAAGACGAAACTATTGTG---GTTGGT 1105

Qy        155 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 174
                |||   |||          :::      |||||||||          |||
Db       1106 GTCCTCAGTAATTTATTCACTCACATTCAGCCATTCTTGGGT------GCTCACGAACTA 1159

Qy        175 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGly------------IleThrAlaAla 190
              |||    :::      |||      :::|||                |||
Db       1160 TACTTTCCTGAGAGAGTGATGCAATGTCATCTTAATGGAGCGACTGTGAAAACTGATGTG 1219

Qy        191 CysLysLeuArg---------------------------------------------- 194
              |||:::::::::
Db       1220 TGTAGAATGAAAGAACACATGGAAGATAGAGTAAATGTGGCAGATTTCAGAAAACTAGAA 1279

Qy        195 --------------------------PheAspArgIlePhe------------------ 199
                                         ||||||:::::
Db       1280 TGGCTTTTCCCAGAAACAACAGCAAATTTTGATAAACTGTTAATTCAATATCGGGGATTT 1339

Qy        200 ------PheArgAlaAlaAlaGluGluGlyHisIleIlePro 211
                  :::      ||||||   :::|||   ::::::|||
Db       1340 TGTGCTTACACGTTTGCTGCAACAGATGGTCTTCTCCTTCCA 1381

RESULT 57
US-10-330-773-911
; Sequence 911, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 911
;   LENGTH: 69648
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-911

Alignment Scores:
Pred. No.:              1.01e+04        Length:         69648
Score:                  80.50           Matches:        43
Percent Similarity:     42.3%           Conservative:   39
Best Local Similarity:  22.2%           Mismatches:     79
Query Match:            6.8%            Indels:         33
DB:                     7               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-330-773-911 (1-69648)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::         :::         :::   :::::::::
Db       3160 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 3219

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::||||||                    ::::::|||
Db       3220 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 3261

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::          |||   |||      :::   |||
Db       3262 ACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 3321

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                    ||||||:::   |||   :::         :::             |||
Db       3322 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 3381

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::  |||:::         ::::::   |||   |||:::
Db       3382 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 3441

Qy          99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                   ||||||   ::::::      :::      :::|||             |||   |||
Db       3442 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 3492

Qy         118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                       :::|||   ||||||      :::|||          :::|||   |||:::
Db       3493 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 3546

Qy         138 IlePheAlaGlyAspThrAsnLeu----------------------------------- 145
               |||      ||||||   |||
Db       3547 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 3603

Qy         146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                     :::|||   :::                          :::::::::|||:::
Db       3604 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 3663

Qy         164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
               :::      |||      ||||||      :::|||||:::
Db       3664 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 3702

RESULT 58
US-10-928-446A-1
; Sequence 1, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825234)
;   OTHER INFORMATION: the 'n' at position 825234 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825270)
;   OTHER INFORMATION: the 'n' at position 825270 may be 'c' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825401)
;   OTHER INFORMATION: the 'n' at position 825401 may be 'c' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825428)
;   OTHER INFORMATION: the 'n' at position 825428 may be 'g' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825473)
;   OTHER INFORMATION: the 'n' at position 825473 may be 'g' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825765)
;   OTHER INFORMATION: a "c" may be deleted at this position
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (825828)
;   OTHER INFORMATION: the 'n' at position 825828 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (826041)
;   OTHER INFORMATION: the 'n' at position 826041 may be 'a' or 'g'
;   FEATURE:
```

```
;   NAME/KEY: allele
;   LOCATION: (826546)
;   OTHER INFORMATION: the 'n' at position 826546 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (826654)
;   OTHER INFORMATION: the 'n' at position 826654 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (826826)
;   OTHER INFORMATION: the 'n' at position 826826 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (826863)
;   OTHER INFORMATION: the 'n' at position 826863 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (827008)
;   OTHER INFORMATION: the 'n' at position 827008 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (843055)
;   OTHER INFORMATION: the 'n' at position 843055 may be 'g' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (843118)
;   OTHER INFORMATION: the 'n' at position 843118 may be 't' or 'c'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (871027)
;   OTHER INFORMATION: the 'n' at position 871027 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (871140)
;   OTHER INFORMATION: the 'n' at position 871140 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (871168)
;   OTHER INFORMATION: the 'n' at position 871168 may be 'c' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (872678)
;   OTHER INFORMATION: the 'n' at position 872678 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (872742)
;   OTHER INFORMATION: the 'n' at position 872742 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (925859)
;   OTHER INFORMATION: the 'n' at position 925859 may be 'c' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (993220)
;   OTHER INFORMATION: the 'n' at position 993220 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (993254)
;   OTHER INFORMATION: the 'n' at position 993254 may be 'g' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1006462)
;   OTHER INFORMATION: the 'n' at position 1006462 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1007820)..(1007823)
;   OTHER INFORMATION: "ttct" may be deleted at this position
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1018038)
;   OTHER INFORMATION: the 'n' at position 1018038 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1018704)
;   OTHER INFORMATION: the 'n' at position 1018704 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1018718)..(1018720)
;   OTHER INFORMATION: "gtt" may be deleted at this position
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1026786)
;   OTHER INFORMATION: the 'n' at position 1026786 may be 'c' or 'a'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1047134)
;   OTHER INFORMATION: the 'n' at position 1047134 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1047159)
;   OTHER INFORMATION: the 'n' at position 1047159 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1047378)
;   OTHER INFORMATION: the 'n' at position 1047378 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1047739)
;   OTHER INFORMATION: the 'n' at position 1047739 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1050133)..(1050137)
;   OTHER INFORMATION: "ttaaa" may be deleted at this position
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1050539)
;   OTHER INFORMATION: the 'n' at position 1050539 may be 'c' or 't'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1062808)
;   OTHER INFORMATION: the 'n' at position 1062808 may be 'c' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1066392)
;   OTHER INFORMATION: the 'n' at position 1066392 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1073711)
;   OTHER INFORMATION: the 'n' at position 1073711 may be 'c' or 't'
US-10-928-446A-1

Alignment Scores:
Pred. No.:              3.68e+05        Length:         1080000
Score:                  80.50           Matches:        43
Percent Similarity:     42.3%           Conservative:   39
Best Local Similarity:  22.2%           Mismatches:     79
Query Match:            6.8%            Indels:         33
DB:                     8               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-1 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||      :::||||||                        ::::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||       :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||   :::         :::         |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426
```

```
Qy          82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
               ||||||:::   |||:::        ::::::   |||   |||:::
Db      399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy          99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                  ||||||   |||:::      :::      :::|||            |||   |||
Db      399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy         118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   :::|||      :::|||          :::|||   |||:::
Db      399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy         138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
               |||     |||||   |||
Db      399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy         146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                          ::::::::::|||:::
Db      399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy         164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
               :::     |||     ||||||     :::||||||:::
Db      399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 59
US-10-928-446A-181
; Sequence 181, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 181
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (826985)..(827008)
;   FEATURE:
;   OTHER INFORMATION: full exon 1 range is 826667-827008
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (827008)..(827008)
;   OTHER INFORMATION: the 'n' at position 827008 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (843242)..(843315)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (922549)..(922630)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (926021)..(926059)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (929123)..(929176)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (993104)..(993154)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999547)..(999608)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full exon 30 range is 1073289-1075279
US-10-928-446A-181

Alignment Scores:
Pred. No.:                3.68e+05        Length:          1080000
Score:                    80.50           Matches:         43
Percent Similarity:       42.3%           Conservative:    39
Best Local Similarity:    22.2%           Mismatches:      79
Query Match:              6.8%            Indels:          33
DB:                       8               Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-181 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::         :::       :::  :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::||||||                   ::::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::          |||   |||       :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||   :::         :::         |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::           ::::::   |||   |||:::
Db     399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   |||:::      :::      :::|||             |||   |||
Db     399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                :::|||   :::|||       :::|||         :::|||   |||:::
Db     399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db     399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                 :::|||   :::                         :::::::::||:::
Db     399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
```

```
                 :::       |||       ||||||       :::||||||:::
Db     399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 60
US-10-928-446A-183
; Sequence 183, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 183
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (826985)..(827017)
;   FEATURE:
;   OTHER INFORMATION: full exon 1 range is 826667-827008
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (843242)..(843315)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (922549)..(922630)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (926021)..(926059)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (929123)..(929176)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (993104)..(993154)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999547)..(999608)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
```

```
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full exon 30 range is 1073289-1075279
US-10-928-446A-183
```

```
Alignment Scores:
Pred. No.:                3.68e+05        Length:          1080000
Score:                    80.50           Matches:         43
Percent Similarity:       42.3%           Conservative:    39
Best Local Similarity:    22.2%           Mismatches:      79
Query Match:              6.8%            Indels:          33
DB:                       8               Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-183 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::      :::      :::   :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                 |||   |||      :::||||||                    :::::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::          |||   |||      :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                 ||||||:::  |||  :::      :::                  |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
                |||||:::   |||:::            ::::::   |||   |||:::
Db     399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                  ||||||   |||:::      :::      :::|||                |||   |||
Db     399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                 :::|||  :::|||     :::|||          :::||||   |||:::
Db     399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
                |||      ||||||   |||
Db     399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                    ::::::::::|||:::
Db     399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
                ::   |||    ||||||     :::|||||:::
Db     399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 61
US-10-928-446A-185
; Sequence 185, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 185
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (722487)..(722534)
;   FEATURE:
```

```
;   OTHER INFORMATION: full length exon 1 range is 722213-722534
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (843242)..(843315)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (922549)..(922630)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (926021)..(926059)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (929123)..(929176)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (993104)..(993154)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999547)..(999608)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-185

Alignment Scores:
Pred. No.:              3.68e+05        Length:         1080000
Score:                  80.50           Matches:        43
Percent Similarity:     42.3%           Conservative:   39
Best Local Similarity:  22.2%           Mismatches:     79
Query Match:            6.8%            Indels:         33
DB:                     8               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-185 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||      :::||||||                    ::::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::       |||   |||      :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                ||||||:::   |||   :::         :::         |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426
```

```
Qy          82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
               ||||||:::   |||:::          ::::::   |||   |||:::
Db      399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy          99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   |||:::     :::      :::|||             |||   |||
Db      399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy         118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                   :::|||   :::|||      :::|||          :::|||   |||:::
Db      399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy         138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
               |||      ||||||   |||
Db      399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy         146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                          ::::::::::|||:::
Db      399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy         164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
               :::      |||     ||||||     :::||||||:::
Db      399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 62
US-10-928-446A-187
; Sequence 187, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 187
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-187

Alignment Scores:
Pred. No.:              3.68e+05        Length:         1080000
Score:                  80.50           Matches:        43
```

```
Percent Similarity:     42.3%          Conservative:  39
Best Local Similarity:  22.2%          Mismatches:    79
Query Match:            6.8%           Indels:        33
DB:                     8              Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-187 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::   :::       :::   :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||      :::|||||||                  :::::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||      :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                |||||:::   |||   :::          :::          |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::          ::::::   |||   |||:::
Db     399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   |||:::      :::      :::|||            |||   |||
Db     399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                :::|||   :::|||      :::|||         :::|||   |||:::
Db     399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db     399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                  :::|||   :::                        :::::::::|||:::
Db     399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::|||||:::
Db     399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 63
US-10-928-446A-189
; Sequence 189, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 189
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (723508)..(723516)
;   FEATURE:
;   OTHER INFORMATION: full length exon 1 range is 723483-723516
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
```

```
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-189

Alignment Scores:
Pred. No.:                3.68e+05        Length:          1080000
Score:                    80.50           Matches:         43
Percent Similarity:       42.3%           Conservative:    39
Best Local Similarity:    22.2%           Mismatches:      79
Query Match:              6.8%            Indels:          33
DB:                       8               Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-189 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||      :::||||||                  ::::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||  ::::::          |||   |||     :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||   :::         :::         |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::             ::::::   |||   |||:::
Db     399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   |||:::     :::     :::|||         |||   |||
Db     399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   :::|||      :::|||         :::|||   |||:::
Db     399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db     399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                        :::::::::|||:::
Db     399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::||||||:::
Db     399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 64
US-10-928-446A-191
; Sequence 191, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 191
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-191

Alignment Scores:
Pred. No.:              3.68e+05      Length:          1080000
Score:                  80.50         Matches:         43
Percent Similarity:     42.3%         Conservative:    39
Best Local Similarity:  22.2%         Mismatches:      79
Query Match:            6.8%          Indels:          33
DB:                     8             Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-191 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::         :::       :::  :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                 |||  |||      :::||||||                  ::::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||      :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||  :::       :::       |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::           ::::::   |||   |||:::
Db     399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   |||:::      :::      :::|||           |||   |||
Db     399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   :::|||      :::|||         :::|||   |||:::
Db     399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
                |||       ||||||   |||
Db     399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                    :::|||   :::                          :::::::::|||:::
Db     399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::       |||       ||||||       :::|||||:::
Db     399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 65
US-10-928-446A-193
; Sequence 193, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 193
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   OTHER INFORMATION: exon
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   OTHER INFORMATION: exon
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-193

Alignment Scores:
Pred. No.:                3.68e+05        Length:          1080000
Score:                    80.50           Matches:         43
Percent Similarity:       42.3%           Conservative:    39
Best Local Similarity:    22.2%           Mismatches:      79
Query Match:              6.8%            Indels:          33
DB:                       8               Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-193 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::         :::         :::   :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::|||||||                    :::::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                   |||   ::::::          |||   |||   :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||   :::         :::          |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::               ::::::   |||   |||:::
Db     399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   |||:::       :::      :::|||              |||   |||
Db     399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   :::|||       :::|||          :::|||   |||:::
Db     399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy        138 IlePheAlaGlyAspThrAsnLeu----------------------------------- 145
              |||       ||||||  |||
Db     399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                          :::::::::|||:::
Db     399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::       |||       ||||||      :::||||||:::
Db     399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 66
US-10-928-446A-195
; Sequence 195, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 195
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-195

Alignment Scores:
Pred. No.:              3.68e+05        Length:         1080000
Score:                  80.50           Matches:        43
Percent Similarity:     42.3%           Conservative:   39
Best Local Similarity:  22.2%           Mismatches:     79
Query Match:            6.8%            Indels:         33

DB:                     8               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-195 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::||||||                     ::::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||       :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||   :::         :::         |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::           ::::::   |||   |||:::
Db     399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   |||:::       :::       :::|||           |||   |||
Db     399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   :::|||       :::|||           :::|||   |||:::
Db     399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
                |||       ||||||   |||
Db     399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                                 ::::::::|||:::
Db     399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::       |||       ||||||       :::||||||:::
Db     399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 67
US-10-928-446A-197
; Sequence 197, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 197
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   FEATURE:
;   NAME/KEY: CDS
```

```
;   LOCATION: (1002118)..(1002284)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-197

Alignment Scores:
Pred. No.:              3.68e+05        Length:         1080000
Score:                  80.50           Matches:        43
Percent Similarity:     42.3%           Conservative:   39
Best Local Similarity:  22.2%           Mismatches:     79

Query Match:            6.8%            Indels:         33
DB:                     8               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-197 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::||||||                  :::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                |||   ::::::         |||   |||       :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||   :::         :::         |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::               ::::::   |||   |||:::
Db     399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   |||:::       :::       :::|||           |||   |||
Db     399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   :::|||       :::|||         :::|||   |||:::
Db     399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||       ||||||   |||
Db     399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                 :::|||   :::                         :::::::::|||:::
Db     399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||     :::|||||:::
Db     399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 68
US-10-928-446A-199
; Sequence 199, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 199
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   FEATURE:
;   OTHER INFORMATION: full length exon 7 range is 999547-999608
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
;   FEATURE:
;   OTHER INFORMATION: full length exon 30 range is 1073289-1075279
US-10-928-446A-199

Alignment Scores:
Pred. No.:              3.68e+05        Length:         1080000
Score:                  80.50           Matches:        43
Percent Similarity:     42.3%           Conservative:   39
Best Local Similarity:  22.2%           Mismatches:     79
Query Match:            6.8%            Indels:         33
DB:                     8               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-199 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||      :::||||||                       :::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::          |||   |||     :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                   |||||:::   |||   :::         :::         |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::             ::::::   |||   |||:::
Db     399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   |||:::   :::     :::|||                 |||   |||
Db     399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                   :::|||   :::|||       :::|||         :::|||   |||:::
Db     399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||       ||||||   |||
Db     399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                 :::|||   :::                         ::::::::::|||:::
Db     399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::       |||       ||||||      :::||||||:::
Db     399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 69
US-10-928-446A-201
; Sequence 201, Application US/10928446A
; Publication No. US20050277123A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF UTAH RESEARCH FOUNDATION
;  TITLE OF INVENTION: VARIANTS OF NEDD4L ASSOCIATED WITH HYPERTENSION AND
;  TITLE OF INVENTION:  VIRAL BUDDING
;  FILE REFERENCE: 0274-5785.1US
;  CURRENT APPLICATION NUMBER: US/10/928,446A
;  CURRENT FILING DATE:  2004-08-26
;  PRIOR APPLICATION NUMBER: 60/359,741
;  PRIOR FILING DATE: 2002-02-26
;  NUMBER OF SEQ ID NOS: 202
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 201
;   LENGTH: 1080000
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (827008)..(827008)
;   OTHER INFORMATION: the 'n' at position 827008 may be 'a' or 'g'
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (872617)..(872860)
;   OTHER INFORMATION: Exon 2a
```

```
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (904952)..(905030)
;   OTHER INFORMATION: Exon 2d
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (922549)..(922630)
;   OTHER INFORMATION: Exon 3
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (926021)..(926059)
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (929123)..(929176)
;   OTHER INFORMATION: Exon 5
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (993104)..(993154)
;   OTHER INFORMATION: Exon 6
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (999562)..(999608)
;   OTHER INFORMATION: Full exon range is 999547-999608
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1000354)..(1000456)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1000354)..(1000456)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1002118)..(1002284)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1002118)..(1002284)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1006117)..(1006249)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1006117)..(1006249)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1007860)..(1008036)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1007860)..(1008036)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1010940)..(1011014)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1010940)..(1011014)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1018160)..(1018291)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018160)..(1018291)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1018800)..(1018919)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1018800)..(1018919)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1020028)..(1020225)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1020028)..(1020225)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1026659)..(1026736)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1026659)..(1026736)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1028113)..(1028167)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1028113)..(1028167)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1034316)..(1034374)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1034316)..(1034374)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1041390)..(1041455)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1041390)..(1041455)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1043121)..(1043350)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1043121)..(1043350)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1044868)..(1044989)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1044868)..(1044989)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1047519)..(1047589)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1047519)..(1047589)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1050296)..(1050391)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1050296)..(1050391)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1060368)..(1060441)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1060368)..(1060441)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1062648)..(1062708)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1062648)..(1062708)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1064561)..(1064620)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1064561)..(1064620)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1066207)..(1066314)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1066207)..(1066314)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1067768)..(1067864)
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1067768)..(1067864)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1068609)..(1068681)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1068609)..(1068681)
;   FEATURE:
;   NAME/KEY: exon
;   LOCATION: (1073289)..(1075279)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1073289)..(1073388)
US-10-928-446A-201

Alignment Scores:
Pred. No.:                3.68e+05        Length:          1080000
Score:                    80.50           Matches:         43
Percent Similarity:       42.3%           Conservative:    39
Best Local Similarity:    22.2%           Mismatches:      79
Query Match:              6.8%            Indels:          33
DB:                       8               Gaps:            8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-928-446A-201 (1-1080000)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::     :::         :::   :::::::::
Db     399205 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 399264

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                   |||   |||      :::||||||                   ::::::|||
Db     399265 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 399306

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                  |||   ::::::         |||   |||       :::   |||
Db     399307 ACGTGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 399366

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProProSerSer 81
                  ||||||:::   |||  :::      :::         |||
Db     399367 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 399426

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::           ::::::   |||   |||:::
Db     399427 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCGACAA 399486

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   |||:::      :::      :::|||          |||   |||
Db     399487 GAAGAGCTAACTCTCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 399537

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                  :::|||   :::|||      :::|||         :::|||   |||:::
Db     399538 TTC------ATAAAGCAAATCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 399591

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db     399592 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 399648

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::                 :::::::::|||:::
Db     399649 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 399708

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::|||||:::
Db     399709 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 399747

RESULT 70
US-09-925-065A-861077
; Sequence 861077, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 861077
;   LENGTH: 664
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-861077

Alignment Scores:
Pred. No.:                25.3            Length:          664
Score:                    80.00           Matches:         44
Percent Similarity:       46.2%           Conservative:    36
Best Local Similarity:    25.4%           Mismatches:      59
Query Match:              6.7%            Indels:          34
DB:                       6               Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-861077 (1-664)

Qy         25 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 44
              |||   |||   ::::::||||||           :::   :::
Db        100 CCATCGGTGTGCTATATTCAGGAGACC---------CATCTCACATGCAAAGACACAATA 150

Qy         45 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet--- 63
                 |||   |||                  :::||||||   ||||||       |||
Db        151 GGCTCAAAATAT--------------AGGGATGAAGGAATATTTACCAAGCAAATGGAA 195

Qy         64 ---LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
                 ||||||:::   |||   :::       :::         |||   ||||||
Db        196 AGCAAAAAAAAAGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAATCAACA 255

Qy         83 LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------------ 98
              |||:::   ::::::           :::       |||   |||:::
Db        256 AAGATCAAAAAAGACAAAGAAGGGCATTACACAATGGTAAAGGGATCAATGCAACAAGAA 315

Qy         99 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 118
              ||||||   ::::::      ::::::   :::|||          |||   |||
Db        316 GAGCTAACTATCCTAAATATATATATACCAAATACA---------GGAGCACCCAGATTC 366

Qy        119 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 138
                 :::|||   ||||||:::   :::|||          :::|||   |||   :::
Db        367 ------ATAAAGCAAGTTCTTAGAGACCTACAAAGAGACTTAGACTCCCACACARTATTG 420

Qy        139 PheAlaGly--------------------AspThrAsnLeuArgAspArgGluValThr 151
              :::                          |||:::   :::||||||::::::
Db        421 TGGGAGCCTTTAGTACCCCACTGTCAATATTAGACAGATCAACAAGACAGAAAATTAGCA 480

Qy        152 ArgCys-GlyGlyLeu---------------ProAsn-AsnIleValAspValTrpGluP 166
              |||      ||||||                |||::: ::::::::::|||::::::
Db        481 AGGATATTCAGGACTTGAACTCAGCTCTGGACCAAGTGGACCTAATAGACATATACAGAA 540

Qy        166 heLeuGlyLysProLysHisCysGlnTyrThrTrp 177
                |||      ||||||      :::|||||:::
Db        541 CTCTC---CACCCCAAATCAACAGAATATACATTC 572

RESULT 71
US-11-177-987-9
; Sequence 9, Application US/11177987
; Publication No. US20050271619A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Dumoutier, Laure
;  APPLICANT:  Renauld, Jean-Christophe
;  TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac
;  TITLE OF INVENTION:  Interleukin-21, The Proteins Encoded, and Uses Thereof
;  FILE REFERENCE: LUD 5664
;  CURRENT APPLICATION NUMBER: US/11/177,987
;  CURRENT FILING DATE:  2005-07-07
;  PRIOR APPLICATION NUMBER: US/09/626,617
;  PRIOR FILING DATE: 2000-07-27
;  PRIOR APPLICATION NUMBER: US09/419,568
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: US09/354,243
;  PRIOR FILING DATE: 1999-07-16
;  PRIOR APPLICATION NUMBER: US09/178,973
;  PRIOR FILING DATE: 1998-10-26
;  NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 9
;   LENGTH: 1111
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
US-11-177-987-9

Alignment Scores:
Pred. No.:              49.8          Length:         1111
Score:                  80.00         Matches:        36
Percent Similarity:     40.8%         Conservative:   24
Best Local Similarity:  24.5%         Mismatches:     43
Query Match:            6.7%          Indels:         44
DB:                     12            Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-11-177-987-9 (1-1111)

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 59
Db        317 TACCTGATGAAGCAGGTGCTCAACTTCACCCTG-------------------------- 349

Qy         60 ThrAlaIleMetLeuLysLysSer---ArgValLysLeuLysSerGlnGluIleIlePro 78
Db        350 GAAGACATTCTGCTCCCCCAGTCAGACAGGTTCCGGCCCTACATGCAGGAGGTGGTGCCT 409

Qy         79 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 98
Db        410 TTC---CTGACCAAACTCAGCAATCAGCTCAGCTCCTGTCAC------ATCAGTGGTGAC 460

Qy         99 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 118
Db        461 GAC----------------------------------CAGAACATCCAGAAGAATGTC 484

Qy        119 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 138
Db        485 AGAAGGCTGAAGGAGACAGTGAAAAAGCTTGGAGAGAGCGGAGAGATCAAAGCGATC--- 541

Qy        139 PheAlaGlyAspThrAsnLeu------------------------------------- 145
Db        542 ------GGGGAACTGGACCTGCTGTTTATGTCTCTGAGAAATGCTTGCGTCTGAGCGAGA 595

Qy        146 ArgAspArgGluValThrArgCysGlyGlyLeuPro---------AsnAsnIleValAsp 162
Db        596 AGAAGCTAGAAAACGAAGAACTGCTCCTTCCTGCCTTCTAAAAAGAACAATAAGATCCCT 655

Qy        163 ValTrpGluPheLeuGlyLys 169
Db        656 GAATGGACTTTTTTACTAAAG 676

RESULT 72
US-11-121-086-94
; Sequence 94, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 94
;   LENGTH: 100387
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-94

Alignment Scores:
Pred. No.:              2.14e+04      Length:         100387
Score:                  79.50         Matches:        44
Percent Similarity:     37.6%         Conservative:   33
Best Local Similarity:  21.5%         Mismatches:     69
Query Match:            6.7%          Indels:         59
DB:                     12            Gaps:           12

US-10-757-745-2_COPY_121_347 (1-227) x US-11-121-086-94 (1-100387)

Qy         36 ProTyrTyr------SerTyrLeuLysLysArgSerSerAsn-----------TyrGlu 49
Db      90046 CCCTGTTATGCCATGTCTTTTAGGGACAAAAGGGCTCAGAATTTATGTCTCCTCTACTCT 90105

Qy         50 IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 69
Db      90106 GTTCTTAGAAGCCATCAT---------------------------------------- 90123

Qy         70 LysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu 89
Db      90124 ---CTAAGGGGGCCTTGATTGCTCCCATTTCTATCCAGCTCCCAAATTCTC---TTGCTC 90177

Qy         90 CysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSer 109
Db      90178 TGTAGCTCTGTGTCTCCTGCAAGCACGCCAGTCTGGAACATGACAACTATTTTGAATGTA 90237

Qy        110 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 129
Db      90238 ACTAAAGGT---ATCACTGAACCAGTTCATTATTTACAAATTATCTTTCCTTGCATGGAA 90294

Qy        130 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 149
Db      90295 TATTCAGAT--------------------GCTGGAGTCATTGGGAACAGAGATGGCCAA 90333

Qy        150 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLys 169
Db      90334 GCAGCAACTCTGGGT------TGTGAAACACTTACAGCCCCGCTATTCTTCTGTGGCAAC 90387

Qy        170 ProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 189
Db      90388 GAGAGACCCTGTAAA--------------------AAGGACAATTTTGCTTTGGAGACT 90426

Qy        190 ---------------------------AlaCysLysLeuArgPheAspArgIlePhePhe 200
Db      90427 AAGCAGCTCCCCACCAGTCAGGAATGACTCTGC------------------TTTTACTTT 90468

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
Db      90469 AGGCTGAGAGGTCAGAAGGGTCAC------------ACCCTCAATATTCTC---CTCTCC 90513

Qy        221 LysLeuAspCysGly 225
Db      90514 AAGACGGAGTGTGGC 90528

RESULT 73
US-11-121-086-51
; Sequence 51, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51
;   LENGTH: 176760
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-51

Alignment Scores:
Pred. No.:              4.5e+04       Length:          176760
Score:                  79.50         Matches:         48
Percent Similarity:     42.0%         Conservative:    44
Best Local Similarity:  21.9%         Mismatches:      84
Query Match:            6.7%          Indels:          43
DB:                     12            Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-11-121-086-51 (1-176760)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::|||   |||:::         :::         :::   :::::::::
Db      81361 TTAAATGTAAATGGACTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 81420

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||       :::||||||                 ::::::|||
Db      81421 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 81462

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::         |||   |||       :::   |||
Db      81463 ACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGGGGAAGATCTACCAAGCAAAT 81522

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                   ||||||:::   |||   :::         :::         |||
Db      81523 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 81582

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::         ::::::   |||   |||:::
Db      81583 ACAAAGATCAAAAGAGACAAAGAAAGCCATTACATAATGGTAAAGGGATCAATTCAACAA 81642

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                |||||||   ::::::      :::      :::|||             |||   |||
Db      81643 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 81693

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                 :::|||   |||||||       :::|||         :::|||   |||:::
Db      81694 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTA 81747

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||       ||||||   |||
Db      81748 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAA 81804

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                  :::|||   :::                         :::::::::|||:::
Db      81805 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 81864

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp------------------ 177
              :::       |||       ||||||      :::||||||:::
Db      81865 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTTTTTTCAGCACCACACCAC 81921

Qy        178 ---AspThrGlnMetAsnSerAsnLeuGly---------IleThrAlaAlaCysLys 192
                :::::::|||:::       ||||||         :::   :::   ||||||
Db      81922 ACCTGTTCCAAAATGGACCACATACTTGGAAGTAAAGCTCTCCTCAGCAAATGTAAA 81978

RESULT 74
US-11-121-086-83/c
; Sequence 83, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 83
;   LENGTH: 187745
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-83

Alignment Scores:
Pred. No.:              4.87e+04      Length:          187745
Score:                  79.50         Matches:         37
Percent Similarity:     37.4%         Conservative:    24
Best Local Similarity:  22.7%         Mismatches:      71
Query Match:            6.7%          Indels:          31
DB:                     12            Gaps:            6

US-10-757-745-2_COPY_121_347 (1-227) x US-11-121-086-83 (1-187745)

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
                |||   :::|||||||||||||||:::   |||       :::|||       |||
Db      33426 CCTTGTTTCTCCTATTTAAAAAAAAAAAAATCTGGGCCTAAGATT------GGATCATTA 33373

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
                 |||       |||:::   |||       |||:::           |||   |||
Db      33372 TGGGGGAAAGGAAGAGCATTAGAATTAACAAATAGCAAGGAATCATGGAAGCGACAAAAT 33313

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeu---------LeuCysValHis 92
                  |||:::         :::   :::       |||         |||   |||
Db      33312 ---ACTCCATGGAATAAGGAGCAGGCTATAAACTATTTGCAACGTAAACTTCTGGTTAGT 33256

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
                 |||                 ||||||   :::   |||
Db      33255 AGAAACTTTCAGCGTGGTGGCTCATGCCTATAATCCCAGCAC----------------- 33214

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
                                :::|||:::       |||             |||:::|||
Db      33213 ---------------TTTGGGAGGCTGAGGCGACGGCTCACCTCTGAAGTCGAGAGTCCG 33169

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
                          :::   :::             :::   ::::::   ::::::
Db      33168 AGACCAGTCTGACTAACGTGGCGAAAACCTGTCTCTACTAAAAATACAAAATTAGCCGAG 33109

Qy        153 CysGlyGlyLeu-----------------------------------ProAsnAsnIle 160
              |||||||||                                     |||   :::
Db      33108 TGTGGTGGCGCAGGCCTGTAATCCCAGCTATTCGGGAGGCTGAGGCAGGAGAATCGCTTG 33049

Qy        161 ValAspValTrpGluPheLeuGlyLysPro------LysHisCysGlnTyrThrTrpAsp 178
                                ||||:::|||       :::||||||         |||
Db      33048 AACCCAGGAGGTAGAGGTTGCGGTGAGCCGAGATCGCGCCATTGCACTCCAGCCTGGGCG 32989

Qy        179 ThrGlnMet 181
              |||::::::
Db      32988 ACAAGAGTG 32980

RESULT 75
US-09-925-065A-81444/c
; Sequence 81444, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
```

```
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 81444
;   LENGTH: 1267
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-81444

Alignment Scores:
Pred. No.:                77.5          Length:          1267
Score:                    79.00         Matches:         19
Percent Similarity:       56.9%         Conservative:    14
Best Local Similarity:    32.8%         Mismatches:      23
Query Match:              6.6%          Indels:          2
DB:                       6             Gaps:            1

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-81444 (1-1267)

Qy         51 IleThrGlyHisGluGluGly------TyrPheThrAlaIleMetLeuLysLysSerArg 68
              :::   |||::::::   |||       |||       ||||||:::         :::
Db        321 GTTATGGGATACCAATATGGTGCATGATATTCTCTGGCAATTCTGGGCACTGGCAATGAA 262

Qy         69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 88
              ::::::         :::   ||||||        ||||||   ::::::    |||::: 
Db        261 CTTCAGCCCCCAGTCAGGCACATAATCATGAGTAGCAGTACACTGATACTCTACAATATA 202

Qy         89 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 106
              ||||||      |||::::::   |||       ||||||:::   |||||||||
Db        201 CTGTGTTACCAGGTTAGTATTTTTGGCATTGTGCTTTGTGTTTTCACATCCCAT 148


Search completed: March 11, 2006, 13:49:50
Job time : 4697.94 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 00:16:23 ; Search time 1181.57 Seconds
(without alignments)
1588.688 Million cell updates/sec

Title: US-10-757-745-2_COPY_121_347
Perfect score: 1192
Sequence: 1 IDGLDLNNLSERARGVCSYL.....HIIPRSLDLLGLEKLDCGRF 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170337_4279/app_query.fasta_1
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss08
-USER=US10757745_@CGN_1_1_2715_@runat_09032006_170337_4279 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1192	100.0	1920	9	US-10-757-745-1	Sequence 1, Appli
2	1192	100.0	1940	9	US-10-783-271-28	Sequence 28, Appl
3	1192	100.0	1948	5	US-10-037-270-889	Sequence 889, App
4	1192	100.0	1948	6	US-10-117-722-889	Sequence 889, App
5	1192	100.0	1948	9	US-10-122-851-889	Sequence 889, App
6	1192	100.0	3152	3	US-09-981-353-88	Sequence 88, Appl
7	1184	99.3	1296	3	US-09-925-299-170	Sequence 170, App
8	1184	99.3	1296	3	US-09-925-299-170	Sequence 170, App
9	1116	93.6	1088	9	US-10-450-763-22058	Sequence 22058, A
10	1102	92.4	674	5	US-10-066-543-1067	Sequence 1067, Ap
c 11	1090	91.4	1079	3	US-09-745-288-63	Sequence 63, Appl
c 12	1090	91.4	1079	7	US-10-453-919-63	Sequence 63, Appl
13	1069	89.7	644	5	US-10-066-543-1070	Sequence 1070, Ap
14	1054	88.4	633	5	US-10-066-543-1084	Sequence 1084, Ap
15	1031	86.5	625	5	US-10-066-543-548	Sequence 548, App
16	1020	85.6	625	5	US-10-066-543-1031	Sequence 1031, Ap
17	945	79.3	1312	9	US-10-757-745-3	Sequence 3, Appli
18	902	75.7	553	5	US-10-066-543-983	Sequence 983, App
19	651	54.6	401	5	US-10-066-543-1363	Sequence 1363, Ap
20	590	49.5	400	7	US-10-242-535A-8082	Sequence 8082, Ap
21	590	49.5	400	7	US-10-085-783A-8082	Sequence 8082, Ap
22	465	39.0	1227	9	US-10-450-763-22059	Sequence 22059, A
23	298	25.0	176	3	US-09-878-722-121	Sequence 121, App
c 24	298	25.0	176	3	US-09-878-178-50	Sequence 50, Appl
25	298	25.0	176	3	US-09-904-456-121	Sequence 121, App
c 26	298	25.0	176	5	US-10-046-935-50	Sequence 50, Appl
c 27	298	25.0	176	5	US-10-146-502-50	Sequence 50, Appl
c 28	298	25.0	176	5	US-10-060-036-1367	Sequence 1367, Ap
29	191.5	16.1	1686	7	US-10-424-599-132360	Sequence 132360,
30	182.5	15.3	444	3	US-09-864-761-16762	Sequence 16762, A
31	182.5	15.3	691	3	US-09-864-761-33229	Sequence 33229, A
32	164.5	13.8	1497	7	US-10-437-963-42449	Sequence 42449, A
33	158.5	13.3	525	7	US-10-021-323-14543	Sequence 14543, A
c 34	134	11.2	268	3	US-09-876-889-312	Sequence 312, App
35	121.5	10.2	898	8	US-10-425-115-105489	Sequence 105489,
36	119	10.0	1236	7	US-10-437-963-47256	Sequence 47256, A
37	114.5	9.6	634	4	US-09-925-065A-783536	Sequence 783536,
38	111	9.3	606	4	US-09-925-065A-783537	Sequence 783537,
c 39	108	9.1	932	8	US-10-425-115-66072	Sequence 66072, A
40	108	9.1	168050	9	US-10-939-321-1	Sequence 1, Appli
c 41	101.5	8.5	1877	6	US-10-264-237-1230	Sequence 1230, Ap
42	98.5	8.3	1644	6	US-10-369-493-24191	Sequence 24191, A
43	97.5	8.2	9737	9	US-10-764-420-1725	Sequence 1725, Ap
c 44	97.5	8.2	90043	5	US-10-087-192-1141	Sequence 1141, Ap
45	97	8.1	1692	7	US-10-282-122A-22362	Sequence 22362, A
46	97	8.1	1830121	7	US-10-329-670-1	Sequence 1, Appli
47	97	8.1	1830121	8	US-10-158-865-1	Sequence 1, Appli
48	97	8.1	1830121	9	US-10-981-687-1	Sequence 1, Appli
49	96.5	8.1	800	8	US-10-425-115-105488	Sequence 105488,
c 50	95	8.0	5386	9	US-10-795-159-468	Sequence 468, App
c 51	95	8.0	908766	9	US-10-795-159-685	Sequence 685, App
52	94.5	7.9	1650	7	US-10-282-122A-30809	Sequence 30809, A
53	94.5	7.9	1767	9	US-10-998-197-33	Sequence 33, Appl
54	94.5	7.9	1971	9	US-10-887-553A-839	Sequence 839, App
55	94	7.9	3285	3	US-09-954-456-541	Sequence 541, App
56	94	7.9	3285	3	US-09-880-107-3745	Sequence 3745, Ap
57	94	7.9	3285	7	US-10-283-975A-269	Sequence 269, App
58	94	7.9	3285	9	US-10-843-641A-3568	Sequence 3568, Ap
59	94	7.9	3397	3	US-09-880-107-2403	Sequence 2403, Ap
60	94	7.9	3397	3	US-09-873-367C-166	Sequence 166, App
61	94	7.9	3397	7	US-10-283-975A-219	Sequence 219, App
62	94	7.9	3397	9	US-10-843-641A-166	Sequence 166, App
63	94	7.9	3423	6	US-10-172-118-832	Sequence 832, App
64	94	7.9	3423	7	US-10-342-887-832	Sequence 832, App
65	94	7.9	3423	9	US-10-887-553A-436	Sequence 436, App
66	94	7.9	3738	3	US-09-767-041-43	Sequence 43, Appl
67	94	7.9	5299	7	US-10-473-574-33	Sequence 33, Appl
68	94	7.9	5322	7	US-10-473-574-31	Sequence 31, Appl
69	92	7.7	1644	6	US-10-369-493-44898	Sequence 44898, A
c 70	92	7.7	42348	7	US-10-715-066-2	Sequence 2, Appli
71	91	7.6	1077	6	US-10-369-493-24652	Sequence 24652, A
72	90	7.6	108359	5	US-10-191-807-3	Sequence 3, Appli
73	90	7.6	108359	9	US-10-959-243-3	Sequence 3, Appli
c 74	89.5	7.5	3309400	3	US-09-738-626-1	Sequence 1, Appli
75	89	7.5	738	7	US-10-389-647-119	Sequence 119, App

ALIGNMENTS

RESULT 1

```
US-10-757-745-1
; Sequence 1, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1627)..(1627)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1108)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1849)..(1849)
;   OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

Alignment Scores:
Pred. No.:              6.05e-153       Length:         1920
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-757-745-1 (1-1920)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTT 1060

RESULT 2
US-10-783-271-28
; Sequence 28, Application US/10783271
; Publication No. US20050186577A1
; GENERAL INFORMATION:
;  APPLICANT: Veridex, LLC
;  APPLICANT:  Wang, Yixin
;  TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
;  FILE REFERENCE: VDX-5003 USNP
;  CURRENT APPLICATION NUMBER: US/10/783,271
;  CURRENT FILING DATE:  2004-02-20
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: human
US-10-783-271-28

Alignment Scores:
Pred. No.:              6.15e-153       Length:         1940
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-783-271-28 (1-1940)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 445

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTT 685

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 745
```

```
Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 805

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 865

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 925

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 985

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1045

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1046 AAACTGGACTGTGGTAGATTT 1066
```

RESULT 3
US-10-037-270-889
; Sequence 889, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-037-270-889

Alignment Scores:

Pred. No.:	6.19e-153	Length:	1948
Score:	1192.00	Matches:	227
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-037-270-889 (1-1948)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079
```

RESULT 4
US-10-117-722-889
; Sequence 889, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104

```
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-117-722-889

Alignment Scores:
Pred. No.:              6.19e-153       Length:         1948
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-117-722-889 (1-1948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079

RESULT 5
US-10-122-851-889
; Sequence 889, Application US/10122851
; Publication No. US20050239060A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BDV3
;  CURRENT APPLICATION NUMBER: US/10/122,851
;  CURRENT FILING DATE:  2002-04-12
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-122-851-889

Alignment Scores:
Pred. No.:              6.19e-153       Length:         1948
Score:                  1192.00         Matches:        227
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     9               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-122-851-889 (1-1948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938
```

```
Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1059 AAACTGGACTGTGGTAGATTT 1079

RESULT 6
US-09-981-353-88
; Sequence 88, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE: PA-0038 US
;  CURRENT APPLICATION NUMBER: US/09/981,353
;  CURRENT FILING DATE:  2001-10-11
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020160382A1 232992.1
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

Alignment Scores:
Pred. No.:              1.32e-152      Length:         3152
Score:                  1192.00        Matches:        227
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-981-353-88 (1-3152)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 1671

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1672 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 1731

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1732 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 1791

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1792 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 1851

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1852 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 1911

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1912 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 1971

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1972 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 2031

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2032 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 2091

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2092 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 2151

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2152 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 2211

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2212 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 2271

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       2272 AAACTGGACTGTGGTAGATTT 2292

RESULT 7
US-09-925-299-170
; Sequence 170, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:              4.14e-152      Length:         1296
Score:                  1184.00        Matches:        226
Percent Similarity:     99.6%          Conservative:   0
Best Local Similarity:  99.6%          Mismatches:     1
Query Match:            99.3%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-299-170 (1-1296)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 443

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 503

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACA 563
```

```
Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         564 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 623

Qy          81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         624 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTT 683

Qy         101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         684 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 743

Qy         121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         744 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 803

Qy         141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         804 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 863

Qy         161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         864 GTGGATGTCTGGGAGTTTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAA 923

Qy         181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         924 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 983

Qy         201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         984 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1043

Qy         221 LysLeuAspCysGlyArgPhe 227
               |||||||||||||||||||||
Db        1044 AAACTGGACTGTGGTAGATTT 1064

RESULT 8
US-09-925-299-170
; Sequence 170, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:              4.14e-152       Length:         1296
Score:                  1184.00         Matches:        226
Percent Similarity:     99.6%           Conservative:   0
Best Local Similarity:  99.6%           Mismatches:     1
Query Match:            99.3%           Indels:         0
DB:                     3               Gaps:           0
```

```
US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-299-170 (1-1296)

Qy           1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         384 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 443

Qy          21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         444 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 503

Qy          41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         504 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACA 563

Qy          61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         564 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 623

Qy          81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         624 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTT 683

Qy         101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         684 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 743

Qy         121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         744 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 803

Qy         141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         804 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 863

Qy         161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         864 GTGGATGTCTGGGAGTTTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAA 923

Qy         181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         924 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 983

Qy         201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         984 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1043

Qy         221 LysLeuAspCysGlyArgPhe 227
               |||||||||||||||||||||
Db        1044 AAACTGGACTGTGGTAGATTT 1064

RESULT 9
US-10-450-763-22058
; Sequence 22058, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22058
;   LENGTH: 1088
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
```

```
;   NAME/KEY: SIMILAR
;   LOCATION: (1)..(687)
;   OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=1199.
US-10-450-763-22058

Alignment Scores:
Pred. No.:              7.6e-143      Length:          1088
Score:                  1116.00       Matches:         212
Percent Similarity:     99.1%         Conservative:    1
Best Local Similarity:  98.6%         Mismatches:      2
Query Match:            93.6%         Indels:          0
DB:                     9             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-450-763-22058 (1-1088)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy        213 SerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 645

RESULT 10
US-10-066-543-1067
; Sequence 1067, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1067
;   LENGTH: 674
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11, 377, 640, 654
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1067

Alignment Scores:
Pred. No.:              3.05e-141     Length:          674
Score:                  1102.00       Matches:         213
Percent Similarity:     98.6%         Conservative:    0
Best Local Similarity:  98.6%         Mismatches:      2
Query Match:            92.4%         Indels:          1
DB:                     5             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-1067 (1-674)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 77

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        133 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

Qy        192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProAr 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy        212 gSerLeuAspLeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||   |||||||   |||||||||
```

```
Db        618 AAGTTTGGACCTTCTTGGATTANAAAAACTGGACTGNGGTAGATTT 663

RESULT 11
US-09-745-288-63/c
; Sequence 63, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  TITLE OF INVENTION:  COMPOUNDS FOR IMMUNOTHERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE:  210121.446D1
;  CURRENT APPLICATION NUMBER: US/09/745,288
;  CURRENT FILING DATE:  2000-12-19
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-745-288-63

Alignment Scores:
Pred. No.:              2.91e-139      Length:          1079
Score:                  1090.00        Matches:         211
Percent Similarity:     99.5%          Conservative:    0
Best Local Similarity:  99.5%          Mismatches:      1
Query Match:            91.4%          Indels:          1
DB:                     3              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-745-288-63 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 12
US-10-453-919-63/c
; Sequence 63, Application US/10453919
; Publication No. US20040033230A1
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.446C7
;  CURRENT APPLICATION NUMBER: US/10/453,919
;  CURRENT FILING DATE:  2003-06-03
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-453-919-63

Alignment Scores:
Pred. No.:              2.91e-139      Length:          1079
Score:                  1090.00        Matches:         211
Percent Similarity:     99.5%          Conservative:    0
Best Local Similarity:  99.5%          Mismatches:      1
Query Match:            91.4%          Indels:          1
DB:                     7              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-453-919-63 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538
```

```
Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||||||||||||
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 13
US-10-066-543-1070
; Sequence 1070, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1070
;   LENGTH: 644
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-066-543-1070

Alignment Scores:
Pred. No.:              1.02e-136       Length:         644
Score:                  1069.00         Matches:        203
Percent Similarity:     100.0%          Conservative:   1
Best Local Similarity:  99.5%           Mismatches:     0
Query Match:            89.7%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-1070 (1-644)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy        213 SerLeuAspLeu 216
              ||||||||||||
Db        632 AGTTTGGACCTT 643

RESULT 14
US-10-066-543-1084
; Sequence 1084, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1084
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1084

Alignment Scores:
Pred. No.:              1.16e-134       Length:         633
Score:                  1054.00         Matches:        200
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            88.4%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-1084 (1-633)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213
```

```
Qy          73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy          93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy         113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy         133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy         153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy         173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy         193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArg 212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 633

RESULT 15
US-10-066-543-548
; Sequence 548, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 19, 71
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-548

Alignment Scores:
Pred. No.:              1.7e-131      Length:         625
Score:                  1031.00       Matches:        198
Percent Similarity:     99.0%         Conservative:   0
Best Local Similarity:  99.0%         Mismatches:     1
Query Match:            86.5%         Indels:         1
DB:                     5             Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-548 (1-625)

Qy          13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
               ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||

Db          26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy          33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy          53 -GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 72
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy          72 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 92
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy          92 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy         112 yHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy         132 oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy         152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy         172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLy 192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy         192 sLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIlePro 211
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCC 623

RESULT 16
US-10-066-543-1031
; Sequence 1031, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1031
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 10
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1031

Alignment Scores:
Pred. No.:              5.62e-130     Length:         625
```

```
Score:                  1020.00       Matches:       194
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            85.6%         Indels:        0
DB:                     5             Gaps:          0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-1031 (1-625)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy         33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy         53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        153 CysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHis 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy        173 CysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 581

Qy        193 LeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGlu 206
              ||||||||||||||||||||||||||||||||||||||||||
Db        582 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAG 623

RESULT 17
US-10-757-745-3
; Sequence 3, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (122)..(1234)

;   OTHER INFORMATION:
US-10-757-745-3

Alignment Scores:
Pred. No.:              4.05e-119     Length:        1312
Score:                  945.00        Matches:       179
Percent Similarity:     89.0%         Conservative:  23
Best Local Similarity:  78.9%         Mismatches:    23
Query Match:            79.3%         Indels:        2
DB:                     9             Gaps:          1

US-10-757-745-2_COPY_121_347 (1-227) x US-10-757-745-3 (1-1312)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||||||||||||||  ||||||  ||||||||||||||||||||||||||  |||
Db        512 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 571

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||:::||||||||||||||||||||||||||||||   :::|||
Db        572 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTAC 631

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::::::::|||   ||||||||||||:::||||||||||||||||||||
Db        632 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 691

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||:::||||||||   ||||||||   ||||||||||||||||||||||||||||||
Db        692 GCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 751

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::|||||||||||||||||||||||||||||:::|||::::::   |||||||||
Db        752 AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 811

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||   |||:::|||||||:::   |||
Db        812 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 871

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||   ||||||   ||||||||||||||||||:::|||   ||||||||||||||||
Db        872 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 931

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||:::|||||||   :::|||||||||||||||:::|||:::
Db        932 GGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTT 991

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        992 TTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAA 1051

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||:::||||||   |||   ||||||   |||   ||||||||||||||||||||||
Db       1052 GCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTC 1111

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||      ||||||||||||||:::||||||:::||||||||||||:::||||||||||
Db       1112 AGA------GCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 1165

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1166 AAACTGGACTGTGGTAGATTT 1186

RESULT 18
US-10-066-543-983
; Sequence 983, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
```

```
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 983
;   LENGTH: 553
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 6, 35
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-983

Alignment Scores:
Pred. No.:               8.91e-114      Length:          553
Score:                   902.00         Matches:         171
Percent Similarity:      98.3%          Conservative:    0
Best Local Similarity:   98.3%          Mismatches:      3
Query Match:             75.7%          Indels:          0
DB:                      5              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-983 (1-553)

Qy           9 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 28
               |||          ||||||||||||||||||||||||||||||||||||||||||||||||
Db          30 CTCTANATGCATGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATA 89

Qy          29 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 48
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          90 TTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTAT 149

Qy          49 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 68
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         150 GAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 209

Qy          69 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 88
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         210 GTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTT 269

Qy          89 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         270 TTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAG 329

Qy         109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         330 AGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATG 389

Qy         129 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 148
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         390 CAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGA 449

Qy         149 GluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGly 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         450 GAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGC 509

Qy         169 LysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 182
               ||||||||||||||||||||||||||||||||||||||||||
Db         510 AAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 551

RESULT 19
US-10-066-543-1363
; Sequence 1363, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1363
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-066-543-1363

Alignment Scores:
Pred. No.:               2.34e-79       Length:          401
Score:                   651.00         Matches:         127
Percent Similarity:      100.0%         Conservative:    0
Best Local Similarity:   100.0%         Mismatches:      0
Query Match:             54.6%          Indels:          0
DB:                      5              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-066-543-1363 (1-401)

Qy          13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          19 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 78

Qy          33 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 52
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          79 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 138

Qy          53 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         139 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 198

Qy          73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 92
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         199 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258

Qy          93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         259 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 318

Qy         113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         319 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 378

Qy         133 GluSerAlaThrValIlePhe 139
               |||||||||||||||||||||
Db         379 GAGTCAGCTACAGTTATATTT 399

RESULT 20
US-10-242-535A-8082
; Sequence 8082, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
```

```
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8082
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-8082

Alignment Scores:
Pred. No.:              6.1e-71       Length:          400
Score:                  590.00        Matches:         111
Percent Similarity:     100.0%        Conservative:    1
Best Local Similarity:  99.1%         Mismatches:      0
Query Match:            49.5%         Indels:          0
DB:                     7             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-242-535A-8082 (1-400)

Qy        103 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 122
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GTGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAA 124

Qy        123 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGAT 184

Qy        143 ThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGAT 244

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 GTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 304

Qy        183 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAla 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCA 364

Qy        203 AlaAlaGluGluGlyHisIleIleProArgSerLeu 214
              ||||||||||||||||||||||||||||||||||||
Db        365 GCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 400

RESULT 21
US-10-085-783A-8082
; Sequence 8082, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8082
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-8082

Alignment Scores:
Pred. No.:              6.1e-71       Length:          400
Score:                  590.00        Matches:         111
Percent Similarity:     100.0%        Conservative:    1
Best Local Similarity:  99.1%         Mismatches:      0
Query Match:            49.5%         Indels:          0
DB:                     7             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-085-783A-8082 (1-400)

Qy        103 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 122
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GTGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAA 124

Qy        123 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGAT 184

Qy        143 ThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGAT 244

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsn 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 GTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAAC 304

Qy        183 SerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAla 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCA 364

Qy        203 AlaAlaGluGluGlyHisIleIleProArgSerLeu 214
              ||||||||||||||||||||||||||||||||||||
Db        365 GCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 400

RESULT 22
US-10-450-763-22059
; Sequence 22059, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22059
;   LENGTH: 1227
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (84)..(203)
;   OTHER INFORMATION: 85% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=164.
US-10-450-763-22059

Alignment Scores:
Pred. No.:              6.36e-53      Length:          1227
Score:                  465.00        Matches:         115
Percent Similarity:     54.8%         Conservative:    11
Best Local Similarity:  50.0%         Mismatches:      22
Query Match:            39.0%         Indels:          82
DB:                     9             Gaps:            5
```

```
US-10-757-745-2_COPY_121_347 (1-227) x US-10-450-763-22059 (1-1227)

Qy          2 AspGlyLeuAsp---LeuAsnAsnLeu-SerGlu-ArgAlaArgGlyValCysSerTyrL 20
              ||||||:::      |||||||||||| |||||| ||||||||||||||||||||||||
Db        444 GATGGGATTAGGATCTTAAACAATCTGTTCAGAGGAGGGCTCGAGGGGTGTGTTCCTACT 503

Qy         20 euAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerT 40
              ||||||||        ::::::|||          |||                ||||
Db        504 TAGCTTTGTAAGTATTATCACTTCTATTTAATAGTAATGTG---------------TCTT 548

Qy         40 yrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheT 60
              ||:::
Db        549 ATGTA---------------------------------------------------- 553

Qy         60 hrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheP 80

Db        553 ---------------------------------------------------------- 553

Qy         80 roSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluL 100

Db        553 ---------------------------------------------------------- 553

Qy        100 euCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnG 120
                                                          ||||||||||||||
Db        554 ----------------------------------------TATACGGAACGAATGAATC 572

Qy        120 lnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheA 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 AGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTG 632

Qy        140 laGlyAspThrAsnLeuArgAspArgGluValThr---ArgCysGlyGlyLeuProAsnA 159
              ||||||||||||||||||||||||||||||||:::   :::         :::     :
Db        633 CAGGAGATACAAATCTAAGGGATCGAGAGGTGAGTGACAAAGTCAGTTTGATGTTTGAGG 692

Qy        159 snIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspT 179
              ::   :::   :::|||      :::     |||       ||||||||||||||||||
Db        693 ACAGATTATTGCTATGGAGACTAATTTAGGTTCCT---TTTTGCCAGTATACATGGGATA 749

Qy        179 hrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIleP 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        750 CACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATAT 809

Qy        199 hePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyL 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        810 TTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGAT 869

Qy        219 euGluLysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||||||
Db        870 TAGAAAAACTGGACTGTGGTAGATTT 895

RESULT 23
US-09-878-722-121
; Sequence 121, Application US/09878722
; Publication No. US20020040127A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Clapper, Jonathan
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.524
;  CURRENT APPLICATION NUMBER: US/09/878,722
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 245
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-878-722-121

Alignment Scores:
Pred. No.:              3.31e-31      Length:          176
Score:                  298.00        Matches:         58
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            25.0%         Indels:          0
DB:                     3             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-878-722-121 (1-176)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 62

Qy         44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 122

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176

RESULT 24
US-09-878-178-50/c
; Sequence 50, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527
;  CURRENT APPLICATION NUMBER: US/09/878,178
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 2237
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-878-178-50

Alignment Scores:
Pred. No.:              3.31e-31      Length:          176
Score:                  298.00        Matches:         58
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            25.0%         Indels:          0
DB:                     3             Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-878-178-50 (1-176)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy         44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 25
US-09-904-456-121
; Sequence 121, Application US/09904456
; Publication No. US20030017167A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
```

```
;  FILE REFERENCE: 210121.524C1
;  CURRENT APPLICATION NUMBER: US/09/904,456
;  CURRENT FILING DATE:  2001-07-11
;  NUMBER OF SEQ ID NOS: 247
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-904-456-121

Alignment Scores:
Pred. No.:              3.31e-31       Length:         176
Score:                  298.00         Matches:        58
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            25.0%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-904-456-121 (1-176)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 62

Qy         44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 122

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176

RESULT 26
US-10-046-935-50/c
; Sequence 50, Application US/10046935
; Publication No. US20020156011A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C1
;  CURRENT APPLICATION NUMBER: US/10/046,935
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2239
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-046-935-50

Alignment Scores:
Pred. No.:              3.31e-31       Length:         176
Score:                  298.00         Matches:        58
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            25.0%          Indels:         0
DB:                     5              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-046-935-50 (1-176)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy         44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 27
US-10-146-502-50/c
; Sequence 50, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C2
;  CURRENT APPLICATION NUMBER: US/10/146,502
;  CURRENT FILING DATE:  2002-05-14
;  NUMBER OF SEQ ID NOS: 2241
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-146-502-50

Alignment Scores:
Pred. No.:              3.31e-31       Length:         176
Score:                  298.00         Matches:        58
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            25.0%          Indels:         0
DB:                     5              Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-146-502-50 (1-176)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy         44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 28
US-10-060-036-1367/c
; Sequence 1367, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
;  APPLICANT: Benson, Darin R.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Persing, David H.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF PANCREATIC CANCER
;  FILE REFERENCE: 210121.566
;  CURRENT APPLICATION NUMBER: US/10/060,036
;  CURRENT FILING DATE:  2002-01-30
;  NUMBER OF SEQ ID NOS: 4560
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1367
;   LENGTH: 176
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-060-036-1367
```

```
Alignment Scores:
Pred. No.:              3.31e-31        Length:         176
Score:                  298.00          Matches:        58
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            25.0%           Indels:         0
DB:                     5               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-10-060-036-1367 (1-176)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 43
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 115

Qy         44 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         54 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1

RESULT 29
US-10-424-599-132360
; Sequence 132360, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132360
;   LENGTH: 1686
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_90528C.1
US-10-424-599-132360

Alignment Scores:
Pred. No.:              5.83e-15        Length:         1686
Score:                  191.50          Matches:        65
Percent Similarity:     45.1%           Conservative:   40
Best Local Similarity:  27.9%           Mismatches:     93
Query Match:            16.1%           Indels:         35
DB:                     7               Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-424-599-132360 (1-1686)

Qy          9 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 28
              |||   :::|||   :::  :::   :::     :::|||||||||   |||
Db        658 TTGCACAAGAGGATGAAGGCTATTGGCGACCTTGTTCAGTTCCATTCCCCTGACTTTATC 717

Qy         29 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer------Ser 46
                 |||||||||   |||     |||        |||:::   :::      |||
Db        718 TGTTTCCAGGAGGTTACTCCCAATATATATGACATTTTCAAGCGATCCACTTGGTGGAGT 777

Qy         47 AsnTyrGluIleIleThrGlyHisGluGluGlyTyr---------PheThrAlaIleMet 63
                 |||          |||     |||          :::      :::::
Db        778 GGGTATCGTTGCTCAGTTTCTTCTGAGATGGATTATTCAAGACCCTACTTTTGTATGGTG 837

Qy         64 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 83
              |||   |||       ||||||      |||::::::      ||||||   ::::::
Db        838 TTAAGCAAACTGCCTGTGAAATCATTCAGCAAGAAG------CCCTTCAGCAATTCTATA 891

Qy         84 MetMetArgAsnLeuLeuCysValHisValAsn---ValSerGlyAsnGluLeuCysLeu 102
                 |||   |||   |||             |||          |||||||      |||   :::
Db        892 ATGGGAAGAGAACTTTGCATTGCGGAGGTTGAAGCTGCAAGTGGCAAGCCATTGGTTATT 951

Qy        103 MetThrSerHisLeuGluSer----------------------ThrArgGlyHisAla 114
                 ||||||||||||||||||                      :::   ::::::
Db        952 GCCACTAGCCATCTTGAGAGTCCCTGTCCAGCTCCTCCAAAATGGGATCAGATGTACAGC 1011

Qy        115 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 134
                 |||||:::   |||          :::     :::::::::   |||
Db       1012 AAGGAAAGAGTAGTGCAGGCCAATGAGGCTATAAACCTTCTCAAGAAACAGCCA------ 1065

Qy        135 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 154
                 :::|||:::|||   ||||||   |||      |||::::::      :::
Db       1066 ---AGTGTTGTTTTTGGGGGAGACATGAACTGGAATGACCAACAGGATGGTCAATAT--- 1119

Qy        155 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 174
                 |||   :::      ||||||   |||      |||   :::|||
Db       1120 CCTCTACAAGATGGATGGGTTGATGCCTGGTCTCAGCTA---AGACCAAATGAAAGTGGT 1176

Qy        175 TyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 194
                 :::|||:::||||||:::   |||      |||                 :::   |||
Db       1177 TGGACTTATGACACCAAGTCAAACCAGATGTTGACAGGCAACCGTACTCTCCAAAAGCGA 1236

Qy        195 PheAspArgIlePhe--------------PheArgAlaAlaAlaGluGluGlyHisIle 209
                 ||||||                   |||:::
Db       1237 TTAGATCGCTTTATTTGCCAATTCACTGATTTTAAGATAACC----------------- 1278

Qy        210 IleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 222
                    |||:::|||:::::::|||:::|||   :::
Db       1279 ---------AGTGTTGACATGATTGGGATGGAAGCAATA 1308

RESULT 30
US-09-864-761-16762
; Sequence 16762, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16762
;   LENGTH: 444
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
US-09-864-761-16762

Alignment Scores:
Pred. No.:                1.24e-14        Length:          444
Score:                    182.50          Matches:         37
Percent Similarity:       74.5%           Conservative:    4
Best Local Similarity:    67.3%           Mismatches:      11
Query Match:              15.3%           Indels:          3
DB:                       3               Gaps:            1

US-10-757-745-2_COPY_121_347 (1-227) x US-09-864-761-16762 (1-444)

Qy          15 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 31
               |||:::|||       |||                |||||||||||||||||||||||||
Db         221 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 280

Qy          32 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 51
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         281 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 340

Qy          52 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 66
               ||||||:::       |||   |||   :::      :::|||
Db         341 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 385

RESULT 31
US-09-864-761-33229
; Sequence 33229, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33229
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
;   OTHER INFORMATION: SWISSPROT HIT: P18004, EVALUE 7.00e-01
;   OTHER INFORMATION: NT HIT: AF201687.1, EVALUE 3.00e-45
;   OTHER INFORMATION: EST_HUMAN HIT: AV717253.1, EVALUE 5.00e-45
US-09-864-761-33229

Alignment Scores:
Pred. No.:                2.49e-14        Length:          691
Score:                    182.50          Matches:         37
Percent Similarity:       74.5%           Conservative:    4
Best Local Similarity:    67.3%           Mismatches:      11
Query Match:              15.3%           Indels:          3
DB:                       3               Gaps:            1

US-10-757-745-2_COPY_121_347 (1-227) x US-09-864-761-33229 (1-691)

Qy          15 GlyValCysSerTyrLeuAlaLeu---------TyrSerProAspValIlePheLeuGln 31
               |||:::|||       |||                |||||||||||||||||||||||||
Db          83 GGTATTTGTCATAAATTGATGTTTTTCATTAGGTACAGCCCAGATGTGATATTTCTACAG 142

Qy          32 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 51
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         143 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATT 202

Qy          52 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 66
               ||||||:::       |||   |||   :::      :::|||
Db         203 ACAGGTAACAATTTTTGTTATTCAACTTGTCTTTTCTTTGAAAAA 247

RESULT 32
US-10-437-963-42449
; Sequence 42449, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
```

; SEQ ID NO 42449
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_456C.1
US-10-437-963-42449

Alignment Scores:

Pred. No.:	2.57e-11	Length:	1497
Score:	164.50	Matches:	60
Percent Similarity:	44.0%	Conservative:	28
Best Local Similarity:	30.0%	Mismatches:	81
Query Match:	13.8%	Indels:	31
DB:	7	Gaps:	9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-437-963-42449 (1-1497)

```
Qy         23 TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 42
              ::::::||||||:::|||       |||||||||   |||       |||       |||::: 
Db        772 CACAACCCTGACCTTATATGCTTCCAGGAGGTTACACCAAACATATATCTGCTTCTACAA 831

Qy         43 LysArgSer------SerAsnTyrGluIleIleThrGlyHis-----------GluGlu 56
              |||             |||                  :::               :::  
Db        832 AAATCTGACTGGTGGCAAGAATACCGATGCTCGCTGTCGAATAGCATGGCTATGCAGCGG 891

Qy         57 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 76
                |||:::          ::::::   |||::::::   ::::::|||     :::  
Db        892 AAATATTAC---------TGCATGCAGATGAGCAAATTGCCTGTTGAGTCTTTTGACTGC 942

Qy         77 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 96
              ||||||   ::::::  |||   |||  |||         ||||||
Db        943 ACTCCATTTTCCAACTCAATAATGGGAAGAGAACTGTGTGTAGCACATGTCAAAACTGGA 1002

Qy         97 GlyAsn---GluLeuCysLeuMetThrSerHisLeuGluSer----------------- 109
              |||       :::|||   |||  ||||||||||||||||||
Db       1003 GGTGCAGTTAAGCTAGTGCTGGCCACTAGCCACCTTGAGAGTCCAATGCCTGGGCCTCCT 1062

Qy        110 ------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
                   :::   ::::::  ||||||:::   |||                ||||||
Db       1063 ACTTGGGATCAGATGTACAGCACGGAACGAGTTGCTCAGGCAAACAAATCTCTCAAAATT 1122

Qy        128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 147
              :::                  |||||||||   ||||||   |||      |||
Db       1123 TTG---------GGATCCTTCCGCAATGTAATATTTTGTGGCGATATGAACTGGGATGAC 1173

Qy        148 ArgGluValThrArgCysGlyGly-----------LeuProAsnAsnIleValAspVal 163
              :::                |||            ||||||             |||
Db       1174 AAA--------------GGAGATGGGCCCTTCCCCCTACCAGCTGGCTGGACCGATGCT 1218

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
              |||       |||   ||||||           :::|||:::|||||:::   ||||||
Db       1219 TGGATTGAGTTG---AAGCCGGGTGAAGATGGCTGGACTTATGACACAAAAGCTAACAGT 1275

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAla 203
                 |||              :::   |||  ||||||         :::   |||
Db       1276 ATGCTATCAGCCAACCGTAAGTTACAGAAGAGATTGGACCGATTCGTGTGCAAGCTGGCA 1335
```

RESULT 33
US-10-021-323-14543
; Sequence 14543, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 14543
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-034-Q1-K6-G12
US-10-021-323-14543

Alignment Scores:

Pred. No.:	3.31e-11	Length:	525
Score:	158.50	Matches:	53
Percent Similarity:	46.1%	Conservative:	29
Best Local Similarity:	29.8%	Mismatches:	63
Query Match:	13.3%	Indels:	33
DB:	7	Gaps:	7

US-10-757-745-2_COPY_121_347 (1-227) x US-10-021-323-14543 (1-525)

```
Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              :::|||   :::  :::       :::   |||:::|||||||||:::|||
Db         19 AAGAGGATGAAAGCTATTGGTGACCTTATACAACTTCATTCCCCTGATATCATTTGTTTT 78

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||   |||       |||             |||   ||||||:::
Db         79 CAAGAGGTTACTCCGATGATTTATGACATTTTT------CGTGGGTCTAATTGG------ 126

Qy         51 IleThrGlyHisGluGluGlyTyrPheThrAlaIle---------------------- 62
                        :::||||||       :::|||
Db        127 -----------TGGAAAGGGTACCGGTGCTCTATCTCGGACAACACGGCGAGTTTGAGA 174

Qy         63 ---------MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe 79
                      ::::::   |||::::::  :::||||||                ||||||
Db        175 GCTTACTTTTGTATACAGTTAAGCAAATTGCCTGTGAAATCCTTTCGTCGTGAACCTTTT 234

Qy         80 ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu 99
              ::::::   |||  |||  |||          |||   |||||||||   |||
Db        235 GATAATTCGATAATGGGGCGAGAACTTTGCATGACTGAGGTTGAAGTTTCTGGGGGTGAA 294

Qy        100 ------LeuCysLeuMetThrSerHisLeuGluSer---------------------- 109
                   |||  ::   |||:::|||||||||||||
Db        295 ACAACGCTTGTTGTTGCTACTACCCATCTCGAGAGTCCGTGTCCTGCTCCGCCGAAATGG 354

Qy        110 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 129
                 :::   ::::::  ||||||:::   |||  |||:::   :::      ::::::
Db        355 GATCAAATGTATAGTAAGGAACGTGTAGAACAAGCGAAAGTAGCGATCAATATTCTCGAG 414

Qy        130 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 149
                   |||            |||||||||  ||||||  |||       |||:::
Db        415 AATAACCCG---------AATGTTATTTTTGGGGGTGATATGAATTGGGATGACAAATTA 465

Qy        150 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 |||          |||   :::      :::|||  |||       |||
Db        466 GATGGTCGATTT---CCTTTAACTGATGGATGGATTGATGCTTGGAACGAGTTA 516
```

RESULT 34
US-09-876-889-312/c
; Sequence 312, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C3
; CURRENT APPLICATION NUMBER: US/09/876,889
; CURRENT FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 312

```
;   LENGTH: 268
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-876-889-312

Alignment Scores:
Pred. No.:                2.76e-08        Length:         268
Score:                    134.00          Matches:        26
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              11.2%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-876-889-312 (1-268)

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAA 208

Qy        222 LeuAspCysGlyArgPhe 227
              |||||||||||||||||
Db        207 CTGGACTGTGGTAGATTT 190

RESULT 35
US-10-425-115-105489
; Sequence 105489, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 105489
;   LENGTH: 898
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(898)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_27708C.1
US-10-425-115-105489

Alignment Scores:
Pred. No.:                9.86e-06        Length:         898
Score:                    121.50          Matches:        45
Percent Similarity:       40.5%           Conservative:   19
Best Local Similarity:    28.5%           Mismatches:     59
Query Match:              10.2%           Indels:         35
DB:                       8               Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x US-10-425-115-105489 (1-898)

Qy          9 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 28
              |||       |||         :::         :::       :::|||||||||:::|||
Db        457 CTGATTAGAAGGATGAATGCTATTGGAGATCTTATCCAGCATCACAGCCCAGATCTCATA 516

Qy         29 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer------Ser 46
                 |||||||||    |||       |||          :::|||
Db        517 TGCTTCCAGGAGGTCACTCCAAACATTTACCTGCTTTTCGAGAAATCAGACTGGTGGCAA 576

Qy         47 AsnTyrGluIleIleThrGlyHis-----------GluGluGlyTyrPheThrAlaIle 62
                ||:::             |||            |||       |||:::
Db        577 GCATACAAATGCTCCTTGCCGCATAAGATGGCCATGGAAAGATCATACTAC--------- 627

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
                  ::::::   ||||||:::   :::||||||   :::       ||||||   ::::::
Db        628 TGCATGCAGATGAGCAGGTTGCCGGTGAAGTCGTTTGACCGCAAACCATTCTTGAACTCA 687

Qy         83 LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn---GluLeuCys 101
              ||||||   |||   |||                |||   |||   |||:::   :::|||
Db        688 AAGATGGGACGGGAGCTGTGCGTAGCCGATGTGACCGTTGGAGGTTCGATCAAGCTGGTG 747

Qy        102 LeuMetThrSerHisLeuGluSer---------------------------------- 109
              |||   ||||||||||||||||||
Db        748 CTGGCTACAAGCCACCTGGAGAGCCCATGCCCCGGACCCTCGACGTGNGACCAGATGTTC 807

Qy        110 ---------ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 126
                        |||:::|||   :::|||   |||:::
Db        808 AGCAAGGAAAGAGTAGCTCAGGCAAGCGAGTCTTTGAACACTCTGAGAGCTTTC------ 861

Qy        127 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
                                        |||       ||||||   |||
Db        862 ----------------------CGCAACGNCATATTGTGCGGNGACATGAAC 891

RESULT 36
US-10-437-963-47256
; Sequence 47256, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 47256
;   LENGTH: 1236
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_50042C.1
US-10-437-963-47256

Alignment Scores:
Pred. No.:                3.61e-05        Length:         1236
Score:                    119.00          Matches:        60
Percent Similarity:       41.9%           Conservative:   38
Best Local Similarity:    25.6%           Mismatches:     77
Query Match:              10.0%           Indels:         60
DB:                       7               Gaps:           11

US-10-757-745-2_COPY_121_347 (1-227) x US-10-437-963-47256 (1-1236)

Qy          7 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr---------- 19
              |||||||||:::::::::       :::   :::|||
Db        400 AATAATTTAAACAAGAAGACCTTCAAAATCATGACATACAATGTCTGGATCAGAGAGGAT 459

Qy         20 ----------------------------------------LeuAlaLeuTyrSerProAsp 26
                                                     :::   |||::::::||||||
Db        460 ATTGAACTGCACAGAAGGCTGGGCGCTCTTGGAGATCTTATTCAACTTCACAACCCAGAC 519

Qy         27 ValIlePheLeuGln--------------------GluValIlePro-ProTyr----- 37
                 |||       |||                        |||   |||    |||:::
Db        520 TTCATATGCTTCCAGGTGCCAGGTAACAAGGAGAACGAGAAAATCGAACCCCTTCCTGTT 579

Qy         38 -TyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGl 57
               :::   |||         |||      :::::::::|||:::
Db        580 GTGGTTCTACACCCTCCTGCGAGAGGACTCCCACCAAATCTTGAAG-------------- 625

Qy         57 yTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIl 77
```

```
                    :::::::::   :::   |||:::::::   :::                   ||
Db        626 -------CGAAGCTTGCTGGC-AGGATGAGCAAATTACCGGTGAGCGAATCGAACCCCAT 677

Qy         77 eProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGl 97
              |||||||   |||      :::|||:::      |||||||||   |||       :::||
Db        678 CCCATTT---TCGAAATCAATAATGAAGAGGGAGCTATGCGTAGCGGTCGTCAAAACGGG 734

Qy         97 yAsn---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGl 116
              |       |||   :::   ||||||||||||||||||
Db        735 AGAGATCCATTTGGCTGTGGGCACAAGTCACCTAGAGAGCCCC------TGCCCTCTACC 788

Qy        116 uArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaTh 136
                 :::      ||||||          |||::::::   :::|||   :::|||
Db        789 TCCTCTGTGGGATCTGAAGTACAGCGAGAAGCGAGTAGCTCAGGCAAAGCAGTCTCTGGA 848

Qy        136 rVal--------------------IlePheAlaGlyAspThrAsnLeuArgAspArgGl 149
                 :::                 ||||||   ||||||   |||     |||:::
Db        849 GATTCTTGGACAATTGCGCAATGCAATCTTCTGCGGCGACATGAACTGGGAAGACAAA-- 906

Qy        149 uValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLy 169
              |||             ||||||:::      :::|||   |||       |||   ||
Db        907 -GTAGATGGTCCATTCCCTCTACCTGACGGCTGGATTGATGCTTGGGTGGAGCTG---AA 962

Qy        169 sProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThr-- 188
              ||||           :::|||:::||||||:::   |||:::   |||
Db        963 GCCAGGTGACAATGGTTGGACCTATGATACGAAAGCCAATGCCATGCTGTCAGCCAACTT 1022

Qy        189 -------------------------AlaAlaCysLysLeu 193
                                        |||||||||
Db       1023 CAAGCAGCAGAAGAGGCCAGACCAGTTCGTGTGCAAGTTG 1062
```

RESULT 37
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-783536

Alignment Scores:

Pred. No.:	5.27e-05	Length:	634
Score:	114.50	Matches:	27
Percent Similarity:	69.0%	Conservative:	2
Best Local Similarity:	64.3%	Mismatches:	10
Query Match:	9.6%	Indels:	3
DB:	4	Gaps:	1

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-783536 (1-634)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 432

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||                 ::::::|||          |||           |||   |||
Db        433 GCTTTGTAAGTATTATCACTTCTATTTAATAGTAATGTG---------TCTTATGTATAT 483

Qy         41 LeuLys 42
                  |||
Db        484 ACAAAG 489
```

RESULT 38
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-783537

Alignment Scores:

Pred. No.:	0.000149	Length:	606
Score:	111.00	Matches:	22
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	9.3%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-925-065A-783537 (1-606)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 581

Qy         21 AlaLeu 22
              ||||||
Db        582 GCTTTG 587
```

RESULT 39
US-10-425-115-66072/c
; Sequence 66072, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 66072
; LENGTH: 932

```
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_160258C.1
US-10-425-115-66072

Alignment Scores:
Pred. No.:                0.000763        Length:         932
Score:                    108.00          Matches:        58
Percent Similarity:       43.1%           Conservative:   32
Best Local Similarity:    27.8%           Mismatches:     81
Query Match:              9.1%            Indels:         40
DB:                       8               Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-425-115-66072 (1-932)

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSer---------SerAsnTyrGluIleIleThr 52
              ||||||          |||   :::::|||              |||      :::
Db        928 CCATACATATATATGCTTATGCAAAAATCTGACTGGTGGCAACAATACACATGCTTGTTG 869

Qy         53 GlyHisGluGluGly-----------TyrPheThrAlaIleMetLeuLysLysSerArg 68
                ||||||                 |||:::           ::::::   |||:::
Db        868 TCACATGAGAATGCTATTCAGATGCCGTATTAC---------TGCATGCAGTTGAGCAAA 818

Qy         69 ValLysLeuLysSerGlnGluIleIleProProSerThrLysMetMetArgAsnLeu 88
              :::      |||      :::  |||||||||  ::::::  |||  |||   |||
Db        817 ATGCCGACGAAGCCATCTGACTGCATCCCATTTTCGAATTCAACAATGGGAAGAGAACTA 758

Qy         89 LeuCysValHisValAsnValSerGly---AsnGluLeuCysLeuMetThrSerHisLeu 107
                        |||                ::::|||  |||  |||:::||||||
Db        757 TGCATCGCAAGTGTTAGGACAGGAGAGATCACAAAGTTGATTTTGGCCACAACCCACCTC 698

Qy        108 GluSerThr-ArgGlyHis----------------------AlaAlaGluArgMetAs 119
              ||||||      |||                  :::   ||||||:::
Db        697 GAGAGCCCATGCCCTGCACCTCCCAAGTGGGGATCAGATGTATAGCAAAGAACGTGTGGC 638

Qy        119 nGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePh 139
              |||   |||      |||:::   :::  :::|||  :::         |||
Db        637 TCAGGCAAAGCAATCCCTGGAATTCTGGGAGGCTTCCCCAAT-GCC---------ATACT 588

Qy        139 eAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly---------- 155
              ||||||   |||      |||:::              |||
Db        587 CTGTGGGGACATGAACTGGGACGACAAG--------------GGAGACGGCCCTTTCCC 543

Qy        156 -LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTy 175
              |||   :::         |||   |||      |||      ||||||          ::
Db        542 TCTGCAGGATGGGTGGACTGATGCTTGGGTTGAGCTG---AAACCTGGAGAGGATGGCTG 486

Qy        175 rThrTrpAspThrGlnMetAsnSerAsnLeuGly-IleThrAlaAlaCysLysLeuArgP 195
              :|||:::||||||:::   |||      |||      ||||||:::|||:::   |||
Db        485 GACCTACGACACCAAGGCCAATGGCATGCTGTCAGGCAACCGCAAGCTGCAGAAG-CGAT 427

Qy        195 heAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuA 215
              ||||||         :::   |||                              |||::::
Db        426 TGGACCGGTTTGTGTGCAAGCTAGCA-----------GACTTCAAGATGGACAGCATTG 379

Qy        215 spLeuLeuGlyLeuGluLysLeu 222
              ::::::::|||   |||   :::
Db        378 AGATGATCGGGAAGGAGGCAATA 356

RESULT 40
US-10-939-321-1
; Sequence 1, Application US/10939321
; Publication No. US20050221336A1
; GENERAL INFORMATION:
;  APPLICANT: MARTIGNETTI, John
;  APPLICANT:  DOWLING, Oonagh
;  APPLICANT:  DIFEO, Analisa
;  APPLICANT:  RAMIREZ, Maria Celeste M.
;  APPLICANT:  TUKEL, Turgut
;  APPLICANT:  GLUCKSMAN, Marc J.
;  TITLE OF INVENTION: Mutations in Capillary Morphogenesis Gene-2 (Cmg-2)
;  FILE REFERENCE: 02420/100N211-US1
;  CURRENT APPLICATION NUMBER: US/10/939,321
;  CURRENT FILING DATE:  2004-09-09
;  NUMBER OF SEQ ID NOS: 68
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;   LENGTH: 168050
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-939-321-1

Alignment Scores:
Pred. No.:                2.77            Length:         168050
Score:                    108.00          Matches:        58
Percent Similarity:       41.7%           Conservative:   45
Best Local Similarity:    23.5%           Mismatches:     94
Query Match:              9.1%            Indels:         50
DB:                       9               Gaps:           11

US-10-757-745-2_COPY_121_347 (1-227) x US-10-939-321-1 (1-168050)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
              ||||||::::::      |||                                ||||||
Db      71606 GATGGCATTGAATCTATAAAT----------------------------TACCTTGGG 71635

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                 |||                          ||||||                 :::::|||
Db      71636 CAGTATGGC------------CATTTTCAGGAA-----------------ACCCATCTC 71665

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::          |||   |||      :::   |||
Db      71666 ACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 71725

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                    ||||||:::   |||   :::          :::          |||
Db      71726 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGAATTTAAACCA 71785

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::           ::::::   |||   |||:::
Db      71786 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAACGGTAAAGGGATCAATTCAACAG 71845

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||          |||   |||
Db      71846 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 71896

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                 :::|||   ||||||      :::|||          :::|||   |||:::
Db      71897 TTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 71950

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db      71951 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 72007

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   ::|||  :::                    :::::::::|||:::
Db      72008 GTTAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGACATC 72067

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
              :::      |||      ||||||      :::||||||:::   :::   :::|||
Db      72068 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTTTTTTCAACCCTGAACTTT 72124

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAla 203
              |||   |||:::   ||||||:::|||   :::   ||||||||||:::
Db      72125 TTATAGGGT---ACAAGCAATTGCAAAATAAGAGGCAATGAAATATTCTTTAAAATATTT 72181

Qy        204 AlaGluGluGlyHisIle------------IleProArgSerLeuAspLeuLeuGlyLeu 219
                |||      |||   :::               :::::::::   :::
Db      72182 GCGGTGTTTGGTTTTTTGTTCTTGTGATAGTTTACTCAGAATGTGGGAATTGAACAATGA 72241

Qy        220 GluLysLeuAspCysGlyArg 226
              :::   :::|||   ||||||
Db      72242 GATCACATGGACACAGGAAGG 72262

RESULT 41
```

```
US-10-264-237-1230/c
; Sequence 1230, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA131P1
;  CURRENT APPLICATION NUMBER: US/10/264,237
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/16450
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: US 60/205,515
;  PRIOR FILING DATE: 2000-05-19
;  NUMBER OF SEQ ID NOS: 2876
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1230
;   LENGTH: 1877
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36)..(36)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1814)..(1814)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1840)..(1840)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1849)..(1849)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1865)..(1865)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1875)..(1875)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-1230

Alignment Scores:
Pred. No.:              0.0182          Length:         1877
Score:                  101.50          Matches:        48
Percent Similarity:     32.3%           Conservative:   28
Best Local Similarity:  20.4%           Mismatches:     68
Query Match:            8.5%            Indels:         91
DB:                     6               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-264-237-1230 (1-1877)

Qy         24 SerProAspValIlePheLeuGlnGluVal---------------IleProProTyrTyr 38
              |||   |||::::::   ||||||||||||                :::   |||   ::
Db       1663 AGCGCCGATATTGTTTGCCTGCAGGAAGTGATGGGCGCGCACGAAGTTCATCCGCTGCAT 1604

Qy         39 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly--------------- 53
                          :::|||:::||||||   :::
Db       1603 GTGGAAAACTGGCCCGATACCTCGCACTACGAGTTTCTCGCCGACACTATGTGGAGCGAT 1544

Qy         54 ------------------------HisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
                                    :::   ||||||:::          :::|||
Db       1543 TTTGCCTACGGTCGCAATGCCGTATACCCGGAAGGGCATCACGGCAACGCCGTACTG--- 1487

Qy         66 LysSerArgValLysLeuLysSerGlnGlu----------------------------- 75
                |||||||      ::::::      |||
Db       1486 ---TCGCGTTATCCCATTGAACATTATGAGAATCGCGATGTTTCGGTCGATGGTGCGGAA 1430

Qy         76 ------------------------IleIleProPheProSerThrLysMetMetArgAsn 87
                                     |||:::|||   |||   |||            :::

Db       1429 AAGCGCGGCGTGCTCTACTGCCGCATTGTGCCG---CCGATGACCGGAAAAGCGATTCAT 1373

Qy         88 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 107
              ::::::|||||||||:::   :::
Db       1372 GTGATGTGCGTACATCTGGGCCTG---------------------------------- 1349

Qy        108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
                                  |||   |||       ||||||   |||:::   :::
Db       1348 --------------CGTGAGGCGCACCGTCAGGCGCAGCTTGCGATGCTCGCCGAATGG 1304

Qy        128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 147
              :::   |||   |||:::         |||:::   |||||||||   |||
Db       1303 GTGAATGAGCTACCGGACGGCGAACCGGTATTGGTGGCGGGTGATTTCAAT--------- 1253

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167

Db       1253 ---------------------------------------------------------- 1253

Qy        168 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIle 187
                                        |||       :::   |||       |||   :::
Db       1252 ----------------------GACTGGCGGCAAAAAGCTAATCATCCGTTAAAAGTG 1217

Qy        188 ThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGly 207
                 |||                 |||   ||||||   ||||||
Db       1216 CAGGCCGGA-----------CTGGATGAGATTTTTACCCGCGCCCACGGACGCCCGGCG 1169

Qy        208 HisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 222
                       |||               ||||||   |||::::::::
Db       1168 CGCACGTTTCCGGTGCAATTTCCTCTACTACGACTGGACAGGATC 1124

RESULT 42
US-10-369-493-24191
; Sequence 24191, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24191
;   LENGTH: 1644
;   TYPE: DNA
;   ORGANISM: Xenorhabdus nematophilus
US-10-369-493-24191

Alignment Scores:
Pred. No.:              0.0382          Length:         1644
Score:                  98.50           Matches:        57
Percent Similarity:     40.9%           Conservative:   49
Best Local Similarity:  22.0%           Mismatches:     90
Query Match:            8.3%            Indels:         63
DB:                     6               Gaps:           14

US-10-757-745-2_COPY_121_347 (1-227) x US-10-369-493-24191 (1-1644)

Qy          3 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 22
              |||       :::|||      :::||||||          ::::::||||||
Db        253 GGTGAAAAGATTAACCGTACTGAAGATCGCGCT---------GTATTGCATATTGCTTTG 303

Qy         23 ------------TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 38
                          ||||||      :::      ::::::|||:::|||
Db        304 CGTAACCGCAGTAACTCTCCGATTAAGGTTGATGAGGAAGATGTTATGCCGCAGGTCAAT 363

Qy         39 SerTyrLeu---LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGly 57
```

```
Db        364 GCGGTACTGGCAAAAATGAAATCATTCAGTGAGCGCATTATTAGTGGGGAGTGGAAAGGT 423

Qy         58 TyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle 77

Db        424 TATACCGGTAAAGCTATTACTGATGTGGTAAACATCGGCATTGGTGGTTCAGATCTTGGC 483

Qy         78 ProPheProSerThr---------LysMetMetArgAsnLeuLeuCysValHisVal--- 93

Db        484 CCATATATGGTGACAGAGGCACTGAAAGCCTATAAGAACCACCTTAATATGCACTTTGTA 543

Qy         94 ---AsnValSerGlyAsnGluLeuCys-------------------------------- 101

Db        544 TCCAATGTGGATGGCACTCATATCGCGGAAACACTGAAGTCTCTGGATCCGGAAACCACG 603

Qy        102 --------------------LeuMetThrSerHisLeuGluSerThrArgGlyHisAla 114

Db        604 CTGTTCCTGATTGCATCGAAGACTTTTACCACGCAGGAAACGATGACTAACGCACATTCT 663

Qy        115 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 134

Db        664 GCC---CGTGATTGGTTCTTGCAAAGTGCAGTTGATGAAGTACATGTTGCCAAACATTTT 720

Qy        135 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 154

Db        721 GCGGCACTT-----------TCAACGAAT------GAACAGGAAGTCAGCAAGTTCGGT 762

Qy        155 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 174

Db        763 ATTGATCCGCAAAACATGTTTGAGTTCTGGGACTGGGTCGGCGGC-----------CGT 810

Qy        175 TyrThr---TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeu 193

Db        811 TATTCTCTTTGG--------------TCTGCGATTGGTTTATCCATTGTCCTGTCCATT 855

Qy        194 ArgPheAspArgIle-------------------------------------PhePhe 200

Db        856 GGATATGATAATTTCGAGCAATTGCTGAATGGTGCCCATGCTATGGATAAGCACTTTGCA 915

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeu 219

Db        916 CAAACTGCTCCTGAACAGAAT------ATTCCGGTTATTCTGGCGCTGATTGGCATC 966
```

RESULT 43
US-10-764-420-1725
; Sequence 1725, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
; APPLICANT: Lum, Pek Yee
; APPLICANT: Tan, Yejun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methods For Determining Whether An Agent
; TITLE OF INVENTION: Possesses A Defined Biological Activity
; FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764,420
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1725
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2002, 2030, 2068, 2082, 2164, 2240, 4178, 4212, 4214, 4385,
; LOCATION: 9487
; OTHER INFORMATION: n = A,T,C or G
US-10-764-420-1725

Alignment Scores:

Pred. No.:	0.87	Length:	9737
Score:	97.50	Matches:	51
Percent Similarity:	37.1%	Conservative:	40
Best Local Similarity:	20.8%	Mismatches:	111
Query Match:	8.2%	Indels:	43
DB:	9	Gaps:	9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-764-420-1725 (1-9737)

```
Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30

Db       7856 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 7915

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50

Db       7916 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 7963

Qy         51 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 59

Db       7964 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 8020

Qy         60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72

Db       8021 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 8080

Qy         73 SerGlnGluIleIleProProProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91

Db       8081 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 8140

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108

Db       8141 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 8200

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128

Db       8201 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 8260

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146

Db       8261 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 8320

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164

Db       8321 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 8380

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181

Db       8381 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 8428

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201

Db       8429 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 8482

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221

Db       8483 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 8533

Qy        222 LeuAspCysGlyArg 226

Db       8534 ATTGGGCCCAACCGG 8548
```

RESULT 44
US-10-087-192-1141/c
; Sequence 1141, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192

```
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1141
;    LENGTH: 90043
;    TYPE: DNA
;    ORGANISM: Mus musculus
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)...(90043)
;    OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1141

Alignment Scores:
Pred. No.:              29.1          Length:         90043
Score:                  97.50         Matches:        51
Percent Similarity:     37.1%         Conservative:   40
Best Local Similarity:  20.8%         Mismatches:     111
Query Match:            8.2%          Indels:         43
DB:                     5             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-087-192-1141 (1-90043)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              ||||||     :::         :::      |||   |||||:::::   |||
Db      85949 GAGAGGAAGTGCCTGATCCTGGAGGAGATCCTGGCTTACCAGCCAGACATACTGTGCCTC 85890

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||          |||:::        :::          |||
Db      85889 CAGGAAGTG---GACCACTACTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 85842

Qy         51 IleThrGlyHisGluGluGlyTyrPhe-------------------------------- 59
                 |||::::::      :::|||
Db      85841 ---CTGGGCTACCAAGGCACGTTTTTCCCCAAGCCCTGGTCACCATGTCTAGATGTGGAA 85785

Qy         60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                                  |||:::        ::::::|||   ||||||
Db      85784 CACAACAACGGTCCAGATGGCTGTGCCTTATTTTTTCTCCAAAACCGATTCAAGCTTATC 85725

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
              |||          |||         :::      :::       |||   :::   :::
Db      85724 AGCAGCACCAATATTAGGCTGACAGCCATGACCCTAAAAACCAATCAGGTGGCCATCGCA 85665

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                       ||||||   :::   |||:::      :::||||||:::
Db      85664 CAGACCCTGGAGTGCAAGGAGTCTGGGCGACAGTTCTGCATTGCTGTCACCCACTTAAAA 85605

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              :::   |||        |||      |||         :::|||:::   :::
Db      85604 GCCCGCACTGGCTGGGAGCGGTTCCGGTCAGCTCAGGGCTGTGACCTCCTCCAGAACCTG 85545

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
              |||         :::   |||   :::      |||      ||||||   |||
Db      85544 CAGAACATCACCCAGGGAGCAAAGATCCCCCTGATCGTCTGCGGGGACTTCAACGCAGAG 85485

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              |||                                  |||:::         :::
Db      85484 CCAACCGAAGAGGTCTACAAACACTTTGCGTCCTCCAGCCTCAACCTCAACAGCGCCTAC 85425

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
              :::   |||         |||:::   :::          ||||||
Db      85424 AAGCTGCTGAGTCCCGATGGACAGTCGGAGCCTCCCTACACAACCTGG------------ 85377

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                   :::      :::       |||:::          |||   |||::::::
Db      85376 ------AAGATCCGGACCTCAGGCGAGTGTCGCCACACGCTGGACTATATCTGGTACTCC 85323

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                 |||                 :::      :::|||||||||||||      |||:::
Db      85322 AGACATGCTCTGAGT---------GTCACGTCCGCCCTGGATCTACTCACTGAAGAACAG 85272

Qy        222 LeuAspCysGlyArg 226
                :::         |||
Db      85271 ATTGGGCCCAACCGG 85257

RESULT 45
US-10-282-122A-22362
; Sequence 22362, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22362
;    LENGTH: 1692
;    TYPE: DNA
;    ORGANISM: Haemophilus influenzae
US-10-282-122A-22362

Alignment Scores:
Pred. No.:              0.0645        Length:         1692
Score:                  97.00         Matches:        51
Percent Similarity:     41.0%         Conservative:   43
Best Local Similarity:  22.3%         Mismatches:     89
Query Match:            8.1%          Indels:         46
DB:                     7             Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-282-122A-22362 (1-1692)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||   ::::::      ::::::|||::|||         :::   |||   |||
Db        361 ACGCCAGTGCTTGTTGATGGCAAAGATGTTATGCCTGAAGTCAATGCTGTGCTAGCTAAA 420

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                 :::                 ||||||:::|||       :::|||||||
Db        421 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGTAAAGCC 480
```

```
Qy          63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
               :::          :::   :::         ::::::   |||:::      |||
Db         481 ATTACGGATGTTGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 540

Qy          83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
               :::   :::|||         |||   |||   :::|||         ||||||   |||
Db         541 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 600

Qy          98 AsnGlu--------------------------------------LeuCysLeuMetThr 104
                                                          |||   |||:::
Db         601 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 660

Qy         105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
               |||         :::|||:::                 :::             :::
Db         661 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCTCAAAGTGCGCGTGATTGGTTA 720

Qy         125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
               ||||||      ::::::          |||         |||      ||||||
Db         721 CTAAAAGCAGCGAAAGATGAAAGTGCAGTTGCAAAACATTTTGCAGCATTATCAACCAAT 780

Qy         145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                  :::|||      |||   :::   |||         ||||||:::   :::   |||
Db         781 GCTAAAGAT------GTAGAAAAATTTGGCATTGATACCAATAATATGTTTGAATTTTGG 834

Qy         165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
               :::::::::|||           :::|||:::   |||               |||
Db         835 GATTGGGTTGGCGGT-----------CGTTACTCTTTATGG--------------TCA 867

Qy         184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe----------------------- 195
                  :::|||::::::   |||      :::   |||
Db         868 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAACTTTGAAGCGTTATTAAAT 927

Qy         196 --------------AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210
                           |||:::   |||               |||:::         |||
Db         928 GGTGCGCATGAAATGGATAAACATTTCCTCTCTACTCCAATCGAACAAAAT------ATC 981

Qy         211 ProArgSerLeuAspLeuLeuGlyLeu 219
               |||   :::|||   |||:::||||||
Db         982 CCAACCACTTTAGCATTAGTTGGTTTA 1008
```

RESULT 46
US-10-329-670-1
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature

```
;   LOCATION: (65313)..(65313)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (80024)..(80024)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (100091)..(100091)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (102696)..(102696)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (105121)..(105121)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (107248)..(107248)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (117136)..(117136)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (119750)..(119750)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (119924)..(119924)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (120038)..(120038)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (121344)..(121344)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (122167)..(122167)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (122336)..(122336)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131340)..(131340)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131360)..(131360)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (139910)..(139910)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (140398)..(140398)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (142750)..(142750)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145058)..(145058)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145171)..(145171)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145942)..(145942)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (147197)..(147197)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (150841)..(150841)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152500)..(152500)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152530)..(152530)

Alignment Scores:
Pred. No.:              3.96e+03        Length:         1830121
Score:                  97.00           Matches:        51
Percent Similarity:     41.0%           Conservative:   43
Best Local Similarity:  22.3%           Mismatches:     89
Query Match:            8.1%            Indels:         46
DB:                     7               Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-329-670-1 (1-1830121)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||  ::::::      ::::::|||:::|||       :::  |||   |||
Db    1644110 ACGCCAGTGCTTGTTGATGGCAAAGATGTTATGCCTGAAGTCAATGCTGTGCTAGCTAAA 1644169

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
              :::          ||||||:::|||      :::||||||
Db    1644170 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGTAAAGCC 1644229

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::              :::   :::      ::::::   |||:::      |||
Db    1644230 ATTACGGATGTTGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 1644289

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
              :::  :::|||         |||  |||  :::|||        ||||||   |||
Db    1644290 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 1644349

Qy         98 AsnGlu---------------------------------------LeuCysLeuMetThr 104
                                                         |||   |||:::
Db    1644350 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 1644409

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              |||        :::|||:::                 :::          :::
Db    1644410 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCTCAAAGTGCGCGTGATTGGTTA 1644469

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||      ::::::            |||          |||     ||||||
Db    1644470 CTAAAAGCAGCGAAAGATGAAAGTGCAGTTGCAAAACATTTTGCAGCATTATCAACCAAT 1644529

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              :::|||      |||   :::   |||         ||||||:::   :::   |||
Db    1644530 GCTAAAGAT------GTAGAAAAATTTGGCATTGATACCAATAATATGTTTGAATTTTGG 1644583

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
              :::::::::|||           :::|||:::  |||                 |||
Db    1644584 GATTGGGTTGGCGGT-----------CGTTACTCTTTATGG---------------TCA 1644616

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe---------------------- 195
              :::|||::::::   |||     :::   |||
```

```
Db    1644617 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAACTTTGAAGCGTTATTAAAT 1644676

Qy        196 --------------AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210
                           |||:::   |||               |||:::         |||
Db    1644677 GGTGCGCATGAAATGGATAAACATTTCCTCTCTACTCCAATCGAACAAAAT------ATC 1644730

Qy        211 ProArgSerLeuAspLeuLeuGlyLeu 219
              |||   :::|||   |||:::||||||
Db    1644731 CCAACCACTTTAGCATTAGTTGGTTTA 1644757

RESULT 47
US-10-158-865-1
; Sequence 1, Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
;  TITLE OF INVENTION:  Thereof, and Uses Thereof
;  FILE REFERENCE: PB186P2C1D1
;  CURRENT APPLICATION NUMBER: US/10/158,865
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US 09/557,884
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US 08/476,102
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/426,787
;  PRIOR FILING DATE: 1995-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1830121
;   TYPE: DNA
;   ORGANISM: Haemophilus influenzae
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4747)..(4747)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9921)..(9921)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10150)..(10150)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (29298)..(29298)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36543)..(36543)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36551)..(36551)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36636)..(36636)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (40808)..(40810)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44416)..(44416)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44905)..(44905)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44975)..(44975)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45593)..(45593)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45732)..(45732)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (47036)..(47036)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51334)..(51334)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51602)..(51602)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51786)..(51786)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51805)..(51805)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (55369)..(55369)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65309)..(65309)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65313)..(65313)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (80024)..(80024)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (100091)..(100091)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (102696)..(102696)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (105121)..(105121)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (107248)..(107248)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (117136)..(117136)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (119750)..(119750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (119924)..(119924)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (120038)..(120038)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (121344)..(121344)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (122167)..(122167)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (122336)..(122336)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131340)..(131340)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131360)..(131360)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (139910)..(139910)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (140398)..(140398)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (142750)..(142750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145058)..(145058)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145171)..(145171)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145942)..(145942)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (147197)..(147197)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (150841)..(150841)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152500)..(152500)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152530)..(152530)

Alignment Scores:
Pred. No.:                3.96e+03        Length:          1830121
Score:                    97.00           Matches:         51
Percent Similarity:       41.0%           Conservative:    43
Best Local Similarity:    22.3%           Mismatches:      89
Query Match:              8.1%            Indels:          46

DB:                       8               Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-158-865-1 (1-1830121)

Qy          24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
               :::|||   ::::::      :::::::|||:::|||        :::   |||   |||
Db     1644110 ACGCCAGTGCTTGTTGATGGCAAAGATGTTATGCCTGAAGTCAATGCTGTGCTAGCTAAA 1644169

Qy          43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
               :::                    ||||||:::|||       :::||||||
Db     1644170 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGTAAAGCC 1644229

Qy          63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
               :::                 :::   :::         ::::::   |||:::      |||
Db     1644230 ATTACGGATGTTGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 1644289

Qy          83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
               :::  :::|||           |||   |||   :::|||           ||||||   |||
Db     1644290 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 1644349

Qy          98 AsnGlu-------------------------------------LeuCysLeuMetThr 104
                                                          |||   |||:::
Db     1644350 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 1644409

Qy         105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
               |||        :::|||:::                  :::           :::
Db     1644410 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCTCAAAGTGCGCGTGATTGGTTA 1644469

Qy         125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
               ||||||      ::::::            |||           |||     ||||||
Db     1644470 CTAAAAGCAGCGAAAGATGAAAGTGCAGTTGCAAAACATTTTGCAGCATTATCAACCAAT 1644529

Qy         145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
               :::|||       |||   :::   |||          ||||||:::   :::   |||
Db     1644530 GCTAAAGAT------GTAGAAAAATTTGGCATTGATACCAATAATATGTTTGAATTTTGG 1644583

Qy         165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
               :::::::::|||           :::|||:::  |||                   |||
Db     1644584 GATTGGGTTGGCGGT------------CGTTACTCTTTATGG---------------TCA 1644616

Qy         184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe------------------------ 195
               :::|||::::::   |||      :::   |||
Db     1644617 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAACTTTGAAGCGTTATTAAAT 1644676

Qy         196 --------------AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210
                           |||:::   |||          ||:::          |||
Db     1644677 GGTGCGCATGAAATGGATAAACATTTCCTCTCTACTCCAATCGAACAAAAT------ATC 1644730

Qy         211 ProArgSerLeuAspLeuLeuGlyLeu 219
               |||   :::|||  |||:::||||||
Db     1644731 CCAACCACTTTAGCATTAGTTGGTTTA 1644757

RESULT 48
US-10-981-687-1
; Sequence 1, Application US/10981687
; Publication No. US20050131222A1
; GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
;  TITLE OF INVENTION:  Thereof, and Uses Thereof
;  FILE REFERENCE: PB186P2C1D12
;  CURRENT APPLICATION NUMBER: US/10/981,687
;  CURRENT FILING DATE:  2004-11-05
;  PRIOR APPLICATION NUMBER: US 10/158,865
;  PRIOR FILING DATE: 2002-06-03
;  PRIOR APPLICATION NUMBER: US 09/557,884
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US 08/476,102
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/426,787
;  PRIOR FILING DATE: 1995-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
```

; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature

```
;   LOCATION: (139910)..(139910)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (140398)..(140398)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (142750)..(142750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145058)..(145058)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145171)..(145171)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145942)..(145942)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (147197)..(147197)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (150841)..(150841)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152500)..(152500)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:

Alignment Scores:
Pred. No.:              3.96e+03        Length:         1830121
Score:                  97.00           Matches:        51
Percent Similarity:     41.0%           Conservative:   43
Best Local Similarity:  22.3%           Mismatches:     89
Query Match:            8.1%            Indels:         46
DB:                     9               Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-981-687-1 (1-1830121)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||  ::::::    :::::::|||:::|||     :::   |||   |||
Db    1644110 ACGCCAGTGCTTGTTGATGGCAAAGATGTTATGCCTGAAGTCAATGCTGTGCTAGCTAAA 1644169

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                :::              ||||||:::|||     :::||||||
Db    1644170 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGTAAAGCC 1644229

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::           :::   :::       ::::::   |||:::     |||
Db    1644230 ATTACGGATGTTGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 1644289

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
              :::   :::|||         |||   |||   :::|||         ||||||   |||
Db    1644290 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 1644349

Qy         98 AsnGlu---------------------------------------LeuCysLeuMetThr 104
                                                        |||   |||:::
Db    1644350 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 1644409

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              |||          :::|||:::          :::             :::
Db    1644410 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCTCAAAGTGCGCGTGATTGGTTA 1644469

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||      ::::::          |||          |||     ||||||
Db    1644470 CTAAAAGCAGCGAAAGATGAAAGTGCAGTTGCAAAACATTTTGCAGCATTATCAACCAAT 1644529

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                 :::|||     |||   :::   |||          ||||||:::   :::   |||
Db    1644530 GCTAAAGAT------GTAGAAAAATTTGGCATTGATACCAATAATATGTTTGAATTTTGG 1644583

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
              :::::::::|||              :::|||:::   |||                |||
Db    1644584 GATTGGGTTGGCGGT-----------CGTTACTCTTTATGG---------------TCA 1644616

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe---------------------- 195
                 :::|||::::::   |||     :::   |||
Db    1644617 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAACTTTGAAGCGTTATTAAAT 1644676

Qy        196 --------------AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210
                           |||:::   |||          |||:::         |||
Db    1644677 GGTGCGCATGAAATGGATAAACATTTCCTCTCTACTCCAATCGAACAAAAT------ATC 1644730

Qy        211 ProArgSerLeuAspLeuLeuGlyLeu 219
              |||   :::|||   |||:::||||||
Db    1644731 CCAACCACTTTAGCATTAGTTGGTTTA 1644757

RESULT 49
US-10-425-115-105488
; Sequence 105488, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 105488
;   LENGTH: 800
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_27707C.1
US-10-425-115-105488

Alignment Scores:
Pred. No.:              0.0232          Length:         800
Score:                  96.50           Matches:        35
Percent Similarity:     48.1%           Conservative:   17
Best Local Similarity:  32.4%           Mismatches:     46
Query Match:            8.1%            Indels:         11
DB:                     8               Gaps:           4

US-10-757-745-2_COPY_121_347 (1-227) x US-10-425-115-105488 (1-800)

Qy          9 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 28
              |||     |||          :::          :::  :::|||||||||:::|||
Db        477 CTGATTAGAAGGATGAATGCTATTGGAGATCTTATCCAGCATCACAGCCCAGATCTCATA 536

Qy         29 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer------Ser 46
                    ||||||:::   |||       |||          :::|||
Db        537 TGCTTCCAGGAGATCACTCCAAACATTTACCTGCTTTTTGAGAAATCAGATTGGTGGCAA 596

Qy         47 AsnTyrGluIleIleThrGlyHis-----------GluGluGlyTyrPheThrAlaIle 62
              |||:::          |||          |||      |||:::
Db        597 GCATACAAATGCTCCTTGCCGCATAATATGGCCATGGAAAGACCATACTAC--------- 647

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              ::::::   ||||||:::   :::||||||   :::       ||||||   ::::::
Db        648 TGCATGCAGATGAGCAGGTTGCCGGTGAAGTCGTTTGACCGCAAACCATTCTTGAACTCA 707

Qy         83 LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn---GluLeuCys 101
              ||||||   |||   |||          |||   |||   |||:::   :::|||
Db        708 AAGATGGGACGGGAGCTGTGCGTAGCCGATGTGACCGTTGGAGGTTCGATCAAGCTGGTG 767
```

```
Qy        102 LeuMetThrSerHisLeuGluSer 109
              |||   |||   ||||||||||||
Db        768 CTGGCTACAGC-CACCTGGAGAGC 790

RESULT 50
US-10-795-159-468/c
; Sequence 468, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
;  APPLICANT: BAKALETZ et al.
;  TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
;  FILE REFERENCE: 28335/38815A
;  CURRENT APPLICATION NUMBER: US/10/795,159
;  CURRENT FILING DATE:  2004-03-05
;  PRIOR APPLICATION NUMBER: US 60/453,134
;  PRIOR FILING DATE: 2003-03-06
;  NUMBER OF SEQ ID NOS: 771
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 468
;   LENGTH: 5386
;   TYPE: DNA
;   ORGANISM: H. influenzae
US-10-795-159-468

Alignment Scores:
Pred. No.:              0.757          Length:          5386
Score:                  95.00          Matches:         51
Percent Similarity:     40.6%          Conservative:    42
Best Local Similarity:  22.3%          Mismatches:      90
Query Match:            8.0%           Indels:          46
DB:                     9              Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-795-159-468 (1-5386)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||   ::::::      ::::::|||:::|||         :::   |||   |||
Db       2429 ACGCCAGTACTTGTTGATGGTAAAGATGTTATGCCTGAAGTCAATGCTGTGTTAGCTAAA 2370

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                :::                |||||:::|||      :::||||||
Db       2369 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGCAAAGCC 2310

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::               :::   :::         ::::::   |||:::      |||
Db       2309 ATTACGGATGTAGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 2250

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
              :::   :::|||         |||   |||   :::|||         ||||||   |||
Db       2249 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 2190

Qy         98 AsnGlu-----------------------------------------LeuCysLeuMetThr 104
                                                           |||   |||:::
Db       2189 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 2130

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              |||         :::|||:::                        :::           :::
Db       2129 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCACAAAGTGCGCGTGATTGGTTA 2070

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||      ::::::             |||           |||      ||||||
Db       2069 CTAAAAGCAGCGAAAGATGAAAGTGCGGTTGCAAAACATTTTGCAGCATTATCAACCAAT 2010

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              :::|||      |||   :::   |||         ||||||:::   :::   |||
Db       2009 GCTAAAGAT------GTAGAAAAATTTGGTATTGATACCAATAACATGTTTGAATTTTGG 1956

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
              :::::::::|||               :::|||:::   |||                |||
Db       1955 GATTGGGTTGGCGGT-----------CGTTACTCTTTATGG--------------TCA 1923

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAla 203
                 :::|||::::::   |||      :::   |||:::         |||   |||
```

```
Db       1922 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAAC------TTTGAAGCGTTA 1869

Qy        204 AlaGluGluGlyHis---------------------------------------IleIle 210
                           |||                                       |||
Db       1868 TTAAATGGCGCGCATGAAATGGATAAACATTTCCGCACTACTCCAATTGAAAAAAATATT 1809

Qy        211 ProArgSerLeuAspLeuLeuGlyLeu 219
              |||   :::|||   |||:::||||||
Db       1808 CCAACTACTTTAGCCTTAGTTGGTTTA 1782

RESULT 51
US-10-795-159-685/c
; Sequence 685, Application US/10795159
; Publication No. US20050221439A1
; GENERAL INFORMATION:
;  APPLICANT: BAKALETZ et al.
;  TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
;  FILE REFERENCE: 28335/38815A
;  CURRENT APPLICATION NUMBER: US/10/795,159
;  CURRENT FILING DATE:  2004-03-05
;  PRIOR APPLICATION NUMBER: US 60/453,134
;  PRIOR FILING DATE: 2003-03-06
;  NUMBER OF SEQ ID NOS: 771
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 685
;   LENGTH: 908766
;   TYPE: DNA
;   ORGANISM: H. influenzae
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9223)..(9223)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (39640)..(39640)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (292404)..(292404)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (292495)..(292495)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (908604)..(908641)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-795-159-685

Alignment Scores:
Pred. No.:              2.48e+03       Length:          908766
Score:                  95.00          Matches:         51
Percent Similarity:     40.6%          Conservative:    42
Best Local Similarity:  22.3%          Mismatches:      90
Query Match:            8.0%           Indels:          46
DB:                     9              Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-795-159-685 (1-908766)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||   ::::::      ::::::|||:::|||         :::   |||   |||
Db     272652 ACGCCAGTACTTGTTGATGGTAAAGATGTTATGCCTGAAGTCAATGCTGTGTTAGCTAAA 272593

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                :::                |||||:::|||      :::||||||
Db     272592 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGCAAAGCC 272533

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::               :::   :::         ::::::   |||:::      |||
Db     272532 ATTACGGATGTAGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 272473

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
```

```
Db     272472 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 272413

Qy         98 AsnGlu-----------------------------------------LeuCysLeuMetThr 104

Db     272412 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 272353

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124

Db     272352 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCACAAAGTGCGCGTGATTGGTTA 272293

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144

Db     272292 CTAAAAGCAGCGAAAGATGAAAGTGCGGTTGCAAAACATTTTGCAGCATTATCAACCAAT 272233

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164

Db     272232 GCTAAAGAT------GTAGAAAAATTTGGTATTGATACCAATAACATGTTTGAATTTTGG 272179

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183

Db     272178 GATTGGGTTGGCGGT------------CGTTACTCTTTATGG---------------TCA 272146

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAla 203

Db     272145 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAAC------TTTGAAGCGTTA 272092

Qy        204 AlaGluGluGlyHis-----------------------------------------IleIle 210

Db     272091 TTAAATGGCGCGCATGAAATGGATAAACATTTCCGCACTACTCCAATTGAAAAAAATATT 272032

Qy        211 ProArgSerLeuAspLeuLeuGlyLeu 219

Db     272031 CCAACTACTTTAGCCTTAGTTGGTTTA 272005
```

```
RESULT 52
US-10-282-122A-30809
; Sequence 30809, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30809
;   LENGTH: 1650
;   TYPE: DNA
;   ORGANISM: Pasteurella multocida
US-10-282-122A-30809

Alignment Scores:
Pred. No.:                0.137          Length:          1650
Score:                    94.50          Matches:           56
Percent Similarity:       38.8%          Conservative:      44
Best Local Similarity:    21.7%          Mismatches:        97
Query Match:              7.9%           Indels:            61
DB:                       7              Gaps:              13

US-10-757-745-2_COPY_121_347 (1-227) x US-10-282-122A-30809 (1-1650)

Qy          3 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 22

Db        253 GGGGCAAAAATCAATAAAACGGAAGATCGCGCC---------GTATTACATACCGCACTT 303

Qy         23 ----------TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyr 38

Db        304 CGTAATCGTTCTAATTCGCCTGTGTTGGTTGATGGGAAAGATGTGATGCCAGAAGTGAAT 363

Qy         39 SerTyrLeu---LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGly 57

Db        364 GCGGTTTTAGCCAAAATGAAAGATTTTTGCCATCGTGTGATTTCTGGTGAATGGAAAGGT 423

Qy         58 TyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle 77

Db        424 TATACTGGTAAAGCGATTACAGATGTGGTGAACATCGGAATTGGTGGGTCTGATTTAGGA 483

Qy         78 ProPheProSerThrLysMetMetArg---------AsnLeuLeuCysValHisVal--- 93

Db        484 CCTTATATGGTAACAGAAGCCTTACGTCCTTATAAAAATCATTTGAATCTACATTTCGTG 543

Qy         94 ---AsnValSerGlyAsnGluLeuCys---------------------------------- 101

Db        544 TCTAACGTAGATGGTACACATATTGCGGAGACCTTAAAGAAAGTTAATCCAGAAACCACG 603

Qy        102 ----------------------LeuMetThrSerHisLeuGluSerThrArgGlyHisAla 114

Db        604 CTATTTTTGGTTGCGTCAAAAACCTTTACCACACAAGAAACTATGACCAATGCGCATTCT 663

Qy        115 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 134

Db        664 GCGCGT---------------AATTGGTTCTTAGCCACGGCAAAAGACGAAAGTCATGTG 708

Qy        135 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGly 154

Db        709 GCAAAACACTTTGCAGCACTGTCCACAAAC------AGTAAAGCCGTTGCTGAATTTGGT 762

Qy        155 GlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 174

Db        763 ATTGATACGAATAATATGTTTGAATTTTGGGATTGGGTAGGGGGT------------CGT 810

Qy        175 TyrThr---TrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeu 193

Db        811 TATTCATTATGG---------------TCAGCGATTGGTTTATCTATTGCCCTTTCAATC 855

Qy        194 ArgPheAspArgIle-----------------------------------PhePheArg 201

Db        856 GGTTTTGAGCATTTTGAAGCGCTTTTAGCGGGTGCACACGAAATGGATAAACATTTCCGT 915

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeu 219

Db        916 ACTGCCCCAATTGAGCAAAAT---ATTCCGACCACATTAGCGTTAATTGGTTTG 966
```

```
RESULT 53
US-10-998-197-33
; Sequence 33, Application US/10998197
; Publication No. US20050222070A1
; GENERAL INFORMATION:
;  APPLICANT: Dohrmann, Cord
;  APPLICANT:  Austen, Matthias
;  TITLE OF INVENTION: Pancreas-Specific Proteins
;  FILE REFERENCE: 2923-669
;  CURRENT APPLICATION NUMBER: US/10/998,197
;  CURRENT FILING DATE:  2004-11-29
;  PRIOR APPLICATION NUMBER: PCT/EP03/05700
;  PRIOR FILING DATE: 2003-05-30
;  PRIOR APPLICATION NUMBER: EP 02020829.4
;  PRIOR FILING DATE: 2002-09-17
;  PRIOR APPLICATION NUMBER: EP 02011963.2
;  PRIOR FILING DATE: 2002-05-29
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
;   LENGTH: 1767
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-998-197-33

Alignment Scores:
Pred. No.:              0.153          Length:          1767
Score:                  94.50          Matches:         50
Percent Similarity:     37.1%          Conservative:    41
Best Local Similarity:  20.4%          Mismatches:      111
Query Match:            7.9%           Indels:          43
DB:                     9              Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-998-197-33 (1-1767)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              ||||||          :::         :::        |||   |||||:::::   |||
Db        713 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 772

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||          |||:::         :::         |||
Db        773 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 820

Qy         51 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 59
                    |||::::::      :::|||
Db        821 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 877

Qy         60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                               |||:::         ::::::|||   ||||||
Db        878 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 937

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
              :::         |||         :::       :::       |||   :::   :::
Db        938 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 997

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                       ||||||   :::   |||:::       :::||||||:::
Db        998 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 1057

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              :::       |||         |||       |||         :::|||:::   :::
Db       1058 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 1117

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
              |||         :::   |||   :::       |||       ||||||   |||
Db       1118 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 1177

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                   |||                                 |||:::         :::
Db       1178 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 1237

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
```

```
                  :::   |||         |||:::   :::             ||||||
Db       1238 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 1285

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                        :::       :::       |||:::         |||   |||::::::
Db       1286 ------AAGATCCGGACCTCAGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 1339

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                   |||                  :::       :::||||||||||||       |||:::
Db       1340 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 1390

Qy        222 LeuAspCysGlyArg 226
                  :::   |||
Db       1391 ATTGGACCCAACAGG 1405

RESULT 54
US-10-887-553A-839
; Sequence 839, Application US/10887553A
; Publication No. US20050085436A1
; GENERAL INFORMATION:
;  APPLICANT: Garza, Dan
;  APPLICANT:  Li, Hao
;  TITLE OF INVENTION: Method to treat conditions associated
;  TITLE OF INVENTION:  with insulin signalling dysregulation
;  FILE REFERENCE: 4-33262
;  CURRENT APPLICATION NUMBER: US/10/887,553A
;  CURRENT FILING DATE:  2004-07-08
;  PRIOR APPLICATION NUMBER: 60/485,883
;  PRIOR FILING DATE: 2003-08-07
;  NUMBER OF SEQ ID NOS: 1208
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 839
;   LENGTH: 1971
;   TYPE: DNA
;   ORGANISM: human
US-10-887-553A-839

Alignment Scores:
Pred. No.:              0.182          Length:          1971
Score:                  94.50          Matches:         50
Percent Similarity:     37.1%          Conservative:    41
Best Local Similarity:  20.4%          Mismatches:      111
Query Match:            7.9%           Indels:          43
DB:                     9              Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-887-553A-839 (1-1971)

Qy         11 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 30
              ||||||          :::         :::        |||   |||||:::::   |||
Db        713 GAAAGGAAATGTCTCATCCTGGAAGAAATCCTGGCCTACCAGCCTGATATATTGTGCCTC 772

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 50
              |||||||||          |||:::         :::         |||
Db        773 CAAGAGGTG---GACCACTATTTTGACACCTTCCAGCCACTCCTCAGTAGA--------- 820

Qy         51 IleThrGlyHisGluGluGlyTyrPhe------------------------------- 59
                    |||::::::      :::|||
Db        821 ---CTAGGCTATCAAGGCACGTTTTTCCCCAAACCCTGGTCACCTTGTCTAGATGTAGAA 877

Qy         60 --------------------ThrAlaIleMetLeuLysLysSerArgValLysLeuLys 72
                               |||:::         ::::::|||   ||||||
Db        878 CACAACAATGGACCAGATGGTTGTGCCTTATTTTTTCTTCAAAACCGATTCAAGCTAGTC 937

Qy         73 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal--- 91
              :::         |||         :::       :::       |||   :::   :::
Db        938 AACAGTGCCAATATTAGGCTGACAGCCATGACATTGAAAACCAACCAGGTGGCCATTGCA 997

Qy         92 ---------HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 108
                       ||||||   :::   |||:::       :::||||||:::
Db        998 CAGACCCTGGAGTGCAAGGAGTCAGGCCGACAGTTCTGCATCGCTGTTACCCATCTAAAA 1057

Qy        109 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 128
              :::       |||         |||       |||         :::|||:::   :::
```

```
Db       1058 GCACGCACTGGCTGGGAGCGGTTTCGATCAGCTCAAGGCTGTGACCTCCTTCAGAACCTG 1117

Qy        129 GlnGluAlaProGluSerAlaThrVal------IlePheAlaGlyAspThrAsnLeuArg 146
              |||      :::   |||   :::        |||   |||||   |||
Db       1118 CAAAACATCACCCAAGGAGCCAAGATTCCCCTTATTGTGTGTGGGGACTTCAATGCAGAG 1177

Qy        147 AspArgGlu------ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
                  |||                                  |||:::         :::
Db       1178 CCAACAGAAGAGGTCTACAAACACTTTGCTTCCTCCAGCCTCAACCTGAACAGCGCCTAC 1237

Qy        165 GluPheLeu---------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMet 181
              :::   |||         |||:::   :::              ||||||
Db       1238 AAGCTGCTGAGTGCTGATGGGCAGTCAGAACCCCCATACACTACCTGG------------ 1285

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArg 201
                  :::   :::   :::          |||:::      |||   |||::::::
Db       1286 ------AAGATCCGGACCTCAGGGGAGTGCAGGCACACCCTGGATTACATCTGGTATTCT 1339

Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
                  |||                  :::      :::||||||||||||      |||:::
Db       1340 AAACATGCTCTAAAT---------GTAAGGTCAGCTCTCGATCTGCTCACTGAAGAACAG 1390

Qy        222 LeuAspCysGlyArg 226
              :::       |||
Db       1391 ATTGGACCCAACAGG 1405

RESULT 55
US-09-954-456-541
; Sequence 541, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 541
;    LENGTH: 3285
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-954-456-541

Alignment Scores:
Pred. No.:                0.476          Length:          3285
Score:                    94.00          Matches:           63
Percent Similarity:       37.7%          Conservative:      37
Best Local Similarity:    23.8%          Mismatches:        96
Query Match:              7.9%           Indels:            70
DB:                       3              Gaps:              10
```

```
US-10-757-745-2_COPY_121_347 (1-227) x US-09-954-456-541 (1-3285)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||      |||   :::|||                    :::|||||||
Db        703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy         33 ValIleProProTyr----------------------------------------TyrSer 39
              :::      |||                                            |||
Db        763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::      |||              :::||||||    ||||||       ||
Db        823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu----------- 75
              ||||      :::   :::|||||||||         :::|||   |||
Db        868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy         76 ----------------------------------------------IleIleProPhePr 80
                                                           |||:::      ||
Db        928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::     ||||||      ::::::||:::   :::   ::
Db        988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::||||||                |||        ||
Db       1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |        :::|||   :::|||   :::|||                  ::::::   ||
Db       1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||           ||||||:::||||||      |||        |||        ||
Db       1155 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |   ::::::   :::   |||           |||   :::|||      |||
Db       1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
               :::|||      :::
Db       1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||          ::::::||||||   :::          |||
Db       1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                    |||   ::::::   |||   |||:::               |||
Db       1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy        221 sLeuAspCysGly 225
               |||   ||||||
Db       1431 TCTAATTTGTGGT 1443

RESULT 56
US-09-880-107-3745
; Sequence 3745, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
```

```
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3745
;    LENGTH: 3285
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    OTHER INFORMATION: Genbank Accession No. US20020142981A1 X58528
US-09-880-107-3745

Alignment Scores:
Pred. No.:              0.476          Length:         3285
Score:                  94.00          Matches:        63
Percent Similarity:     37.7%          Conservative:   37
Best Local Similarity:  23.8%          Mismatches:     96
Query Match:            7.9%           Indels:         70
DB:                     3              Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-09-880-107-3745 (1-3285)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
Db        703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy         33 ValIleProProTyr----------------------------------------TyrSer 39
Db        763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
Db        823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu----------- 75
Db        868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy         76 ---------------------------------------------IleIleProPhePr 80
Db        928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
Db        988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
Db       1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
Db       1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
Db       1155 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
Db       1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
Db       1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
Db       1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
Db       1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy        221 sLeuAspCysGly 225
Db       1431 TCTAATTTGTGGT 1443

RESULT 57
US-10-283-975A-269
; Sequence 269, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
;  APPLICANT: Ortho-Clinical Diagnostics, Inc.
;  TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
;  FILE REFERENCE: CDS 293 PCT
;  CURRENT APPLICATION NUMBER: US/10/283,975A
;  CURRENT FILING DATE:  2002-10-30
;  PRIOR APPLICATION NUMBER: 60/340,938
;  PRIOR FILING DATE: 2001-10-30
;  PRIOR APPLICATION NUMBER: 60/338,997
;  PRIOR FILING DATE: 2001-10-30
;  PRIOR APPLICATION NUMBER: 60/340,081
;  PRIOR FILING DATE: 2001-10-30
;  PRIOR APPLICATION NUMBER: 60/341,012
;  PRIOR FILING DATE: 2001-10-30
;  NUMBER OF SEQ ID NOS: 900
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 269
;    LENGTH: 3285
;    TYPE: DNA
;    ORGANISM: HUMAN
US-10-283-975A-269

Alignment Scores:
Pred. No.:              0.476          Length:         3285
Score:                  94.00          Matches:        63
Percent Similarity:     37.7%          Conservative:   37
Best Local Similarity:  23.8%          Mismatches:     96
Query Match:            7.9%           Indels:         70
DB:                     7              Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-283-975A-269 (1-3285)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
Db        703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy         33 ValIleProProTyr----------------------------------------TyrSer 39
Db        763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
Db        823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu----------- 75
Db        868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy         76 ---------------------------------------------IleIleProPhePr 80
Db        928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
Db        988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
Db       1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
Db       1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154
```

```
Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||                 ||||||:::||||||     |||              ||
Db       1155 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |   ::::::   :::   |||                 |||   ::::||      |||
Db       1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
               :::|||          :::
Db       1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||         ::::::||||||   :::             |||
Db       1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                  |||   ::::::   |||   |||:::                |||
Db       1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy        221 sLeuAspCysGly 225
               |||   ||||||
Db       1431 TCTAATTTGTGGT 1443
```

RESULT 58
US-10-843-641A-3568
; Sequence 3568, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3568
; LENGTH: 3285
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-3568

Alignment Scores:

Pred. No.:	0.476	Length:	3285
Score:	94.00	Matches:	63
Percent Similarity:	37.7%	Conservative:	37
Best Local Similarity:	23.8%	Mismatches:	96
Query Match:	7.9%	Indels:	70
DB:	9	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-843-641A-3568 (1-3285)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||     |||  :::|||                  :::||||||
Db        703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy         33 ValIleProProTyr---------------------------------------TyrSer 39
              :::      |||                                          |||
Db        763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::     |||              :::||||||     ||||||       ||
Db        823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu----------- 75
              ||||      :::   :::||||||||         ::::|||   |||
Db        868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy         76 ---------------------------------------------IleIleProPhePr 80
                                                           |||:::      ||
Db        928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||     :::        ||||||        ::::::|||:::   :::   ::
Db        988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::||||||                |||      ||
Db       1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |         :::|||   :::|||   :::|||                :::::   ||
Db       1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||                 ||||||:::||||||     |||              ||
Db       1155 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |   ::::::   :::   |||                 |||   ::::||      |||
Db       1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
               :::|||          :::
Db       1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||         ::::::||||||   :::             |||
Db       1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                  |||   ::::::   |||   |||:::                |||
Db       1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy        221 sLeuAspCysGly 225
               |||   ||||||
Db       1431 TCTAATTTGTGGT 1443
```

RESULT 59
US-09-880-107-2403
; Sequence 2403, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054

```
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2403
;   LENGTH: 3397
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 M81182
US-09-880-107-2403

Alignment Scores:
Pred. No.:              0.502          Length:          3397
Score:                  94.00          Matches:         63
Percent Similarity:     37.7%          Conservative:    37
Best Local Similarity:  23.8%          Mismatches:      96
Query Match:            7.9%           Indels:          70
DB:                     3              Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-09-880-107-2403 (1-3397)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||      |||   :::|||                    :::||||||
Db        679 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 738

Qy         33 ValIleProProTyr-------------------------------------TyrSer 39
              :::      |||                                        |||
Db        739 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 798

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::      |||              :::||||||      ||||||       ||
Db        799 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 843

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
              ||||      :::   :::||||||||         :::|||   |||
Db        844 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 903

Qy         76 ---------------------------------------------IleIleProPhePr 80
                                                           |||:::      ||
Db        904 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 963

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::      ||||||         ::::::|||:::   :::   ::
Db        964 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1017

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::||||||                |||         ||
Db       1018 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1076

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |          :::|||   :::|||   :::|||                :::::   ||
Db       1077 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1130

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||                ||||||:::||||||      |||              ||
Db       1131 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1172

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |   ::::::   :::   |||          |||   :::|||      |||
Db       1173 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1232

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
               :::|||          :::
Db       1233 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1292

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||         ::::::||||||   :::              |||
Db       1293 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1346

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                   |||   ::::::   |||   |||:::                    |||
Db       1347 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1406

Qy        221 sLeuAspCysGly 225
                 |||   ||||||
Db       1407 TCTAATTTGTGGT 1419

RESULT 60
US-09-873-367C-166
; Sequence 166, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 166
;   LENGTH: 3397
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-166

Alignment Scores:
Pred. No.:              0.502          Length:          3397
Score:                  94.00          Matches:         63
Percent Similarity:     37.7%          Conservative:    37
Best Local Similarity:  23.8%          Mismatches:      96
Query Match:            7.9%           Indels:          70
DB:                     3              Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-09-873-367C-166 (1-3397)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||      |||   :::|||                    :::||||||
Db        679 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 738

Qy         33 ValIleProProTyr-------------------------------------TyrSer 39
              :::      |||                                        |||
Db        739 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 798

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::      |||              :::||||||      ||||||       ||
Db        799 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 843

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
              ||||      :::   :::||||||||         :::|||   |||
Db        844 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 903

Qy         76 ---------------------------------------------IleIleProPhePr 80
                                                           |||:::      ||
Db        904 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 963

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::      ||||||         ::::::|||:::   :::   ::
Db        964 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1017

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::||||||                |||         ||
Db       1018 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1076
```

```
Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
Db       1077 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1130

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
Db       1131 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1172

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
Db       1173 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1232

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
Db       1233 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1292

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
Db       1293 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1346

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
Db       1347 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1406

Qy        221 sLeuAspCysGly 225
Db       1407 TCTAATTTGTGGT 1419
```

RESULT 61
US-10-283-975A-219
; Sequence 219, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283,975A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340,938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338,997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340,081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341,012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 219
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-219

Alignment Scores:

Pred. No.:	0.502	Length:	3397
Score:	94.00	Matches:	63
Percent Similarity:	37.7%	Conservative:	37
Best Local Similarity:	23.8%	Mismatches:	96
Query Match:	7.9%	Indels:	70
DB:	7	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-283-975A-219 (1-3397)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
Db        679 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 738

Qy         33 ValIleProProTyr-------------------------------------TyrSer 39
Db        739 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 798
```

```
Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
Db        799 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 843

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu----------- 75
Db        844 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 903

Qy         76 ----------------------------------------------IleIleProPhePr 80
Db        904 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 963

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
Db        964 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1017

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
Db       1018 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1076

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
Db       1077 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1130

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
Db       1131 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1172

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
Db       1173 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1232

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
Db       1233 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1292

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
Db       1293 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1346

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
Db       1347 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1406

Qy        221 sLeuAspCysGly 225
Db       1407 TCTAATTTGTGGT 1419
```

RESULT 62
US-10-843-641A-166
; Sequence 166, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28

```
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 166
;   LENGTH: 3397
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-166

Alignment Scores:
Pred. No.:                0.502          Length:          3397
Score:                    94.00          Matches:         63
Percent Similarity:       37.7%          Conservative:    37
Best Local Similarity:    23.8%          Mismatches:      96
Query Match:              7.9%           Indels:          70
DB:                       9              Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-843-641A-166 (1-3397)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||      |||   :::|||                 :::||||||
Db        679 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 738

Qy         33 ValIleProProTyr---------------------------------------TyrSer 39
              :::      |||                                          |||
Db        739 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 798

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::      |||          :::||||||      ||||||       ||
Db        799 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 843

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu----------- 75
              ||||      :::   :::||||||||||         :::|||   |||
Db        844 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 903

Qy         76 ---------------------------------------------IleIleProPhePr 80
                                                           |||:::      ||
Db        904 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 963

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::      ||||||          ::::::|||:::   :::   ::
Db        964 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1017

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::|||||||                 |||        ||
Db       1018 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1076

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |         :::|||   :::|||   :::|||                 ::::::   ||
Db       1077 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1130

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||                ||||||:::||||||      |||              ||
Db       1131 TGGG----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1172

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |   ::::::   :::   |||             |||   :::|||      |||
Db       1173 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1232

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
              :::|||          :::
Db       1233 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1292

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||         :::::||||||   :::            |||
Db       1293 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1346

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                 |||   ::::::   |||   |||:::                |||
Db       1347 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1406

Qy        221 sLeuAspCysGly 225
              |||   ||||||
Db       1407 TCTAATTTGTGGT 1419

RESULT 63
US-10-172-118-832
; Sequence 832, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Chris
;  APPLICANT:  Van 't Veer, Laura
;  APPLICANT:  Van de Vijver, Marc
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-175-999
;  CURRENT APPLICATION NUMBER: US/10/172,118
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: 60/380,770
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 832
;   LENGTH: 3423
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: NM_002858
;   DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-832

Alignment Scores:
Pred. No.:                0.508          Length:          3423
Score:                    94.00          Matches:         63
Percent Similarity:       37.7%          Conservative:    37
Best Local Similarity:    23.8%          Mismatches:      96
Query Match:              7.9%           Indels:          70
DB:                       6              Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-172-118-832 (1-3423)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||      |||   :::|||                 :::||||||
Db        703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy         33 ValIleProProTyr---------------------------------------TyrSer 39
              :::      |||                                          |||
Db        763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::      |||          :::||||||      ||||||       ||
Db        823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu----------- 75
              ||||      :::   :::||||||||||         :::|||   |||
Db        868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy         76 ---------------------------------------------IleIleProPhePr 80
                                                           |||:::      ||
Db        928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::      ||||||          ::::::|||:::   :::   ::
Db        988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::||||||                 |||        ||
```

```
Db       1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140

Db       1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160

Db       1155 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178

Db       1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186

Db       1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204

Db       1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221

Db       1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy        221 sLeuAspCysGly 225

Db       1431 TCTAATTTGTGGT 1443
```

RESULT 64
US-10-342-887-832
; Sequence 832, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 832
; LENGTH: 3423
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-832

Alignment Scores:

Pred. No.:	0.508	Length:	3423
Score:	94.00	Matches:	63
Percent Similarity:	37.7%	Conservative:	37
Best Local Similarity:	23.8%	Mismatches:	96
Query Match:	7.9%	Indels:	70
DB:	7	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-342-887-832 (1-3423)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32

Db        703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy         33 ValIleProProTyr---------------------------------------TyrSer 39

Db        763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59

Db        823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu----------- 75

Db        868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy         76 ----------------------------------------------IleIleProPhePr 80

Db        928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100

Db        988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120

Db       1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140

Db       1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160

Db       1155 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178

Db       1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186

Db       1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204

Db       1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221

Db       1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy        221 sLeuAspCysGly 225

Db       1431 TCTAATTTGTGGT 1443
```

RESULT 65
US-10-887-553A-436
; Sequence 436, Application US/10887553A
; Publication No. US20050085436A1
; GENERAL INFORMATION:
; APPLICANT: Garza, Dan
; APPLICANT: Li, Hao
; TITLE OF INVENTION: Method to treat conditions associated
; TITLE OF INVENTION: with insulin signalling dysregulation
; FILE REFERENCE: 4-33262
; CURRENT APPLICATION NUMBER: US/10/887,553A
; CURRENT FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/485,883
; PRIOR FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 1208
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 436
; LENGTH: 3423
; TYPE: DNA
; ORGANISM: human

US-10-887-553A-436

Alignment Scores:

Pred. No.:	0.508	Length:	3423
Score:	94.00	Matches:	63
Percent Similarity:	37.7%	Conservative:	37
Best Local Similarity:	23.8%	Mismatches:	96
Query Match:	7.9%	Indels:	70
DB:	9	Gaps:	10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-887-553A-436 (1-3423)

```
Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||      |||   :::|||                    :::||||||
Db        703 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 762

Qy         33 ValIleProProTyr---------------------------------------TyrSer 39
              :::       |||                                        |||
Db        763 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 822

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::     |||              :::||||||     ||||||         ||
Db        823 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 867

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
              ||||      :::   :::|||||||||         :::|||   |||
Db        868 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 927

Qy         76 ---------------------------------------------IleIleProPhePr 80
                                                           |||:::      ||
Db        928 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 987

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::      ||||||         ::::::|||:::   :::   ::
Db        988 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 1041

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::||||||              |||          ||
Db       1042 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 1100

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |          :::|||   :::|||   :::|||                 ::::::   ||
Db       1101 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1154

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||                  ||||||:::||||||     |||          ||
Db       1155 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1196

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |   ::::::  :::   |||          |||   :::|||      |||
Db       1197 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1256

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
              :::|||          :::
Db       1257 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1316

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||          ::::::||||||   :::              |||
Db       1317 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1370

Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
                 |||   ::::::   |||   |||:::                  |||
Db       1371 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1430

Qy        221 sLeuAspCysGly 225
              |||   ||||||
Db       1431 TCTAATTTGTGGT 1443
```

RESULT 66
US-09-767-041-43
; Sequence 43, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilda
; TITLE OF INVENTION: STREPTOCOCCUS SUIS VACCINES AND DIAGNOSTIC TESTS
; FILE REFERENCE: 2183-4726
; CURRENT APPLICATION NUMBER: US/09/767,041
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: PCT/NL99/00460
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: EP98202465.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: EP98202467.1
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
; LENGTH: 3738
; TYPE: DNA
; ORGANISM: Streptococcus suis
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: CPS7
US-09-767-041-43

Alignment Scores:

Pred. No.:	0.584	Length:	3738
Score:	94.00	Matches:	46
Percent Similarity:	35.6%	Conservative:	33
Best Local Similarity:	20.7%	Mismatches:	79
Query Match:	7.9%	Indels:	64
DB:	3	Gaps:	7

US-10-757-745-2_COPY_121_347 (1-227) x US-09-767-041-43 (1-3738)

```
Qy         23 TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 42
              ||||||||||||      ::::::|||:::|||:::|||                |||
Db       1728 TATAGTCCTGATGAAGCATATGTTCAAAAAGTATTACCAGAAAAAATGAAGTACAATTTG 1787

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
              :::          |||:::   ||||||:::        :::             |||   |||
Db       1788 GAATATATCAGAAACTTTGGAATTATTTCTGATTTTAAAGTAATGATTGATACAGTAATT 1847

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhe--------- 79
              :::   |||       :::|||:::      ::::::   |||||||||
Db       1848 AAAGTAATAAAATAGGAGATTAAAATGACAAAAAGACAAAATATTCCATTTTCACCACCA 1907

Qy         79 ------------------------------------------------------------ 79

Db       1908 GATATTACCCAAGCTGAAATTGATGAAGTTATTGACACACTAAAATCTGGTTGGATTACA 1967

Qy         80 ------ProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly 97
                    |||   ||||||   :::          |||   |||            :::
Db       1968 ACAGGACCAAAGACAAAAGAGCTAGAACGTCGGCTATCAGTATTTACAGGA---ACCAAT 2024

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
              :::||||||   :::            :::|||   |||
Db       2025 AAAACTGTGTGTTTAAATTCT-----------GCTACTGCAGGA-------------- 2057

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                 |||::::::||||||:::   :::          |||             :::
Db       2058 ---------TTGGAACTAGTCTTACGAATTCTTGGTGTTGGACCCGGAGATGAAGTTATT 2108

Qy        138 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuPro 157
              :::   |||                        :::|||       |||      |||
Db       2109 GTTCCTGCTATGACCTATACTGCCTCATGTAGTGTCATTACTCATGTAGGAGCAACTCCT 2168

Qy        158 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::
Db       2169 GTGATGGTT---------------------------------------------------- 2177

Qy        178 AspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArg 197
              |||   |||   ||||||                          ::::::   :::|||
Db       2178 GATATTCAAAAAAACAGC--------------------TTTGAGATGGAATATGATGCT 2216

Qy        198 IlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeu 217
```

```
Db       2217 TTGGAAAAAGCGATTACTCCGAAAACAAAAGTTATCATTCCT------GTTGATCTAGCT 2270

Qy        218 GlyLeu 219
Db       2271 GGTATT 2276

RESULT 67
US-10-473-574-33
; Sequence 33, Application US/10473574
; Publication No. US20040116670A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE CORPORATION; HAFALIA, April J.A.;
;  APPLICANT:  TANG, Y. Tom; YUE, Henry;
;  APPLICANT:  KHAN, Farrah A.; ISON, Craig H.;
;  APPLICANT:  BAUGHN, Mariah R.; WARREN, Bridget A.;
;  APPLICANT:  DUGGAN, Brendan M.; THANGAVELU, Kavitha;
;  APPLICANT:  HONCHELL, Cynthia D.; AZIMZAI, Yalda;
;  APPLICANT:  ELLIOTT, Vicki S.; BURFORD, Neil;
;  APPLICANT:  DING, Li; YUE, Huibin;
;  APPLICANT:  BECHA, Shanya; EMERLING, Brooke M.;
;  APPLICANT:  RICHARDSON, Thomas W.; LEE, Soo Yeun;
;  APPLICANT:  BANDMAN, Olga; LAL, Preeti G.;
;  APPLICANT:  LEE, Sally; GIETZEN, Kimberly J.;
;  APPLICANT:  CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;
;  APPLICANT:  LEE, Ernestine A.; SWARNAKAR, Anita;
;  APPLICANT:  RING, Huijun Z.; JONES, Karen Anne
;  TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
;  FILE REFERENCE: PF-0918 USN
;  CURRENT APPLICATION NUMBER: US/10/473,574
;  CURRENT FILING DATE:  2003-09-29
;  PRIOR APPLICATION NUMBER: PCT/US02/09288
;  PRIOR FILING DATE: 2002-03-25
;  PRIOR APPLICATION NUMBER: US 60/294,451
;  PRIOR FILING DATE: 2001-05-29
;  PRIOR APPLICATION NUMBER: US 60/291,870
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: US 60/290,518
;  PRIOR FILING DATE: 2001-05-10
;  PRIOR APPLICATION NUMBER: US 60/288,609
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: US 60/283,769
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: US 60/281,323
;  PRIOR FILING DATE: 2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/280,508
;  PRIOR FILING DATE: 2001-03-29
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PERL Program
; SEQ ID NO 33
;   LENGTH: 5299
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7948948CB1
US-10-473-574-33

Alignment Scores:
Pred. No.:              1.01          Length:         5299
Score:                 94.00          Matches:          33
Percent Similarity:    37.0%          Conservative:     18
Best Local Similarity: 23.9%          Mismatches:       29
Query Match:            7.9%          Indels:           58
DB:                        7          Gaps:              7

US-10-757-745-2_COPY_121_347 (1-227) x US-10-473-574-33 (1-5299)

Qy         65 LysLysSerArgValLysLeuLysSerGlnGluIleIle---ProPheProSerThrLys 83
Db       3995 GAAAAATCAAAGGCCAAGGAACAAAAGCAGGGCATAATCAACCCATTTCCTGCTTCAAAA 4054

Qy         84 MetMetArg------AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCys 101
Db       4055 GGAATCAGAGCTTTTCCACTTCAGTGTATTCACATA---------------------- 4090

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db       4091 ----------------------GCTGAAGGGCATACA-------------------- 4105

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db       4106 ------------AAAGCTGTGCTCTGTGTGGATTCTACTGATGATCTCCTCTTCACTGGA 4153

Qy        142 AspThrAsn----------------------LeuArgAspArgGluValThrArgCys 153
Db       4154 TCAAAAGATCGTACTTGTAAAGTATGGAATCTGGTGACTGGGCAGGAAATAATGTCACTG 4213

Qy        154 GlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCys 173
Db       4214 GGGGGTCATCCCAACAATGTCGTGTCTGTA--------------------AAATACTGT 4252

Qy        174 GlnTyrThr----------------------------------------TrpAsp 178
Db       4253 AATTATACCAGTTTGGTCTTCACTGTATCAACATCTTATATTAAGGTGTGGGAT 4306

RESULT 68
US-10-473-574-31
; Sequence 31, Application US/10473574
; Publication No. US20040116670A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE CORPORATION; HAFALIA, April J.A.;
;  APPLICANT:  TANG, Y. Tom; YUE, Henry;
;  APPLICANT:  KHAN, Farrah A.; ISON, Craig H.;
;  APPLICANT:  BAUGHN, Mariah R.; WARREN, Bridget A.;
;  APPLICANT:  DUGGAN, Brendan M.; THANGAVELU, Kavitha;
;  APPLICANT:  HONCHELL, Cynthia D.; AZIMZAI, Yalda;
;  APPLICANT:  ELLIOTT, Vicki S.; BURFORD, Neil;
;  APPLICANT:  DING, Li; YUE, Huibin;
;  APPLICANT:  BECHA, Shanya; EMERLING, Brooke M.;
;  APPLICANT:  RICHARDSON, Thomas W.; LEE, Soo Yeun;
;  APPLICANT:  BANDMAN, Olga; LAL, Preeti G.;
;  APPLICANT:  LEE, Sally; GIETZEN, Kimberly J.;
;  APPLICANT:  CHAWLA, Narinder K.; GRIFFIN, Jennifer A.;
;  APPLICANT:  LEE, Ernestine A.; SWARNAKAR, Anita;
;  APPLICANT:  RING, Huijun Z.; JONES, Karen Anne
;  TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
;  FILE REFERENCE: PF-0918 USN
;  CURRENT APPLICATION NUMBER: US/10/473,574
;  CURRENT FILING DATE:  2003-09-29
;  PRIOR APPLICATION NUMBER: PCT/US02/09288
;  PRIOR FILING DATE: 2002-03-25
;  PRIOR APPLICATION NUMBER: US 60/294,451
;  PRIOR FILING DATE: 2001-05-29
;  PRIOR APPLICATION NUMBER: US 60/291,870
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: US 60/290,518
;  PRIOR FILING DATE: 2001-05-10
;  PRIOR APPLICATION NUMBER: US 60/288,609
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: US 60/283,769
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: US 60/281,323
;  PRIOR FILING DATE: 2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/280,508
;  PRIOR FILING DATE: 2001-03-29
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PERL Program
; SEQ ID NO 31
;   LENGTH: 5322
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 6260407CB1
US-10-473-574-31
```

```
Alignment Scores:
Pred. No.:                1.02          Length:          5322
Score:                    94.00         Matches:         33
Percent Similarity:       37.0%         Conservative:    18
Best Local Similarity:    23.9%         Mismatches:      29
Query Match:              7.9%          Indels:          58
DB:                       7             Gaps:            7

US-10-757-745-2_COPY_121_347 (1-227) x US-10-473-574-31 (1-5322)

Qy         65 LysLysSerArgValLysLeuLysSerGlnGluIleIle---ProPheProSerThrLys 83
              :::||||||:::   |||   :::   |||   ||||||   |||||||||:::::|||
Db       4077 GAAAAATCAAAGGCCAAGGAACAAAAGCAGGGCATAATCAACCCATTTCCTGCTTCAAAA 4136

Qy         84 MetMetArg------AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCys 101
                 :::|||          |||   |||:::|||:::
Db       4137 GGAATCAGAGCTTTTCCACTTCAGTGTATTCACATA---------------------- 4172

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
                                        ||||||
Db       4173 ----------------------GCTGAAGGGCATACA-------------------- 4187

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
                   |||   :::              :::      ::::::|||   |||
Db       4188 -----------AAAGCTGTGCTCTGTGTGGATTCTACTGATGATCTCCTCTTCACTGGA 4235

Qy        142 AspThrAsn----------------------LeuArgAspArgGluValThrArgCys 153
                 :::                        :::     :::|||:::
Db       4236 TCAAAAGATCGTACTTGTAAAGTATGGAATCTGGTGACTGGGCAGGAAATAATGTCACTG 4295

Qy        154 GlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCys 173
              ||||||   |||||||||:::|||   |||                  |||:::|||
Db       4296 GGGGGTCATCCCAACAATGTCGTGTCTGTA-------------------AAATACTGT 4334

Qy        174 GlnTyrThr-------------------------------------TrpAsp 178
                 ||||||                                     ||||||
Db       4335 AATTATACCAGTTTGGTCTTCACTGTATCAACATCTTATATTAAGGTGTGGGAT 4388

RESULT 69
US-10-369-493-44898
; Sequence 44898, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 44898
;   LENGTH: 1644
;   TYPE: DNA
;   ORGANISM: Xenorhabdus nematophilus
US-10-369-493-44898

Alignment Scores:
Pred. No.:                0.302         Length:          1644
Score:                    92.00         Matches:         49
Percent Similarity:       41.3%         Conservative:    46
Best Local Similarity:    21.3%         Mismatches:      73
Query Match:              7.7%          Indels:          62
DB:                       6             Gaps:            12

US-10-757-745-2_COPY_121_347 (1-227) x US-10-369-493-44898 (1-1644)

Qy         31 GlnGluValIleProProTyrTyrSerTyrLeu---LysLysArgSerSerAsnTyrGlu 49
              ::::::|||:::|||          :::   |||   |||   :::|||   :::
Db       340 GAAGATGTGATGCCGCAAGTCAATGCGGTGCTGGCAAAAATGAAGTCATTCAGTGAACGC 399

Qy         50 IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 69
              ||||||:::|||       :::||||||            :::           :::
Db       400 ATCATCAGCGGAGAATGGAAAGGGTATACCGGTCAAGCCATTACTGACGTTGTCAATATT 459

Qy         70 LysLeuLysSerGlnGluIleIleProPhe---------ProSerThrLysMetMetArg 86
                 :::          ::::::   |||:::          :::   |||      :::
Db       460 GGTATTGGTGGTTCCGATCTCGGCCCTTATATGGTCTCTGAAGCACTGAAAGCGTATAAA 519

Qy         87 AsnLeuLeuCysValHisVal------AsnValSerGlyAsnGluLeuCys------Leu 102
              |||   |||   :::|||          ||||||   |||       :::          |||
Db       520 AATCACCTCAATATGCACTTTGTTTCCAATGTGGATGGAACCCATATCGCAGAAACGTTG 579

Qy        103 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 122
                 |||   :::   |||:::|||                                 |||
Db       580 CAAACATTGAACCCAGAAACGACG-----------------------------CTGTTC 609

Qy        123 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly--- 141
              ::::::      |||                 |||:::   |||       ::::::
Db       610 CTGATTGCTTCCAAGACTTTCACCACCCAGGAAACTATGACCAATGCGTATTCTGCCCGC 669

Qy        141 ------------------------------------------------------------ 141

Db       670 AATTGGTTCCTGCAAAATGCGGAGATTGAGGCTCATGTTGCCAAGCATTTTGCGGCACTT 729

Qy        142 AspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleVal 161
                 ||||||      ::::::||||||:::::   |||       |||   |||:::
Db       730 TCGACCAAT------GAGCAGGAAGTCAGTAAATTCGGTATCGATCCACTGAATATGTTT 783

Qy        162 AspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGln 180
                 :::   |||:::::::::|||                 :::|||:::   |||
Db       784 GAATTCTGGGATTGGGTCGGTGGC-----------CGTTATTCACTGTGG--------- 822

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeu---------ArgPheAspArg 197
                    |||   :::|||::::::   |||       :::             |||::::::
Db       823 ------TCTGCGATTGGTTTATCCATTGCACTTTCCATTGGCTATGAGAATTTCGAACAG 876

Qy        198 IlePhePheArgAlaAlaAlaGluGluGlyHisIle----------------------- 209
                 :::          |||   |||   :::   |||
Db       877 TTGCTGAGTGGCGCTCATGCTATGGATAAACACTTCGAGCAGACTTCTCCTGATCAGAAC 936

Qy        210 IleProArgSerLeuAspLeuLeuGlyLeu 219
              ||||||      |||   |||:::|||:::
Db       937 ATTCCGATCCTGCTGGCGCTGATCGGCATC 966

RESULT 70
US-10-715-066-2/c
; Sequence 2, Application US/10715066
; Publication No. US20040127401A1
; GENERAL INFORMATION:
;  APPLICANT: The Board of Trustees of the University of Arkansas
;  APPLICANT:  O'Brien, Timothy
;  APPLICANT:  Beard, John
;  APPLICANT:  Underwood, Lowell
;  TITLE OF INVENTION: CA125 Gene and Its Use for Diagnostic and Therapeutic
;  TITLE OF INVENTION:  Interventions
;  FILE REFERENCE: 022438.43867
;  CURRENT APPLICATION NUMBER: US/10/715,066
;  CURRENT FILING DATE:  2003-11-17
;  PRIOR APPLICATION NUMBER: 60/284,175
;  PRIOR FILING DATE: 2001-04-17
;  PRIOR APPLICATION NUMBER: 60/299,380
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: 09/965,738
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: 60/345,180
;  PRIOR FILING DATE: 2001-12-02
;  PRIOR APPLICATION NUMBER: 60/427,045
;  PRIOR FILING DATE: 2002-11-15
;  NUMBER OF SEQ ID NOS: 7
```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 42348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(130)
; OTHER INFORMATION: Exon R1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (442)..(510)
; OTHER INFORMATION: Exon R2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2126)..(2250)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5479)..(5652)
; OTHER INFORMATION: Exon R3
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6301)..(6334)
; OTHER INFORMATION: Exon R4
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6593)..(6657)
; OTHER INFORMATION: Exon R5
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7558)..(7683)
; OTHER INFORMATION: Exon R1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8216)..(8284)
; OTHER INFORMATION: Exon R2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8877)..(9050)
; OTHER INFORMATION: Exon R3
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9380)..(9413)
; OTHER INFORMATION: Exon R4
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9675)..(9739)
; OTHER INFORMATION: Exon R5
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10082)..(10200)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10201)..(10291)
; OTHER INFORMATION: Exon R1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10524)..(10592)
; OTHER INFORMATION: Exon R2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11200)..(11373)
; OTHER INFORMATION: Exon R3
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11722)..(11755)
; OTHER INFORMATION: Exon R4
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12016)..(12036)
; OTHER INFORMATION: Exon R5
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12037)..(12150)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12169)..(12295)
; OTHER INFORMATION: Exon R1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12532)..(12600)
; OTHER INFORMATION: Exon R2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13219)..(13392)
; OTHER INFORMATION: Exon R3
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13723)..(13756)
; OTHER INFORMATION: Exon R4
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14016)..(14077)
; OTHER INFORMATION: Exon R5
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15001)..(15126)
; OTHER INFORMATION: Exon R1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15367)..(15435)
; OTHER INFORMATION: Exon R2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15475)..(15601)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15648)..(15773)
; OTHER INFORMATION: Exon R1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (16002)..(16070)
; OTHER INFORMATION: Exon R2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (16653)..(16826)
; OTHER INFORMATION: Exon R3
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (17158)..(17191)
; OTHER INFORMATION: Exon R4
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (17453)..(17517)
; OTHER INFORMATION: Exon R5
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (17866)..(18531)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18532)..(18657)
; OTHER INFORMATION: Exon R1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18658)..(18887)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18888)..(18956)
; OTHER INFORMATION: Exon R2
; FEATURE:

```
;   NAME/KEY: misc_feature
;   LOCATION: (18957)..(18995)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (19633)..(19806)
;   OTHER INFORMATION: Exon R3
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (20093)..(20140)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (20141)..(20176)
;   OTHER INFORMATION: Exon R4
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (20177)..(20386)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (20387)..(20449)
;   OTHER INFORMATION: Exon R5
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (20450)..(20483)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (21609)..(21731)
;   OTHER INFORMATION: Exon R1

Alignment Scores:
Pred. No.:                50.8          Length:         42348
Score:                    92.00         Matches:        41
Percent Similarity:       42.4%         Conservative:   26
Best Local Similarity:    25.9%         Mismatches:     63
Query Match:              7.7%          Indels:         28
DB:                       7             Gaps:           7

US-10-757-745-2_COPY_121_347 (1-227) x US-10-715-066-2 (1-42348)

Qy         32 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle--- 50
              :::|||||||||||||||:::          |||       |||       |||
Db      26592 AAGGTGATACCACCTTACTTCTGC---------AAGAATTGCCACAATCAAAACATCAAA 26542

Qy         51 ---------------IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
                           :::|||   ::::::|||         |||      |||:::
Db      26541 AAATAACAGATGTGTGACTCTGGGGGCCAAGATGGTCGACTATAAGCAGCAGCGCTCAGA 26482

Qy         66 LysSerArg-ValLys--------------------LeuLysSerGlnGluIleIlePr 78
                 |||       |||                  |||         :::   :::
Db      26481 GGCTCACATTGAAAAAAACTGTAATAAGTGTGTGATCCTTCCTGGCAACCAAAATGTCCA 26422

Qy         78 oPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAs 98
                |||||::|||:::   |||   |||         |||
Db      26421 GGTTCCCTCATCAAAATTGACTAGAAGGCTGGCAGTGACCCAC------AGAGAGAAGGA 26368

Qy         98 nGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMe 118
              |||   |||   :::   :::||||||     |||||||||
Db      26367 AGAGCAGTGTGGTGTGGTGGCCCACCTGAGAGACACACGGGGAAGGGGAACCCCCTCCTT 26308

Qy        118 tAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIl 138
                 ::::::|||   ::::::      :::      |||      ||||||::
Db      26307 CCAGCCAAGGGAGGTGGTGAGTGAGCGGGCTACCCAA------GAGGGGGAAACTGTGCT 26254

Qy        138 ePheAla-GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProA 158
              :|||   |||         |||  |||  :::::::::          |||   |
Db      26253 TTTTCCACGGAACTGTGCAACTCACAGATTGGAAGATCTCA-----------CTCGCGA 26206

Qy        158 snAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGln 174
              ||              |||:::   :::   :::|||::::::   |||
Db      26205 ACCCATGC------CACTGGGACCTAGTGTCCCAACCCCAGAACACGCAG 26162
```

```
RESULT 71
US-10-369-493-24652
; Sequence 24652, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24652
;   LENGTH: 1077
;   TYPE: DNA
;   ORGANISM: Archaeoglobus fulgidus
US-10-369-493-24652

Alignment Scores:
Pred. No.:                0.213         Length:         1077
Score:                    91.00         Matches:        44
Percent Similarity:       39.2%         Conservative:   27
Best Local Similarity:    24.3%         Mismatches:     69
Query Match:              7.6%          Indels:         42
DB:                       6             Gaps:           6

US-10-757-745-2_COPY_121_347 (1-227) x US-10-369-493-24652 (1-1077)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||   :::          |||          |||
Db        151 ATTGATGGGCTTACCGTGAGAAGGTTTAAATCTCTCGCCCCTTCCGAGGCGTATTAC--- 207

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                 :::||||||::::::                                  |||   |||
Db        208 ------TTCTCTCCCGAACTC----------------------------TACGACTAT 231

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||||||||   ||||||:::|||:::::::::      |||           |||
Db        232 TTAAAGAAAAATAGCTCGGACTACGATGTAGTTCACGCCCACAACTACCACGCGTTTCCC 291

Qy         61 AlaIleMet------LeuLysLysSerArgValLysLeu-LysSerGlnGluIleIlePr 78
              |||:::         |||  |||   :::   ||||||    |||:::   ||||||
Db        292 GCTCTGTTTGCTGCCTTGACCAAAGGAAAAAACAAATTGATTTTCACGCCGCATTATCAC 351

Qy         78 oPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAs 98
                    |||          ||||||:::|||                :::   |||
Db        352 GGGAGCGGTCACAGC-TTTTTCAGAAACGTCCTTCATAAACCTTACAAGATATTCGGAAG 410

Qy         98 nGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMe 118
               ::::::                  :::              :::
Db        411 GAAGATCTTTAAGAGAGCCGACGCAATCGTTTGCGTTTCCAATTAC-------------- 456

Qy        118 tAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIl 138
                 :::      :::|||||||||       :::   |||   :::          |||||
Db        457 ----GAGAAAAATCTGGTTTTGAAAAATTTCAAGGTCGCCGAGGATAGAACGTATGTAAT 512

Qy        138 ePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAs 158
              |                                                      |||||
Db        513 C------------------------------------------------------CCAAA 518

Qy        158 nAsnIle---ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTr 177
              |   |||   :::|||   ::::::   :::   |||   |||      :::   :::||
Db        519 CGGAATAAATTTAGATGAGTTTAAAGACATTGAAAAAATTAAGAGAAATAAGGAAAGTTG 578

Qy        177 p 177
              |
```

```
Db        579 G 579

RESULT 72
US-10-191-807-3
; Sequence 3, Application US/10191807
; Publication No. US20030068691A1
; GENERAL INFORMATION:
;  APPLICANT: HU, Song et al.
;  TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
;  TITLE OF INVENTION:  USES THEREOF
;  FILE REFERENCE: CL001275-PROV
;  CURRENT APPLICATION NUMBER: US/10/191,807
;  CURRENT FILING DATE:  2002-07-10
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 108359
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(108359)
;   OTHER INFORMATION: n = A,T,C or G
US-10-191-807-3

Alignment Scores:
Pred. No.:              423          Length:        108359
Score:                  90.00        Matches:       46
Percent Similarity:     42.1%        Conservative:  34
Best Local Similarity:  24.2%        Mismatches:    78
Query Match:            7.6%         Indels:        32
DB:                     5            Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-191-807-3 (1-108359)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::    :::|||  |||   :::   |||::::::          |||
Db      78182 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 78238

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::||||||                   :::||||||      |||
Db      78239 TCAGTGTGCTGTATTCAGGAA-----------------ACCTATCTCACGTGCAGAGAC 78280

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       :::   |||                  |||
Db      78281 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 78340

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::  |||  :::        :::          |||       |||||:::
Db      78341 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 78400

Qy         86 Arg-----------AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCys 101
              |||            :::   |||   |||      ::::::          ||||||
Db      78401 AGAGACAAAGAAGGCCATTACCTAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 78460

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||           |||   |||          :::
Db      78461 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 78505

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||     :::|||         :::|||   |||:::|||       |||
Db      78506 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 78562

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db      78563 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 78622

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::                 :::::::::|||::::::      |||
Db      78623 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 78682

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
```

```
                  ||||||      :::||||||:::
Db      78683 ---CACCCCAAATCAACAGAATATACATTT 78709

RESULT 73
US-10-959-243-3
; Sequence 3, Application US/10959243
; Publication No. US20050048560A1
; GENERAL INFORMATION:
;  APPLICANT: HU, Song et al.
;  TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
;  TITLE OF INVENTION:  USES THEREOF
;  FILE REFERENCE: CL001275CON
;  CURRENT APPLICATION NUMBER: US/10/959,243
;  CURRENT FILING DATE:  2004-10-07
;  PRIOR APPLICATION NUMBER: 60/305,158
;  PRIOR FILING DATE: 2001-07-16
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 108359
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(108359)
;   OTHER INFORMATION: n = A,T,C or G
US-10-959-243-3

Alignment Scores:
Pred. No.:              423          Length:        108359
Score:                  90.00        Matches:       46
Percent Similarity:     42.1%        Conservative:  34
Best Local Similarity:  24.2%        Mismatches:    78
Query Match:            7.6%         Indels:        32
DB:                     9            Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-959-243-3 (1-108359)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||  |||   :::   |||::::::          |||
Db      78182 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 78238

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::||||||                  :::||||||      |||
Db      78239 TCAGTGTGCTGTATTCAGGAA-----------------ACCTATCTCACGTGCAGAGAC 78280

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       :::   |||                  |||
Db      78281 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 78340

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::  |||  :::        :::          |||       |||||:::
Db      78341 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 78400

Qy         86 Arg-----------AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCys 101
              |||            :::   |||   |||      ::::::          ||||||
Db      78401 AGAGACAAAGAAGGCCATTACCTAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 78460

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||           |||   |||          :::
Db      78461 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 78505

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||     :::|||         :::|||   |||:::|||       |||
Db      78506 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 78562

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db      78563 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 78622

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::                 :::::::::|||::::::      |||
```

```
Db      78623 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 78682

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                  ||||||      :::|||||||:::
Db      78683 ---CACCCCAAATCAACAGAATATACATTT 78709

RESULT 74
US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
;   LENGTH: 3309400
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Alignment Scores:
Pred. No.:              1.09e+05        Length:         3309400
Score:                  89.50           Matches:        52
Percent Similarity:     38.9%           Conservative:   30
Best Local Similarity:  24.6%           Mismatches:     80
Query Match:            7.5%            Indels:         49
DB:                     3               Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-738-626-1 (1-3309400)

Qy         41 LeuLysLysArg---------SerSerAsnTyrGluIleIleThrGlyHis--------- 54
              ||||||||||||         |||         :::   |||   ||||||
Db    1363037 CTTAAAAAACGGGCGCGCAAAAGCGATACGCTTCAGCACATTGGAGGCCATTTGCCGCGA 1362978

Qy         55 ---------GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 71
                      |||         |||:::             ::::::
Db    1362977 GCTCGGCTGCCAGCCGGGCGACATTTTGCGTTACGACGCCTCCCTCCACAACTAAACCCG 1362918

Qy         72 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysVal 91
              |||   :::|||                 :::::::::      |||      :::   :::
Db    1362917 AAACGCAAAGAGCCCACCTCCCTCAGAAATTCTGAACGGAGACGGGCTTTTGTTCATCTA 1362858

Qy         92 HisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr--- 110
                 |||   :::|||   :::|||      |||:::||||||:::|||::::::
Db    1362857 CAGGTTTTACTCAGCGACCACGAATGGGGTCTGCTCACGAGTAACCTTCAGACGAACCAA 1362798

Qy        111 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 130
              ||||||      |||:::::::::::   ||||||
Db    1362797 CGTGGATTCCTTGCCCAAAAGCTCCATGGACTCAAA---------------------- 1362762

Qy        131 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu--- 149
                  :::                 ||||||||||||||   :::      |||
Db    1362761 ------CAGTGGTGGGGATACAGCCTCGCCGGAGATACCAATGCGCAACGCACCGAACGC 1362708
```

```
Qy        150 ----------------------------------ValThrArgCysGlyGlyLeuPro 157
                                                :::   :::|||
Db    1362707 CACGCGAGGCTTCAGGCCTAGATCCTCAATGAGAGCCTTGTTCAATGC---------- 1362660

Qy        158 AsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
                      :::||||||          |||          ||||||
Db    1362659 ------TGCTTCAATGTTTGGAGT------GGTCCACTCCTCCACTGCCTCCAGCGCTGC 1362612

Qy        178 AspThr---------GlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArg 194
              ||||||          ::::::||||||   |||                      |||
Db    1362611 GATACCGGCGTTGAGAGGCTCAACAGCGGTCTCCTTGAGGTTCTTCTTGGCAGCCTTCTC 1362552

Qy        195 PheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeu 214
                 :::      |||   :::      :::|||          |||||||||
Db    1362551 ATTGAACACCAGATCTTCGTCAGCGGTGACCAGGAACTTCAGCAGATCCCACGCCTC--- 1362495

Qy        215 AspLeuLeuGlyLeuGluLysLeuAspCysGly 225
                        |||   ||||||   |||
Db    1362494 ------------GCTGAGCACCTTGATGCGGGT 1362474

RESULT 75
US-10-389-647-119
; Sequence 119, Application US/10389647
; Publication No. US20040033549A1
; GENERAL INFORMATION:
;  APPLICANT: GREENBERG, E. Peter
;  APPLICANT:  SCHUSTER, Martin
;  APPLICANT:  LOSTROH, Candi
;  TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA
;  FILE REFERENCE: UIZ-038CP
;  CURRENT APPLICATION NUMBER: US/10/389,647
;  CURRENT FILING DATE:  2003-03-14
;  PRIOR APPLICATION NUMBER: 09/653730
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/153022
;  PRIOR FILING DATE: 1999-09-03
;  NUMBER OF SEQ ID NOS: 710
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119
;   LENGTH: 738
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-10-389-647-119

Alignment Scores:
Pred. No.:              0.221           Length:         738
Score:                  89.00           Matches:        47
Percent Similarity:     34.6%           Conservative:   19
Best Local Similarity:  24.6%           Mismatches:     51
Query Match:            7.5%            Indels:         74
DB:                     7               Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-10-389-647-119 (1-738)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              |||      |||   |||      |||   |||
Db         76 CTGCCGGAGTTGCGCGAAGCGGTGCGCAGCGTC----------------------GGTGCC 114

Qy         26 AspValIlePheLeuGlnGluValIle------------------------------- 34
              |||::::::|||||||||||||||:::
Db        115 GACCTGGTGTTCCTCCAGGAGGTCCTCGGCTCCCATAGCCTGCACGCCGCGCGCCTGCCG 174

Qy         35 ------------ProProTyrTyrSerTyrLeuLysLys---------------------- 43
                         |||   |||   ||||||
Db        175 TCCTGGCCGCCGGCGCCGCAATACGAGTATCTGGCCGACAGCATGTGGCCGCAGTTCGCC 234

Qy         44 ------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                    |||::::::   |||                      ||||||:::
Db        235 TACGGGCGCAACGCGGTGTAT-----------------CCCGAAGGCCACCATGGCAAT 276

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                 :::|||   |||      :::
Db        277 GCCGTGCTCTCCAAGCATCCGATCCTCGCCCAC------------------------- 309
```

```
Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu------ 99
                ||||||||           |||:::|||:::||||||||
Db        310 ------------CGCAACCTC---------GACGTCTCGGTGGCGGGCAACGAGGAACGC 348

Qy        100 -----------------------------------LeuCysLeuMetThrSerHisLeu 107
                                                  |||      :::      ||||||
Db        349 GGCCTGTTGCATGCGGTGATCGACATCGGCCGTCCGCTGCACGCGGTCTGCGTGCACCTG 408

Qy        108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
                           |||:::|||      ||||||:::::::::   :::
Db        409 GGATTG---------CGCGAGGCGCAGCGCCAGCGCCAGTTGCAACTGCTCTGCGAGCTG 459

Qy        128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn--------- 144
                :::        ||||||         |||:::   |||||||||   |||
Db        460 GTGGCGGGCCTGCCGGAGCGCGAGCCGGTGGTGGTCGCCGGCGACTTCAACGACTGGCGC 519

Qy        145 ---LeuArgAspArgGluValThrArgCysGly 154
                  ||||||   :::      ||||||
Db        520 CAGCGCGCCGACCGCCTGCTCGCTGGTTGCGGC 552


Search completed: March 11, 2006, 05:42:24
Job time : 2087.57 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 20:06:44 ; Search time 258.151 Seconds
(without alignments)
1563.065 Million cell updates/sec

Title: US-10-757-745-2_COPY_121_347
Perfect score: 1192
Sequence: 1 IDGLDLNNLSERARGVCSYL.....HIIPRSLDLLGLEKLDCGRF 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170331_4111/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_417_@runat_09032006_170331_4111 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1192	100.0	1920	3	US-09-697-863A-1	Sequence 1, Appli
	2	1192	100.0	1948	3	US-09-620-312D-889	Sequence 889, App
c	3	1090	91.4	1079	3	US-09-118-554-63	Sequence 63, Appl
c	4	1090	91.4	1079	3	US-09-118-627-63	Sequence 63, Appl
c	5	1090	91.4	1079	3	US-09-602-877A-63	Sequence 63, Appl
	6	945	79.3	1312	3	US-09-697-863A-3	Sequence 3, Appli
c	7	134	11.2	268	3	US-09-397-787-312	Sequence 312, App
	8	97	8.1	1830121	3	US-09-557-884-1	Sequence 1, Appli
	9	97	8.1	1830121	3	US-09-643-990A-1	Sequence 1, Appli
	10	97	8.1	1830121	3	US-10-158-865-1	Sequence 1, Appli
c	11	95.5	8.0	232547	3	US-09-949-016-16603	Sequence 16603, A
	12	94	7.9	3404	3	US-09-949-016-4002	Sequence 4002, Ap
	13	92.5	7.8	1452	3	US-09-248-796A-4120	Sequence 4120, Ap
c	14	92.5	7.8	421491	3	US-09-949-016-12805	Sequence 12805, A
c	15	92.5	7.8	421494	3	US-09-949-016-14060	Sequence 14060, A
	16	92	7.7	810	3	US-09-489-039A-5801	Sequence 5801, Ap
c	17	89	7.5	1830121	3	US-09-557-884-1	Sequence 1, Appli
c	18	89	7.5	1830121	3	US-09-643-990A-1	Sequence 1, Appli
c	19	89	7.5	1830121	3	US-10-158-865-1	Sequence 1, Appli
	20	88.5	7.4	690	3	US-09-107-532A-2664	Sequence 2664, Ap
c	21	87	7.3	106380	3	US-09-949-016-17553	Sequence 17553, A
c	22	86	7.2	86439	3	US-09-949-016-11945	Sequence 11945, A
c	23	86	7.2	86440	3	US-09-949-016-16990	Sequence 16990, A
c	24	86	7.2	133063	3	US-09-949-016-16293	Sequence 16293, A
c	25	86	7.2	818128	3	US-09-949-016-14546	Sequence 14546, A
c	26	86	7.2	818128	3	US-09-949-016-14547	Sequence 14547, A
c	27	86	7.2	818128	3	US-09-949-016-14548	Sequence 14548, A
c	28	86	7.2	818128	3	US-09-949-016-14549	Sequence 14549, A
c	29	86	7.2	818128	3	US-09-949-016-14550	Sequence 14550, A
c	30	86	7.2	818128	3	US-09-949-016-14551	Sequence 14551, A
c	31	86	7.2	818128	3	US-09-949-016-14552	Sequence 14552, A
c	32	86	7.2	818128	3	US-09-949-016-14553	Sequence 14553, A
c	33	86	7.2	818128	3	US-09-949-016-14554	Sequence 14554, A
c	34	86	7.2	818128	3	US-09-949-016-14555	Sequence 14555, A
c	35	86	7.2	818128	3	US-09-949-016-14556	Sequence 14556, A
c	36	86	7.2	818128	3	US-09-949-016-14557	Sequence 14557, A
c	37	86	7.2	818128	3	US-09-949-016-14558	Sequence 14558, A
c	38	86	7.2	818128	3	US-09-949-016-14559	Sequence 14559, A
c	39	86	7.2	818128	3	US-09-949-016-14560	Sequence 14560, A
c	40	86	7.2	818128	3	US-09-949-016-14561	Sequence 14561, A
c	41	86	7.2	818128	3	US-09-949-016-14562	Sequence 14562, A
c	42	86	7.2	818128	3	US-09-949-016-14564	Sequence 14564, A
c	43	86	7.2	818128	3	US-09-949-016-14565	Sequence 14565, A
c	44	86	7.2	818128	3	US-09-949-016-14566	Sequence 14566, A
c	45	86	7.2	818128	3	US-09-949-016-14567	Sequence 14567, A
c	46	85.5	7.2	1575	3	US-09-774-528-260	Sequence 260, App
c	47	85.5	7.2	1575	3	US-10-120-988-260	Sequence 260, App
	48	85	7.1	756	3	US-09-252-991A-8106	Sequence 8106, Ap
c	49	85	7.1	873	3	US-09-252-991A-8199	Sequence 8199, Ap
c	50	84.5	7.1	410	3	US-08-735-545-20	Sequence 20, Appl
c	51	84.5	7.1	410	3	US-09-449-083-20	Sequence 20, Appl
	52	84.5	7.1	146401	3	US-09-949-016-16151	Sequence 16151, A
	53	84.5	7.1	331814	3	US-09-949-016-12008	Sequence 12008, A
	54	84.5	7.1	331814	3	US-09-949-016-17056	Sequence 17056, A
	55	84	7.0	804	3	US-09-134-000C-2365	Sequence 2365, Ap
c	56	84	7.0	17644	3	US-09-949-016-16342	Sequence 16342, A
c	57	84	7.0	17645	3	US-09-949-016-11805	Sequence 11805, A
c	58	84	7.0	17645	3	US-09-949-016-14188	Sequence 14188, A
c	59	84	7.0	32042	3	US-09-245-281-44	Sequence 44, Appl
c	60	84	7.0	32042	3	US-09-340-620A-63	Sequence 63, Appl
c	61	84	7.0	32042	3	US-09-728-721-63	Sequence 63, Appl
c	62	84	7.0	39553	3	US-09-949-002-810	Sequence 810, App
c	63	84	7.0	58111	3	US-09-949-002-673	Sequence 673, App
c	64	84	7.0	87567	3	US-09-949-016-13335	Sequence 13335, A
c	65	84	7.0	102520	3	US-09-949-016-17367	Sequence 17367, A
c	66	84	7.0	102526	3	US-09-949-016-12448	Sequence 12448, A
c	67	84	7.0	175236	3	US-09-949-016-14353	Sequence 14353, A
c	68	84	7.0	206433	3	US-09-949-016-13527	Sequence 13527, A
c	69	84	7.0	254778	3	US-09-949-016-12417	Sequence 12417, A
c	70	84	7.0	294836	3	US-09-949-016-15974	Sequence 15974, A
	71	83.5	7.0	840	3	US-09-252-991A-14967	Sequence 14967, A
c	72	83.5	7.0	1149	3	US-09-252-991A-14783	Sequence 14783, A
	73	83.5	7.0	1461	3	US-09-252-991A-14833	Sequence 14833, A
	74	83.5	7.0	72843	3	US-09-949-016-12574	Sequence 12574, A
c	75	83.5	7.0	131631	3	US-09-949-016-11757	Sequence 11757, A

ALIGNMENTS

RESULT 1
US-09-697-863A-1
; Sequence 1, Application US/09697863A

```
; Patent No. 6812203
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/09/697,863A
;  CURRENT FILING DATE:  2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1627)..(1627)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1108)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1849)..(1849)
;   OTHER INFORMATION: N stands for any nucleotide.
US-09-697-863A-1

Alignment Scores:
Pred. No.:                2.58e-139       Length:          1920
Score:                    1192.00         Matches:         227
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-697-863A-1 (1-1920)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 439

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 499

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 559

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 619

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 679

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 739

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 799

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 859

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 919

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 979

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039

Qy        221 LysLeuAspCysGlyArgPhe 227
              |||||||||||||||||||||
Db       1040 AAACTGGACTGTGGTAGATTT 1060

RESULT 2
US-09-620-312D-889
; Sequence 889, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-09-620-312D-889

Alignment Scores:
Pred. No.:                2.64e-139       Length:          1948
Score:                    1192.00         Matches:         227
Percent Similarity:       100.0%          Conservative:    0
Best Local Similarity:    100.0%          Mismatches:      0
Query Match:              100.0%          Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-620-312D-889 (1-1948)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 458

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
```

```
Db        459 GCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 518

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60

Db        519 CTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 578

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80

Db        579 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 638

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100

Db        639 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 698

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120

Db        699 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 758

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140

Db        759 TTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 818

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160

Db        819 GGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATT 878

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180

Db        879 GTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAA 938

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200

Db        939 ATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTC 998

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220

Db        999 AGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAA 1058

Qy        221 LysLeuAspCysGlyArgPhe 227

Db       1059 AAACTGGACTGTGGTAGATTT 1079

RESULT 3
US-09-118-554-63/c
; Sequence 63, Application US/09118554A
; Patent No. 6365348
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND
;  TITLE OF INVENTION: METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.450C1
;  CURRENT APPLICATION NUMBER: US/09/118,554A
;  CURRENT FILING DATE: 1998-07-17
;  EARLIER APPLICATION NUMBER: 08/988,255
;  EARLIER FILING DATE: 1997-12-24
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-118-554-63

Alignment Scores:
Pred. No.:              6.91e-127      Length:          1079
Score:                  1090.00        Matches:         211
Percent Similarity:     99.5%          Conservative:    0
Best Local Similarity:  99.5%          Mismatches:      1
Query Match:            91.4%          Indels:          1
DB:                     3              Gaps:            0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-118-554-63 (1-1079)

Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35

Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55

Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75

Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95

Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115

Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135

Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155

Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175

Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195

Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215

Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227

Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443

RESULT 4
US-09-118-627-63/c
; Sequence 63, Application US/09118627A
; Patent No. 6379951
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER
;  TITLE OF INVENTION: AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.446C1
;  CURRENT APPLICATION NUMBER: US/09/118,627A
;  CURRENT FILING DATE: 1998-07-17
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-118-627-63

Alignment Scores:
Pred. No.:              6.91e-127      Length:          1079
Score:                  1090.00        Matches:         211
Percent Similarity:     99.5%          Conservative:    0
Best Local Similarity:  99.5%          Mismatches:      1
Query Match:            91.4%          Indels:          1
DB:                     3              Gaps:            0
```

US-10-757-745-2_COPY_121_347 (1-227) x US-09-118-627-63 (1-1079)

```
Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443
```

RESULT 5
US-09-602-877A-63/c
; Sequence 63, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-63

Alignment Scores:

Pred. No.:	6.91e-127	Length:	1079
Score:	1090.00	Matches:	211
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	91.4%	Indels:	1
DB:	3	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-602-877A-63 (1-1079)

```
Qy         16 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 35
Db       1077 GTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018

Qy         36 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 55
Db       1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958

Qy         56 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 75
Db        957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAG 898

Qy         76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
Db        897 ATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTG 838

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
Db        837 TCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCG 778

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
Db        777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
Db        717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGT 658

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175
Db        657 TTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTAT 598

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
Db        597 ACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTT 538

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
Db        537 GATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-GAC 479

Qy        216 LeuLeuGlyLeuGluLysLeuAspCysGlyArgPhe 227
Db        478 CTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTT 443
```

RESULT 6
US-09-697-863A-3
; Sequence 3, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/09/697,863A
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (122)..(1234)
; OTHER INFORMATION:

US-09-697-863A-3

Alignment Scores:

Pred. No.:	1.47e-108	Length:	1312
Score:	945.00	Matches:	179
Percent Similarity:	89.0%	Conservative:	23
Best Local Similarity:	78.9%	Mismatches:	23
Query Match:	79.3%	Indels:	2
DB:	3	Gaps:	1

US-10-757-745-2_COPY_121_347 (1-227) x US-09-697-863A-3 (1-1312)

```
Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ||||||||||||||||      ||||||   ||||||||||||||||||||||||   |||
Db        512 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 571

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
              ||||||||||||||||||||:::|||||||||||||||||||||||||||   :::|||
Db        572 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTAC 631

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              ||||||||||||:::::::::|||   ||||||||||:::||||||||||||||||||||
Db        632 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 691

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
              ||||||:::||||||||   |||||||||   ||||||||||||||||||||||||||||
Db        692 GCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 751

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeu 100
              :::|||||||||||||||||||||||||||||:::|||::::::   |||||||||
Db        752 AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 811

Qy        101 CysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGln 120
              ||||||||||||||||||||||||||||||||   |||:::|||||||||:::   |||
Db        812 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 871

Qy        121 LeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAla 140
              ||||||   ||||||  ||||||||||||||||||:::|||   ||||||||||||||||
Db        872 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 931

Qy        141 GlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIle 160
              ||||||||||||||||||||:::||||||   :::|||||||||||||||:::|||:::
Db        932 GGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTT 991

Qy        161 ValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGln 180
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        992 TTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAA 1051

Qy        181 MetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe 200
              |||:::||||||   |||  ||||||  |||  ||||||||||||||||||||
Db       1052 GCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTC 1111

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              |||      ||||||||||||||:::||||||:::||||||||||||:::||||||||||
Db       1112 AGA------GCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 1165

Qy        221 LysLeuAspCysGlyArgPhe 227
              ||||||||||||||||||||
Db       1166 AAACTGGACTGTGGTAGATTT 1186
```

RESULT 7
US-09-397-787-312/c
; Sequence 312, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-312

Alignment Scores:

Pred. No.:	8.4e-08	Length:	268
Score:	134.00	Matches:	26
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	11.2%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_121_347 (1-227) x US-09-397-787-312 (1-268)

```
Qy        202 AlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGluLys 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAA 208

Qy        222 LeuAspCysGlyArgPhe 227
              ||||||||||||||||||
Db        207 CTGGACTGTGGTAGATTT 190
```

RESULT 8
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Alignment Scores:

```
Pred. No.:              2.04e+03        Length:         1830121
Score:                  97.00           Matches:        51
Percent Similarity:     41.0%           Conservative:   43
Best Local Similarity:  22.3%           Mismatches:     89
Query Match:            8.1%            Indels:         46
DB:                     3               Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-09-557-884-1 (1-1830121)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||   ::::::      ::::::|||:::|||      :::   |||   |||
Db    1644110 ACGCCAGTGCTTGTTGATGGCAAAGATGTTATGCCTGAAGTCAATGCTGTGCTAGCTAAA 1644169

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                 :::              ||||||:::|||     :::|||||
Db    1644170 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGTAAAGCC 1644229

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::              :::   :::        ::::::   |||:::      |||
Db    1644230 ATTACGGATGTTGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 1644289

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
              :::   :::|||         |||   |||   :::|||         ||||||   |||
Db    1644290 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 1644349

Qy         98 AsnGlu-----------------------------------------LeuCysLeuMetThr 104
                                                           |||   |||:::
Db    1644350 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 1644409

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              |||        :::|||:::                 :::                 :::
Db    1644410 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCTCAAAGTGCGCGTGATTGGTTA 1644469

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||      ::::::           |||           |||       ||||||
Db    1644470 CTAAAAGCAGCGAAAGATGAAAGTGCAGTTGCAAAACATTTTGCAGCATTATCAACCAAT 1644529

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              :::|||      |||   :::   |||         ||||||:::   :::   |||
Db    1644530 GCTAAAGAT------GTAGAAAAATTTGGCATTGATACCAATAATATGTTTGAATTTTGG 1644583

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
              :::::::::|||          :::|||:::  |||              |||
Db    1644584 GATTGGGTTGGCGGT-----------CGTTACTCTTTATGG--------------TCA 1644616

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe---------------------- 195
              :::|||::::::   |||      :::   |||
Db    1644617 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAACTTTGAAGCGTTATTAAAT 1644676

Qy        196 --------------AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210
                           |||:::   |||           |||:::          |||
Db    1644677 GGTGCGCATGAAATGGATAAACATTTCCTCTCTACTCCAATCGAACAAAAT------ATC 1644730

Qy        211 ProArgSerLeuAspLeuGluGlyLeu 219
              |||   :::|||   |||:::||||||
Db    1644731 CCAACCACTTTAGCATTAGTTGGTTTA 1644757

RESULT 9
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
;   GENERAL INFORMATION:
;        APPLICANT: Robert D. Fleischmann
;                   Mark D. Adams
;                   Owen White
;                   Hamilton O. Smith
;                   J. Craig Venter
;        TITLE OF INVENTION: The Nucleotide sequence of
;                            the Haemophilus influenzae Rd Genome, Fragments
;                            Thereof, and Uses Thereof
;        NUMBER OF SEQUENCES: 1
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville,
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3 1/2 inch diskette
;             COMPUTER: Dell Pentium
;             OPERATING SYSTEM: MS DOS v6.22
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/643,990A
;             FILING DATE: 23-Aug-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/487,429
;             FILING DATE: 1995-06-07
;             APPLICATION NUMBER: 08/426,787
;             FILING DATE: 1995-04-21
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB186P1C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-610-5790
;             TELEFAX: 310-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1830121 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Alignment Scores:
Pred. No.:              2.04e+03        Length:         1830121
Score:                  97.00           Matches:        51
Percent Similarity:     41.0%           Conservative:   43
Best Local Similarity:  22.3%           Mismatches:     89
Query Match:            8.1%            Indels:         46
DB:                     3               Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-09-643-990A-1 (1-1830121)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||   ::::::      ::::::|||:::|||      :::   |||   |||
Db    1644110 ACGCCAGTGCTTGTTGATGGCAAAGATGTTATGCCTGAAGTCAATGCTGTGCTAGCTAAA 1644169

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                 :::              ||||||:::|||     :::|||||
Db    1644170 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGTAAAGCC 1644229

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::              :::   :::        ::::::   |||:::      |||
Db    1644230 ATTACGGATGTTGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 1644289

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
              :::   :::|||         |||   |||   :::|||         ||||||   |||
Db    1644290 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 1644349

Qy         98 AsnGlu-----------------------------------------LeuCysLeuMetThr 104
                                                           |||   |||:::
Db    1644350 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 1644409

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              |||        :::|||:::                 :::                 :::
Db    1644410 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCTCAAAGTGCGCGTGATTGGTTA 1644469

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||      ::::::           |||           |||       ||||||
Db    1644470 CTAAAAGCAGCGAAAGATGAAAGTGCAGTTGCAAAACATTTTGCAGCATTATCAACCAAT 1644529
```

```
Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              :::|||      |||   :::   |||         ||||||:::   :::    |||
Db    1644530 GCTAAAGAT------GTAGAAAAATTTGGCATTGATACCAATAATATGTTTGAATTTTGG 1644583

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
              :::::::::|||                :::|||:::   |||               |||
Db    1644584 GATTGGGTTGGCGGT-----------CGTTACTCTTTATGG--------------TCA 1644616

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe----------------------- 195
              :::|||::::::   |||      :::   |||
Db    1644617 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAACTTTGAAGCGTTATTAAAT 1644676

Qy        196 --------------AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210
                           |||:::   |||               |||:::       |||
Db    1644677 GGTGCGCATGAAATGGATAAACATTTCCTCTCTACTCCAATCGAACAAAAT------ATC 1644730

Qy        211 ProArgSerLeuAspLeuLeuGlyLeu 219
              |||   :::|||   |||:::||||||
Db    1644731 CCAACCACTTTAGCATTAGTTGGTTTA 1644757

RESULT 10
US-10-158-865-1
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
;  TITLE OF INVENTION:  Thereof, and Uses Thereof
;  FILE REFERENCE: PB186P2C1D1
;  CURRENT APPLICATION NUMBER: US/10/158,865
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US 09/557,884
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US 08/476,102
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/426,787
;  PRIOR FILING DATE: 1995-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1830121
;   TYPE: DNA
;   ORGANISM: Haemophilus influenzae
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4747)..(4747)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9921)..(9921)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10150)..(10150)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (29298)..(29298)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36543)..(36543)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36551)..(36551)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36636)..(36636)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (40808)..(40810)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44416)..(44416)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44905)..(44905)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44975)..(44975)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45593)..(45593)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45732)..(45732)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (47036)..(47036)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51334)..(51334)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51602)..(51602)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51786)..(51786)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51805)..(51805)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (55369)..(55369)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65309)..(65309)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65313)..(65313)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (80024)..(80024)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (100091)..(100091)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (102696)..(102696)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (105121)..(105121)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (107248)..(107248)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (117136)..(117136)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (119750)..(119750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (119924)..(119924)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (120038)..(120038)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (121344)..(121344)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (122167)..(122167)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (122336)..(122336)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131340)..(131340)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131360)..(131360)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (139910)..(139910)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (140398)..(140398)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (142750)..(142750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145058)..(145058)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145171)..(145171)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145942)..(145942)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (147197)..(147197)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (150841)..(150841)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152500)..(152500)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature

Alignment Scores:
Pred. No.:              2.04e+03        Length:         1830121
Score:                  97.00           Matches:        51
Percent Similarity:     41.0%           Conservative:   43
Best Local Similarity:  22.3%           Mismatches:     89
Query Match:            8.1%            Indels:         46
DB:                     3               Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-10-158-865-1 (1-1830121)

Qy         24 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu---Lys 42
              :::|||   :::::::      :::::::|||:::|||         :::   |||   ||| 
Db    1644110 ACGCCAGTGCTTGTTGATGGCAAAGATGTTATGCCTGAAGTCAATGCTGTGCTAGCTAAA 1644169

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
              :::                 |||||||:::|||       :::|||||||          
Db    1644170 ATGAAAGATTTCTGTCAGCGTATTATTTCTGGCGAATGGAAAGGCTATACAGGTAAAGCC 1644229

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::          :::  :::          ::::::   |||:::          |||  
Db    1644230 ATTACGGATGTTGTGAATATTGGTATTGGTGGCTCTGACTTAGGCCCTTATATGGTAACC 1644289

Qy         83 LysMetMetArg---------AsnLeuLeuCysValHisVal------AsnValSerGly 97
              :::   :::|||         |||   |||   :::|||         ||||||   |||  
Db    1644290 GAAGCACTTCGCCCGTATAAAAATCATCTAAATATGCACTTTGTTTCAAATGTCGATGGT 1644349

Qy         98 AsnGlu-------------------------------------LeuCysLeuMetThr 104
                                                        |||   |||:::      
Db    1644350 ACACATATTGCGGAAACCTTAAAAAAAGTCAATCCAGAAACGACTCTTTTCTTAGTAGCA 1644409

Qy        105 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetVal 124
              |||       :::|||:::                 :::             :::      
Db    1644410 TCGAAAACTTTTACAACTCAAGAAACCATGACAAATGCTCAAAGTGCGCGTGATTGGTTA 1644469

Qy        125 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 144
              ||||||      ::::::            |||           |||      ||||||   
Db    1644470 CTAAAAGCAGCGAAAGATGAAAGTGCAGTTGCAAAACATTTTGCAGCATTATCAACCAAT 1644529

Qy        145 LeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrp 164
              :::|||      |||   :::   |||         ||||||:::   :::   |||  
Db    1644530 GCTAAAGAT------GTAGAAAAATTTGGCATTGATACCAATAATATGTTTGAATTTTGG 1644583

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThr---TrpAspThrGlnMetAsnSer 183
              :::::::::|||           :::|||:::   |||               |||    
Db    1644584 GATTGGGTTGGCGGT-----------CGTTACTCTTTATGG---------------TCA 1644616

Qy        184 AsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe----------------------- 195
              :::|||::::::   |||       :::   |||                          
Db    1644617 GCTATTGGTCTTTCAATTGCACTATCAATTGGCTTTGAAAACTTTGAAGCGTTATTAAAT 1644676

Qy        196 --------------AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIle 210
                           |||:::   |||           |||:::          |||    
Db    1644677 GGTGCGCATGAAATGGATAAACATTTCCTCTCTACTCCAATCGAACAAAAT------ATC 1644730

Qy        211 ProArgSerLeuAspLeuLeuGlyLeu 219
              |||   :::|||   |||:::|||||| 
Db    1644731 CCAACCACTTTAGCATTAGTTGGTTTA 1644757

RESULT 11
US-09-949-016-16603/c
; Sequence 16603, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
```

```
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16603
;   LENGTH: 232547
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16603

Alignment Scores:
Pred. No.:                144             Length:         232547
Score:                    95.50           Matches:        44
Percent Similarity:       43.3%           Conservative:   40
Best Local Similarity:    22.7%           Mismatches:     77
Query Match:              8.0%            Indels:         33
DB:                       3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-16603 (1-232547)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||:::::::|||   |||:::         :::       :::   :::::::::
Db      71923 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGAAACAGACTGGCAAATTGGATAAAG 71864

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                    |||   |||      :::||||||                  ::::::|||
Db      71863 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACCCATCTC 71822

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                  |||   ::::::         |||   |||       :::   |||
Db      71821 ACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 71762

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                  ||||||:::   |||   :::       :::       |||
Db      71761 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 71702

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
              ||||||:::   |||:::      |||   ::::::   |||   |||:::
Db      71701 ACAAAGATCAAAAGAGACAAAGAATGCCATTACATAATGGTAAAGGGATCAATTCAACAA 71642

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                 ||||||   ::::::      :::      :::|||          |||   |||
Db      71641 GAAGAGCTAATTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 71591

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
              :::      :::|||   ||||||      :::|||       :::|||   |||:::
Db      71590 TTG------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACAATA 71537

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db      71536 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAA 71480

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                   :::|||   :::          :::::::::|||:::
Db      71479 GTTAATAAGGATACCCAGGAATTGAACTCAGCTCTGAACCAAGCAGACCTAATAGATATC 71420

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||     :::||||||:::
Db      71419 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 71381

RESULT 12
US-09-949-016-4002
; Sequence 4002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4002
;   LENGTH: 3404
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-4002

Alignment Scores:
Pred. No.:                0.396           Length:         3404
Score:                    94.00           Matches:        63
Percent Similarity:       37.7%           Conservative:   37
Best Local Similarity:    23.8%           Mismatches:     96
Query Match:              7.9%            Indels:         70
DB:                       3               Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-4002 (1-3404)

Qy         13 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 32
              |||:::|||      |||   :::|||                  :::||||||
Db       569 GCTCAGGGCCCAGCGAGCATGATGGCCTACTTGGTTGTTTCTGGGCTATTCCTAACTCGA 628

Qy         33 ValIleProProTyr-------------------------------------TyrSer 39
              :::      |||                                        |||
Db       629 CTTCGAAGACCCATTGGTAAGATGACAATAACTGAGCAAAAGTATGAAGGAGAATATAGA 688

Qy         40 TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-Ph 59
              |||:::      |||                :::||||||      ||||||      ||
Db       689 TATGTTAATTCTCGG--------------CTCATCACAAACAGTGAAGAAATTGCCTTT 733

Qy         59 eThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu---------- 75
              ||||      :::   :::||||||||          ::|||   |||
Db       734 TACAATGGGAATAAAAGAGAAAAGCAGACAGTCCACTCAGTCTTCCGAAAACTGGTGGAA 793

Qy         76 ------------------------------------------IleIleProPhePr 80
                                                       |||:::      ||
Db       794 CACCTACATAATTTCATTTTGTTTCGGTTTTCAATGGGCTTCATTGATAGTATTATTGCC 853

Qy         80 oSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLe 100
              |:::|||      :::      ||||||          ::::::|||:::   :::   ::
Db       854 AAATACCTTGCCACTG------TTGTTGGTTACCTAGTTGTCAGTCGCCCTTTCTTAGAT 907

Qy        100 uCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGl 120
              :||||||:::      ||||||:::||||||                |||       ||
Db       908 TTGTCTCATCCT-CGACATCTCAAGAGTACACATTCGGAACTTCTAGAGGATTACTACCA 966

Qy        120 nLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAl 140
              |          :::|||   :::|||   :::|||                ::::::   ||
Db       967 AAGTGGAAGAATGCTTTTGCGAATGTCTCAAGCT------CTGGGTCGAATAGTTTTGGC 1020

Qy        140 aGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIl 160
              ||||                ||||||:::||||||      |||           ||
Db      1021 TGGG-----------------CGTGAAATGACTAGATTGGCCGGTTTTACTGCTCGGAT 1062

Qy        160 eValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr------ThrTrpAs 178
              |   ::::::   :::   |||          |||   :::|||      |||
Db      1063 TACAGAATTAATGCAAGTACTGAAGGATTTAAATCATGGCAAATATGAGCGCACAATGGT 1122

Qy        178 pThrGlnMetAsnSerAsnLeuGly---------------------------------- 186
               :::|||          :::
Db      1123 CTCACAACAGGAAAAGGGTATTGAAGGAGTACAAGTCATTCCCTTGATACCTGGTGCTGG 1182

Qy        187 -------IleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                    |||   |||        ::::::||||||   :::         |||
Db      1183 AGAAATCATTATTGCAGATAACATTATAAAGTTTGATCATGTT------CCTTTAGCAAC 1236
```

```
Qy        204 aGluGluGlyHisIleIleProArgSerLeuAspLeu---------LeuGlyLeuGluLy 221
Db       1237 GCCAAATGGAGATGTTTTGATCCGAGACCTTAATTTTGAAGTTCGATCTGGGGCTAATGT 1296

Qy        221 sLeuAspCysGly 225
Db       1297 TCTAATTTGTGGT 1309

RESULT 13
US-09-248-796A-4120
; Sequence 4120, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4120
;   LENGTH: 1452
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-4120

Alignment Scores:
Pred. No.:                0.171          Length:          1452
Score:                    92.50          Matches:         49
Percent Similarity:       38.1%          Conservative:    42
Best Local Similarity:    20.5%          Mismatches:      75
Query Match:              7.8%           Indels:          73
DB:                       3              Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-248-796A-4120 (1-1452)

Qy         18 SerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyr 37
Db        103 AGTTACTCATCATATTTAACACCCACTTTAATTTTACTTATTAGGTATATG------ATA 156

Qy         38 TyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGly 57
Db        157 AAGACATTGCTCCGAAGAATGTCAAGTAATACCACTATACCAACCCCAAATGGCTCCAAT 216

Qy         58 TyrPheThrAlaIleMetLeuLysLysSerArgValLys---LeuLysSerGlnGluIle 76
Db        217 CATTGGACTGCCTCAAAAGTTAGATCTACATTTTTAGACTATTTCAAGAAACAGCAACAC 276

Qy         77 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 96
Db        277 ACTTATGTCCCATCTTCATCAGTTGTT------------------------------- 303

Qy         97 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
Db        304 ----------------------CCTCACAATGACCCAACTTTATTGTTTGCCAATGCT 339

Qy        117 ArgMetAsnGlnLeuLysMetVal---LeuLysLysMetGlnGluAlaProGluSerAla 135
Db        340 GGTATGAATCAATATAAACCAATTTTTTTGGGAACAGTTGATCCAGCTAGTGATTTTGCT 399

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
Db        400 AGTTTAAAAAGAGCTGCCAATTCT-----------CAAAAATGTATTAGAGCTGGTGGT 447

Qy        156 LeuProAsnAsnIleValAspVal-------------------------TrpGluPhe 166
Db        448 AAACATAACGATTTAGAAGATGTTGGTAGAGACTCTTATCATCATACATTTTTTGAAATG 507

Qy        167 LeuGly-----------------------LysProLysHisCysGlnTyrThrTrpAsp 178
Db        508 CTTGGTAATTGGTCATTTGGAGATTATTTCAAAAAAGAGGCCATTGACTATTCATGGGAA 567

Qy        179 ThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIle 198
Db        568 TTG--------------------TTGACCAAAGTTTATGGATTACAAGAAGATAGATTG 606

Qy        199 Phe------------------------------------------------------- 199
Db        607 TATGTCACTTATTTTGGTGGTGATGAAAAACAAGGATTAGAACCAGATTTAGAAGCCAAA 666

Qy        200 ---------PheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
Db        667 AATTTTTGGTTAAAGGTTGGTGTACCAGAGGATCATATTTTACCAGGTTCAGTTGAA 723

RESULT 14
US-09-949-016-12805/c
; Sequence 12805, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12805
;   LENGTH: 421491
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(421491)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12805

Alignment Scores:
Pred. No.:                833            Length:          421491
Score:                    92.50          Matches:         45
Percent Similarity:       45.4%          Conservative:    43
Best Local Similarity:    23.2%          Mismatches:      73
Query Match:              7.8%           Indels:          33
DB:                       3              Gaps:            10

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-12805 (1-421491)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
Db     289545 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 289486

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
Db     289485 AATCAAGACCCATCAGTGTGCTATATTCAGGAG----------------ACCCATCTC 289444

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
Db     289443 ACATGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAAATCTACCAAGCAAAT 289384

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
Db     289383 GGAAAACAAAAGAAAGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 289324

Qy         82 ThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnValSerGly 97
```

```
Db     289323 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACTTAATGGTAAAGGGATCAATTCAACAA 289264

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
              ||||||   ::::::      :::      :::||||||          |||   |||
Db     289263 GAAGAGCTAACTATCCTAAATATATATGCACACAATACAAGA---------GCACCCAGA 289213

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                 :::|||   ||||||:::   :::|||:::      :::|||   |||:::
Db     289212 TTC------ATAAAGCAAGTCCTTAGAGACCTACAAAAAGACTTAGATTCCCACACAATA 289159

Qy        138 IlePheAlaGlyAspThrAsn------------LeuArgAspArg--------------- 148
              |||      ||||||  |||          :::   ||||||
Db     289158 ATAATG---GGAGACTTTAACACCCCACTGTCAATATTAGACAGATCAACCAGACAGAAA 289102

Qy        149 ---------GluValThrArgCysGlyGlyLeuProAsn------AsnIleValAspVal 163
                      ::::::                  |||:::      :::::::::   :::
Db     289101 GTTAACAAGGATATCCAGGACTTAAACTCAGCTCCACACCAAGCAGACCTAATACACATC 289042

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||     :::||||||:::
Db     289041 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 289003

RESULT 15
US-09-949-016-14060/c
; Sequence 14060, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14060
;   LENGTH: 421494
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(421494)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14060

Alignment Scores:
Pred. No.:              833          Length:         421494
Score:                  92.50        Matches:        45
Percent Similarity:     45.4%        Conservative:   43
Best Local Similarity:  23.2%        Mismatches:     73
Query Match:            7.8%         Indels:         33
DB:                     3            Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14060 (1-421494)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db     289545 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 289486

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                  |||   |||   ::::::||||||                  ::::::|||
Db     289485 AATCAAGACCCATCAGTGTGCTATATTCAGGAG-----------------ACCCATCTC 289444

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::         |||   |||      :::   |||
Db     289443 ACATGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAAATCTACCAAGCAAAT 289384

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                 ||||||:::   |||   :::          :::          |||
Db     289383 GGAAAACAAAAGAAAGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 289324

Qy         82 ThrLysMetMetArg-----------AsnLeuLeuCysValHisValAsnValSerGly 97
              ||||||:::   |||            :::   |||   |||       ::::::
Db     289323 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACTTAATGGTAAAGGGATCAATTCAACAA 289264

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
              ||||||   ::::::      :::      :::||||||          |||   |||
Db     289263 GAAGAGCTAACTATCCTAAATATATATGCACACAATACAAGA---------GCACCCAGA 289213

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                 :::|||   ||||||:::   :::|||:::      :::|||   |||:::
Db     289212 TTC------ATAAAGCAAGTCCTTAGAGACCTACAAAAAGACTTAGATTCCCACACAATA 289159

Qy        138 IlePheAlaGlyAspThrAsn------------LeuArgAspArg--------------- 148
              |||      ||||||  |||          :::   ||||||
Db     289158 ATAATG---GGAGACTTTAACACCCCACTGTCAATATTAGACAGATCAACCAGACAGAAA 289102

Qy        149 ---------GluValThrArgCysGlyGlyLeuProAsn------AsnIleValAspVal 163
                      ::::::                  |||:::      :::::::::   :::
Db     289101 GTTAACAAGGATATCCAGGACTTAAACTCAGCTCCACACCAAGCAGACCTAATACACATC 289042

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||     :::||||||:::
Db     289041 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTC 289003

RESULT 16
US-09-489-039A-5801
; Sequence 5801, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 5801
;   LENGTH: 810
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5801

Alignment Scores:
Pred. No.:              0.0823       Length:         810
Score:                  92.00        Matches:        50
Percent Similarity:     32.0%        Conservative:   31
Best Local Similarity:  19.8%        Mismatches:     74
Query Match:            7.7%         Indels:         98
DB:                     3            Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-489-039A-5801 (1-810)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              |||      |||   :::      |||  |||                    |||
Db        136 TTGCCCGAACTGCGCGACGCCGTGCGCAGCGTT--------------------AGCGCG 174

Qy         26 AspValIlePheLeuGlnGluVal--------------IleProProTyrTyrSerTyr 40
              |||::::::  ||||||||||||            :::   |||   :::
Db        175 GATATTGTTTGCCTGCAGGAGGTGATGGGCGCTCATGAGGTACATCCGCTGCACATCGAA 234

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGly---------------------- 53
                        ::::::::::||||||   :::
Db        235 AACTGGCCGGATACGACCCATTATGAATTTCTCGCCGACACCATGTGGAGCGACTACGCC 294

Qy         54 ------------------HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 67
                           :::   ||||||:::          :::|||      |||
Db        295 TACGGGCGCAATGCGGTTTATCCGGAAGGCCACCACGGCAACGCCGTGCTG------TCG 348
```

```
Qy          68 ArgValLysLeuLysSerGlnGlu---------------------------------- 75
               |||      ::::::      |||
Db         349 CGCTTTCCCATTGAATACTACGAAAATCGCGATATCTCGGTCGGCAACGGGGAGAAGCGC 408

Qy          76 ------------------IleIleProPheProSerThrLysMetMetArgAsnLeuLeu 89
                               |||:::|||   |||   :::   :::       :::::::::
Db         409 GGCCTGCTGTACTGCCGCATCGTGCCG---CCACGGAGCGGGATAACTATTCACGTGATT 465

Qy          90 CysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSer 109
               |||||||||:::  :::
Db         466 TGCGTCCATCTCGGTCTA---------------------------------------- 483

Qy         110 ThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGln 129
                        |||   :::|||       ||||||   |||:::   :::   :::
Db         484 ---------CGCGCCGACCAACGGCAGGCGCAGTTGACCATGCTGGCGGAGTGGGTCAAT 534

Qy         130 GluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGlu 149
                  |||         |||:::   |||||||||   |||
Db         535 ACCCTGCCCGCGGGCGAGCCGGTGGTGGTGGCTGGCGATTTTAAC-------------- 579

Qy         150 ValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLys 169

Db         579 ---------------------------------------------------------- 579

Qy         170 ProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAla 189
                           |||      |||   |||      |||         |||
Db         580 ------------------GACTGGCGTCAGCAGGCTAACCAGCCGTTAAAAGCCCAGGCC 621

Qy         190 AlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIle 209
                        :::   ||||||   ||||||
Db         622 GGG------------CTGGAAGAGATCTTTACCCGCGCCCGCGGACGGCCGGCGCGCACC 669

Qy         210 IleProArgSerLeuAspLeuLeuGlyLeuGluLysLeu 222
                  |||   |||      ||||||   |||:::::::::
Db         670 TTTCCGGTCAGTTTTCCGCTGCTGCGCCTCGACCGCATT 708

RESULT 17
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
;   GENERAL INFORMATION:
;        APPLICANT: Fleischmann et al.
;        TITLE OF INVENTION: The Nucleotide sequence of
;                            the Haemophilus influenzae Rd Genome, Fragments
;                            Thereof, and Uses Thereof
;        NUMBER OF SEQUENCES: 1
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3 1/2 inch diskette
;             COMPUTER: Dell Pentium
;             OPERATING SYSTEM: MS DOS v6.22
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/557,884
;             FILING DATE: 25-Apr-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/476,102
;             FILING DATE: JUN-5-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Michelle S. Marks
;             REGISTRATION NUMBER: 41,971
;             REFERENCE/DOCKET NUMBER: PB186P3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-309-8504
;             TELEFAX: 301-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1830121 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Alignment Scores:
Pred. No.:                2.06e+04        Length:          1830121
Score:                    89.00           Matches:         52
Percent Similarity:       36.8%           Conservative:    44
Best Local Similarity:    19.9%           Mismatches:      109
Query Match:              7.5%            Indels:          56
DB:                       3               Gaps:            11

US-10-757-745-2_COPY_121_347 (1-227) x US-09-557-884-1 (1-1830121)

Qy           5 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 24
               ::::::|||   |||       |||         :::   :::   :::      |||
Db       43854 AATATAAATGGGTTACGTGCTCGCCCACATCAACTTGAAGCCATTATTGAAAAATATCAG 43795

Qy          25 ProAspValIlePheLeuGlnGlu------ValIleProProTyrTyrSerTyrLeuLys 42
               ||||||||||||   |||||||||       |||          :::    |||
Db       43794 CCAGATGTGATTGGCTTACAGGAAATTAAAGTTGCTGACGAGGCTTTTCCCTATGAAATT 43735

Qy          43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                     :::       |||   :::           ::::::|||::::::
Db       43734 ACAGAAAACTTAGGTTATCACGTATTTCATCACGGTCAGAAAGGTCATTATGGCGTGGCA 43675

Qy          63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
               :::|||   |||         |||:::     :::         |||      :::
Db       43674 TTATTGACTAAACAAGAACCTAAAGTG---ATTCGTCGTGGTTTTCCAACTGACAATGAA 43618

Qy          83 LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu 102
                        :::   :::::      :::                  |||      |||
Db       43617 GATGCACAAAAACGCATTATTATGGCTGATCTGGAA---------ACTGAATTTGGCTTA 43567

Qy         103 MetThr------------------------------SerHisLeuGluSerThrArgGly 112
               :::|||                            |||             |||:::
Db       43566 TTGACTGTCATTAATGGCTATTTCCCACAGGGGGAAAGCCGTGCGCACGAGACTAAATTT 43507

Qy         113 HisAlaAlaGluArgMet---AsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 131
                  |||   |||:::         |||:::      |||:::|||       |||
Db       43506 CCAGCTAAAGAGAAATTTTATGCAGATCTTCAACAGTATTTAGAAAAA------GAACAC 43453

Qy         132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 151
               :::|||   :::::      ||||||   |||:::       ::::::
Db       43452 GACAAATCGAATCCGATTTTAATTATGGGCGATATGAATATTAGCCCAAGTGATTTAGAT 43393

Qy         152 ------------------------------------------ArgCysGlyGlyLeu 156
                                                        :::|||      |||
Db       43392 ATTGGTATCGGAGATGAAAATCGTAAACGTTGGTTACGCACAGGGAAATGTTCTTTCTTA 43333

Qy         157 ProAsn---------------------------------AsnIleValAspValTrpGlu 165
               |||                                  :::   |||   :::
Db       43332 CCGGAAGAACGAGCGTGGTATCAACGTTTGTATGATTACGGATTAGAAGACAGTTTTCGT 43273

Qy         166 PheLeuGlyLysProLysHisCysGlnTyrThrTrp------------AspThrGlnMet 181
                  |||      :::   ::::::::::|||
Db       43272 AAGCTGAACCCAACAGCAAATGATAAATTCTCGTGGTTTGATTACCGTTCTAAAGGTTTT 43213

Qy         182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe--- 200
               :::   |||   |||                |||||   |||   |||
Db       43212 GATGATAATCGAGGT-----------------TTACGAATTGATCATATTTTGGTAAGC 43171

Qy         201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
               :::    |||||      :::     :::|||||:::       :::|||
Db       43170 CAAAAATTAGCTGAACGATGTGTCGATGTGGGCATTGCATTAGATATTCGAGCAATGGAA 43111

Qy         221 Lys 221
               |||
```

```
Db      43110 AAA 43108

RESULT 18
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
;   GENERAL INFORMATION:
;        APPLICANT: Robert D. Fleischmann
;                   Mark D. Adams
;                   Owen White
;                   Hamilton O. Smith
;                   J. Craig Venter
;        TITLE OF INVENTION: The Nucleotide sequence of
;                            the Haemophilus influenzae Rd Genome, Fragments
;                            Thereof, and Uses Thereof
;        NUMBER OF SEQUENCES: 1
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville,
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3 1/2 inch diskette
;             COMPUTER: Dell Pentium
;             OPERATING SYSTEM: MS DOS v6.22
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/643,990A
;             FILING DATE: 23-Aug-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/487,429
;             FILING DATE: 1995-06-07
;             APPLICATION NUMBER: 08/426,787
;             FILING DATE: 1995-04-21
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB186P1C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-610-5790
;             TELEFAX: 310-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1830121 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Alignment Scores:
Pred. No.:                2.06e+04        Length:          1830121
Score:                    89.00           Matches:         52
Percent Similarity:       36.8%           Conservative:    44
Best Local Similarity:    19.9%           Mismatches:      109
Query Match:              7.5%            Indels:          56
DB:                       3               Gaps:            11

US-10-757-745-2_COPY_121_347 (1-227) x US-09-643-990A-1 (1-1830121)

Qy          5 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 24
              ::::::|||   |||      |||      :::  :::  :::     |||
Db      43854 AATATAAATGGGTTACGTGCTCGCCCACATCAACTTGAAGCCATTATTGAAAAATATCAG 43795

Qy         25 ProAspValIlePheLeuGlnGlu------ValIleProProTyrTyrSerTyrLeuLys 42
              ||||||||||||  |||||||||      |||          :::   |||
Db      43794 CCAGATGTGATTGGCTTACAGGAAATTAAAGTTGCTGACGAGGCTTTTCCCTATGAAATT 43735

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                 :::      |||   :::           ::::::|||::::::

Db      43734 ACAGAAAACTTAGGTTATCACGTATTTCATCACGGTCAGAAAGGTCATTATGGCGTGGCA 43675

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::|||   |||        |||:::    :::       |||      :::
Db      43674 TTATTGACTAAACAAGAACCTAAAGTG---ATTCGTCGTGGTTTTCCAACTGACAATGAA 43618

Qy         83 LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu 102
                    :::   ::::::      :::                  |||     |||
Db      43617 GATGCACAAAAACGCATTATTATGGCTGATCTGGAA---------ACTGAATTTGGCTTA 43567

Qy        103 MetThr----------------------------SerHisLeuGluSerThrArgGly 112
              :::|||                          |||           |||:::
Db      43566 TTGACTGTCATTAATGGCTATTTCCCACAGGGGGAAAGCCGTGCGCACGAGACTAAATTT 43507

Qy        113 HisAlaAlaGluArgMet---AsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 131
                |||   |||:::          |||:::      |||:::|||     |||
Db      43506 CCAGCTAAAGAGAAATTTATGCAGATCTTCAACAGTATTTAGAAAAA------GAACAC 43453

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 151
              :::|||      ::::::     ||||||   |||:::        ::::::
Db      43452 GACAAATCGAATCCGATTTTAATTATGGGCGATATGAATATTAGCCCAAGTGATTTAGAT 43393

Qy        152 ----------------------------------------------ArgCysGlyGlyLeu 156
                                                           :::|||     |||
Db      43392 ATTGGTATCGGAGATGAAAATCGTAAACGTTGGTTACGCACAGGGAAATGTTCTTTCTTA 43333

Qy        157 ProAsn--------------------------------AsnIleValAspValTrpGlu 165
              |||                                 :::   |||   :::
Db      43332 CCGGAAGAACGAGCGTGGTATCAACGTTTGTATGATTACGGATTAGAAGACAGTTTTCGT 43273

Qy        166 PheLeuGlyLysProLysHisCysGlnTyrThrTrp-----------AspThrGlnMet 181
                |||           :::   :::::::::|||
Db      43272 AAGCTGAACCCAACAGCAAATGATAAATTCTCGTGGTTTGATTACCGTTCTAAAGGTTTT 43213

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe--- 200
              :::   |||   |||                |||||   |||   |||
Db      43212 GATGATAATCGAGGT----------------TTACGAATTGATCATATTTTGGTAAGC 43171

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              :::      ||||||          :::    :::||||||:::     :::|||
Db      43170 CAAAAATTAGCTGAACGATGTGTCGATGTGGGCATTGCATTAGATATTCGAGCAATGGAA 43111

Qy        221 Lys 221
              |||
Db      43110 AAA 43108

RESULT 19
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fra
; Patent No. 6846651
;  TITLE OF INVENTION:  Thereof, and Uses Thereof
;  FILE REFERENCE: PB186P2C1D1
;  CURRENT APPLICATION NUMBER: US/10/158,865
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US 09/557,884
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US 08/476,102
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/426,787
;  PRIOR FILING DATE: 1995-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1830121
;   TYPE: DNA
;   ORGANISM: Haemophilus influenzae
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4747)..(4747)
```

; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

```
;   NAME/KEY: misc_feature
;   LOCATION: (142750)..(142750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145058)..(145058)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145171)..(145171)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145942)..(145942)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (147197)..(147197)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (150841)..(150841)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152500)..(152500)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature

Alignment Scores:
Pred. No.:                2.06e+04      Length:          1830121
Score:                    89.00         Matches:         52
Percent Similarity:       36.8%         Conservative:    44
Best Local Similarity:    19.9%         Mismatches:      109
Query Match:              7.5%          Indels:          56
DB:                       3             Gaps:            11

US-10-757-745-2_COPY_121_347 (1-227) x US-10-158-865-1 (1-1830121)

Qy          5 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 24
              ::::::|||   |||      |||         :::   :::   :::      |||
Db      43854 AATATAAATGGGTTACGTGCTCGCCCACATCAACTTGAAGCCATTATTGAAAAATATCAG 43795

Qy         25 ProAspValIlePheLeuGlnGlu------ValIleProProTyrTyrSerTyrLeuLys 42
              ||||||||||||  |||||||||      |||           :::   |||
Db      43794 CCAGATGTGATTGGCTTACAGGAAATTAAAGTTGCTGACGAGGCTTTTCCCTATGAAATT 43735

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                 :::      |||   :::          ::::::|||::::::
Db      43734 ACAGAAAACTTAGGTTATCACGTATTTCATCACGGTCAGAAAGGTCATTATGGCGTGGCA 43675

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              :::|||   |||         |||:::      :::         |||      :::
Db      43674 TTATTGACTAAACAAGAACCTAAAGTG---ATTCGTCGTGGTTTTCCAACTGACAATGAA 43618

Qy         83 LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu 102
                 :::   :::::         :::                |||      |||
Db      43617 GATGCACAAAAACGCATTATTATGGCTGATCTGGAA---------ACTGAATTTGGCTTA 43567

Qy        103 MetThr----------------------------SerHisLeuGluSerThrArgGly 112
              :::|||                          |||          |||:::
Db      43566 TTGACTGTCATTAATGGCTATTTCCCACAGGGGGAAAGCCGTGCGCACGAGACTAAATTT 43507

Qy        113 HisAlaAlaGluArgMet---AsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 131
                |||   |||:::           |||:::      |||:::|||      |||
Db      43506 CCAGCTAAAGAGAAATTTTATGCAGATCTTCAACAGTATTTAGAAAAA------GAACAC 43453

Qy        132 ProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr 151
                :::|||      ::::::      ||||||   |||:::          ::::::
Db      43452 GACAAATCGAATCCGATTTTAATTATGGGCGATATGAATATTAGCCCAAGTGATTTAGAT 43393

Qy        152 ----------------------------------------------ArgCysGlyGlyLeu 156

Db      43392 ATTGGTATCGGAGATGAAAATCGTAAACGTTGGTTACGCACAGGGAAATGTTCTTTCTTA 43333

Qy        157 ProAsn--------------------------------AsnIleValAspValTrpGlu 165
              |||                                 :::   |||   :::
Db      43332 CCGGAAGAACGAGCGTGGTATCAACGTTTGTATGATTACGGATTAGAAGACAGTTTTCGT 43273

Qy        166 PheLeuGlyLysProLysHisCysGlnTyrThrTrp-----------AspThrGlnMet 181
                |||           :::   :::::::::|||
Db      43272 AAGCTGAACCCAACAGCAAATGATAAATTCTCGTGGTTTGATTACCGTTCTAAAGGTTTT 43213

Qy        182 AsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePhe--- 200
              :::   |||   |||                |||||   |||   |||
Db      43212 GATGATAATCGAGGT-----------------TTACGAATTGATCATATTTTGGTAAGC 43171

Qy        201 ArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeuLeuGlyLeuGlu 220
              :::      ||||||          :::      :::||||||:::      :::|||
Db      43170 CAAAAATTAGCTGAACGATGTGTCGATGTGGGCATTGCATTAGATATTCGAGCAATGGAA 43111

Qy        221 Lys 221
              |||
Db      43110 AAA 43108

RESULT 20
US-09-107-532A-2664
; Sequence 2664, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;       APPLICANT: Lynn A Doucette-Stamm  and David Bush
;       TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                           ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;       NUMBER OF SEQUENCES: 7310
;       CORRESPONDENCE ADDRESS:
;           ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;           STREET: 100 Beaver Street
;           CITY: Waltham
;           STATE: Massachusetts
;           COUNTRY: USA
;           ZIP: 02354
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: CD/ROM ISO9660
;           COMPUTER: PC
;           OPERATING SYSTEM: <Unknown>
;           SOFTWARE: ASCII
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/107,532A
;           FILING DATE: 30-Jun-1998
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 60/085,598
;           FILING DATE: 14 May 1998
;           APPLICATION NUMBER: 60/051571
;           FILING DATE: July 2, 1997
;       ATTORNEY/AGENT INFORMATION:
;           NAME: Ariniello, Pamela Deneke
;           REGISTRATION NUMBER: 40,489
;           REFERENCE/DOCKET NUMBER: GTC-012
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: (781)893-5007
;           TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 2664:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 690 base pairs
;           TYPE: nucleic acid
;           STRANDEDNESS: double
;           TOPOLOGY: circular
;       MOLECULE TYPE: DNA (genomic)
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       ORIGINAL SOURCE:
;           ORGANISM: Enterococcus faecium
;       FEATURE:
;           NAME/KEY:  misc_feature
;           LOCATION: (B) LOCATION 1...690
```

```
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2664:
US-09-107-532A-2664

Alignment Scores:
Pred. No.:              0.178          Length:         690
Score:                  88.50          Matches:        47
Percent Similarity:     39.1%          Conservative:   30
Best Local Similarity:  23.9%          Mismatches:     77
Query Match:            7.4%           Indels:         43
DB:                     3              Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-107-532A-2664 (1-690)

Qy          4 LeuAspLeuAsnAsnLeuSerGluArg---AlaArgGlyValCysSerTyrLeuAlaLeu 22
              :::::::::||||||:::      |||      :::           :::|||||| 
Db        163 ATAAATATCAATAACATGGAAACACGCTTGGAGAAGAAAAGTTTTGAAATGATCGCATTG 222

Qy         23 TyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 42
                |||         ||:::   :::|||             :::
Db        223 CAACAGCCAGTTACAACTGATTTGCGTATGATCATTACG--------------GTCATG 267

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
              |||   ||||||:::   |||   :::      |||                 |||::: 
Db        268 AAAGCTAGCTCGGATTTAGAACGTATGGCCGATCAT-----------------GCGGTG 309

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer--- 81
              :::   ||||||   :::::::::|||   |||      ||||||
Db        310 TCTGTTGCTAAATCCACGATTCGCGTAAAAGGGCAGACGAGAATACCGGAAATCGAAAAA 369

Qy         82 --------------------------ThrLysMetMetArgAsnLeuLeuCysValHis 92
                                        |||||:::  |||:::|||    :::
Db        370 GAGATCTCCGATATGTCCGATTATGTGAAAAAAATGGTGGACAACGTCCTAGTTGCTTAT 429

Qy         93 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              |||                 :::                              |||
Db        430 GTCAAAACAGACGAAAAAGAC----------------------GCACGTACGATCGCC 465

Qy        113 HisAlaAlaGluArgMetAsnGln--------------LeuLysMetValLeuLysLys 127
              :::       :::|||:::|||                        :::|||
Db        466 AATATGGATCAAAGAGTCAATGACTATTTCAATCGAATCTATAATGCAACAATCAAAAAT 525

Qy        128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAsp 147
              |||||||||   ||||||      |||||||||       |||
Db        526 ATGCAGGAAAATCCTGAG------ACAGTCATCAGTGGAACAGACTACTTGAATGTTGCC 579

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::   |||   |||             ||||||   :::         :::|||
Db        580 AGTTATCTAGAACGGATCGGAGACTATGTGACAAACATCTGTGAATGGATCGTCTACTTG 639

Qy        168 GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
                    |||                 |||::::::|||:::|||
Db        640 GCTACAGGGAAA-----------------ATCACTGAGTTAAATACGAAC 672

RESULT 21
US-09-949-016-17553/c
; Sequence 17553, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17553
;  LENGTH: 106380
;  TYPE: DNA
;  ORGANISM: Human
US-09-949-016-17553

Alignment Scores:
Pred. No.:              519            Length:         106380
Score:                  87.00          Matches:        45
Percent Similarity:     42.6%          Conservative:   36
Best Local Similarity:  23.7%          Mismatches:     77
Query Match:            7.3%           Indels:         32
DB:                     3              Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-17553 (1-106380)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||  :::  |||::::::          |||
Db      22404 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 22348

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||       :::||||||                   ::::::|||      |||
Db      22347 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCATGTGCAGAGAC 22306

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||       :::   |||                |||
Db      22305 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATAGAAAACAAAAA 22246

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::          :::      |||       |||||:::
Db      22245 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 22186

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::           ::::::   |||   |||:::           ||||||
Db      22185 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 22126

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||            |||   |||         :::
Db      22125 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 22081

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||          :::|||   |||:::|||      |||
Db      22080 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 22024

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db      22023 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAC 21964

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                  :::          :::::::::|||::::::      |||
Db      21963 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAATTCTC 21904

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                    ||||||      :::||||||:::
Db      21903 ---CACCCCAAATCAACAGAATATACATTT 21877

RESULT 22
US-09-949-016-11945/c
; Sequence 11945, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11945
;   LENGTH: 86439
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-11945

Alignment Scores:
Pred. No.:              508          Length:         86439
Score:                  86.00        Matches:        45
Percent Similarity:     42.6%        Conservative:   36
Best Local Similarity:  23.7%        Mismatches:     77
Query Match:            7.2%         Indels:         32
DB:                     3            Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-11945 (1-86439)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||  |||  :::  |||::::::             |||
Db      28224 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 28168

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
               |||      :::||||||                :::::::|||       |||
Db      28167 TCAGTGTGCTGTATTCAGGAA-----------------ACTCATCTCACGTGCAGAGAC 28126

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::       |||  |||      :::  |||                    |||
Db      28125 ACACATAGGCTCAAAATAAAAGGATGGAGAAAGATCTACCAAGCCAATGGAAAACAAAAA 28066

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||  :::         :::          |||       ||||||:::
Db      28065 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 28006

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::            ::::::  |||  |||:::            ||||||
Db      28005 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 27946

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||   |||        :::
Db      27945 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 27901

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db      27900 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 27844

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db      27843 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 27784

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
               :::                        :::::::::|||::::::      |||
Db      27783 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGGGACCTAATAGACATCTACAGAACTCTC 27724

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                ||||||      :::||||||:::
Db      27723 ---CACCCCAAATCAACAGAATATACATTT 27697

RESULT 23
US-09-949-016-16990/c
; Sequence 16990, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16990
;   LENGTH: 86440
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16990

Alignment Scores:
Pred. No.:              508          Length:         86440
Score:                  86.00        Matches:        45
Percent Similarity:     42.6%        Conservative:   36
Best Local Similarity:  23.7%        Mismatches:     77
Query Match:            7.2%         Indels:         32
DB:                     3            Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-16990 (1-86440)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||  |||  :::  |||::::::             |||
Db      28224 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 28168

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
               |||      :::||||||                :::::::|||       |||
Db      28167 TCAGTGTGCTGTATTCAGGAA-----------------ACTCATCTCACGTGCAGAGAC 28126

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::       |||  |||      :::  |||                    |||
Db      28125 ACACATAGGCTCAAAATAAAAGGATGGAGAAAGATCTACCAAGCCAATGGAAAACAAAAA 28066

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||  :::         :::          |||       ||||||:::
Db      28065 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 28006

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::            ::::::  |||  |||:::            ||||||
Db      28005 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 27946

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||   |||        :::
Db      27945 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 27901

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db      27900 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 27844

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db      27843 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 27784

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
               :::                        :::::::::|||::::::      |||
Db      27783 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGGGACCTAATAGACATCTACAGAACTCTC 27724

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                ||||||      :::||||||:::
Db      27723 ---CACCCCAAATCAACAGAATATACATTT 27697

RESULT 24
US-09-949-016-16293/c
; Sequence 16293, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
```

```
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16293
;   LENGTH: 133063
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(133063)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16293

Alignment Scores:
Pred. No.:              969           Length:         133063
Score:                  86.00         Matches:        52
Percent Similarity:     46.7%         Conservative:   40
Best Local Similarity:  26.4%         Mismatches:     68
Query Match:            7.2%          Indels:         37
DB:                     3             Gaps:           10

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-16293 (1-133063)

Qy          4 LeuAspLeuAsnAsnLeuSer---------GluArgAlaArgGlyValCysSerTyrLeu 20
              |||:::::::||   |||:::         ||||||   |||   :::   :::::::::
Db     114630 TTAAATGTAAATGGACTAAATGCTCCAATTGAAAGACACAGA---CTGGCAAATTGGATA 114574

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                        |||   |||      :::||||||                  ::::::
Db     114573 AAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCAT 114532

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||       |||   ::::::          |||   |||       :::   |||
Db     114531 CTCACATGCAGAGACACACACAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCA 114472

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                        ||||||:::   |||   :::          :::          |||
Db     114471 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCAGATAAAACAGACTTTAAA 114412

Qy         81 SerThrLysMetMetArg------------AsnLeuLeuCysValHisValAsnValSer 96
                 ||||||:::   |||            :::   :::   |||   ::::::
Db     114411 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGAGATCAATTCAA 114352

Qy         97 GlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                 ||||||   ::::::      :::      :::|||                |||
Db     114351 CAAGAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCC 114301

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||       :::|||  ||||||      :::|||          :::|||   |||
Db     114300 AGATTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACA 114247

Qy        137 ValIlePheAlaGlyAspThrAsn-Leu-----------------------ArgAspAr 148
              :::|||      ||||||   ||| |||                      ||||||||
Db     114246 ATAATAATG---GGAGACTTTAACACTCCACTGTCAACATTAGACAGATCAACGAGACAG 114190

Qy        148 gGluValThrArgCys-GlyGlyLeu---------------ProAsn-AsnIleValAsp 162
              |::::::||||||       |||:::              |||::: :::::::::|||
Db     114189 AAAGTTAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGAC 114130

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              ::::::   |||       ||||||      :::|||||:::
Db     114129 ATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACTTTT 114088

RESULT 25
US-09-949-016-14546/c
; Sequence 14546, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14546
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14546

Alignment Scores:
Pred. No.:              1.47e+04      Length:         818128
Score:                  86.00         Matches:        45
Percent Similarity:     42.6%         Conservative:   36
Best Local Similarity:  23.7%         Mismatches:     77
Query Match:            7.2%          Indels:         32
DB:                     3             Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14546 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::       :::|||   |||   :::   |||::::::            |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                   |||       :::|||||||                  ::::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::               |||   |||       :::   |||            |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::          |||       ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::       :::::   |||  |||:::                 ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||              |||   |||          :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::                        :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                    ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 26
US-09-949-016-14547/c
```

```
; Sequence 14547, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14547
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14547

Alignment Scores:
Pred. No.:                1.47e+04        Length:         818128
Score:                    86.00           Matches:        45
Percent Similarity:       42.6%           Conservative:   36
Best Local Similarity:    23.7%           Mismatches:     77
Query Match:              7.2%            Indels:         32
DB:                       3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14547 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||  |||   :::   |||::::::            |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
               |||       :::||||||                  ::::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||     :::   |||                  |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::           ::::::   |||   |||:::           ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::     :::|||          |||   |||        :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::          :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                  ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 27
US-09-949-016-14548/c
; Sequence 14548, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14548
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14548

Alignment Scores:
Pred. No.:                1.47e+04        Length:         818128
Score:                    86.00           Matches:        45
Percent Similarity:       42.6%           Conservative:   36
Best Local Similarity:    23.7%           Mismatches:     77
Query Match:              7.2%            Indels:         32
DB:                       3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14548 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||  |||   :::   |||::::::            |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
               |||       :::||||||                  ::::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||     :::   |||                  |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::           ::::::   |||   |||:::           ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::     :::|||          |||   |||        :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
```

```
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
              :::                          :::::::::|||:::::       |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                  ||||||      :::|||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489
```

RESULT 28
US-09-949-016-14549/c
; Sequence 14549, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14549
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14549

Alignment Scores:

Pred. No.:	1.47e+04	Length:	818128
Score:	86.00	Matches:	45
Percent Similarity:	42.6%	Conservative:	36
Best Local Similarity:	23.7%	Mismatches:	77
Query Match:	7.2%	Indels:	32
DB:	3	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14549 (1-818128)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::       :::|||   |||   :::   |||::::::          |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||       :::||||||                ::::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||       :::   |||                 |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::       :::           |||      |||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::           ::::::   |||   |||:::           ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::   :::     :::|||         |||   |||          :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693
```

```
Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||          :::|||   |||:::|||       |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                  :::                          :::::::::|||:::::       |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                  ||||||      :::|||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489
```

RESULT 29
US-09-949-016-14550/c
; Sequence 14550, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14550
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14550

Alignment Scores:

Pred. No.:	1.47e+04	Length:	818128
Score:	86.00	Matches:	45
Percent Similarity:	42.6%	Conservative:	36
Best Local Similarity:	23.7%	Mismatches:	77
Query Match:	7.2%	Indels:	32
DB:	3	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14550 (1-818128)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::       :::|||   |||   :::   |||::::::          |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||       :::||||||                ::::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||       :::   |||                 |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::       :::           |||      |||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::           ::::::   |||   |||:::           ||||||
```

```
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              :::::::      :::      :::|||            |||   |||         :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                            :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::                          ::::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                  ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 30
US-09-949-016-14551/c
; Sequence 14551, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14551
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14551

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14551 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::            |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||       :::||||||                  :::::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||      :::   |||                  |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy          66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
               |||:::   |||   :::          :::          |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy          86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
               |||:::            ::::::   |||   |||:::            ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy         102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
               ::::::      :::      :::|||            |||   |||         :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy         122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
               |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy         142 AspThrAsnLeu-----------------------------------------ArgAsp 147
               |||   |||                                            :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy         148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::                          ::::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy         168 GlyLysProLysHisCysGlnTyrThrTrp 177
                  ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 31
US-09-949-016-14552/c
; Sequence 14552, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14552
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14552

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14552 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::            |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||       :::||||||                  :::::::|||      |||
```

```
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65

Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85

Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101

Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121

Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141

Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147

Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167

Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177

Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 32
US-09-949-016-14553/c
; Sequence 14553, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14553
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14553

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14553 (1-818128)
```

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25

Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45

Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65

Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85

Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101

Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121

Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141

Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147

Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167

Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177

Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 33
US-09-949-016-14554/c
; Sequence 14554, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14554
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14554

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
```

```
Best Local Similarity:  23.7%          Mismatches:    77
Query Match:            7.2%           Indels:        32
DB:                     3              Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14554 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||:::::              |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
               |||      :::||||||                  :::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       ::   |||                   |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::            ::::::   |||   |||:::          ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||   |||          :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||      :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::                          :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                    ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 34
US-09-949-016-14555/c
; Sequence 14555, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14555
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G

US-09-949-016-14555

Alignment Scores:
Pred. No.:              1.47e+04       Length:        818128
Score:                  86.00          Matches:       45
Percent Similarity:     42.6%          Conservative:  36
Best Local Similarity:  23.7%          Mismatches:    77
Query Match:            7.2%           Indels:        32
DB:                     3              Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14555 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||:::::              |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
               |||      :::||||||                  :::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       ::   |||                   |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::            ::::::   |||   |||:::          ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||   |||          :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||      :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::                          :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                    ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 35
US-09-949-016-14556/c
; Sequence 14556, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14556
;   LENGTH: 818128
```

```
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14556

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14556 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::           |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
              |||      :::||||||                :::::::|||          |||
Db     401959 TCAGTGTGCTGTATTCAGGAA----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       :::   |||                |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::          ::::::  |||  |||:::             ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||           |||   |||         :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||       :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::                    :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                 ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 36
US-09-949-016-14557/c
; Sequence 14557, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14557
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14557

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14557 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::           |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
              |||      :::||||||                :::::::|||          |||
Db     401959 TCAGTGTGCTGTATTCAGGAA----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       :::   |||                |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::          ::::::  |||  |||:::             ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||           |||   |||         :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||       :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::                    :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                 ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 37
US-09-949-016-14558/c
; Sequence 14558, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
```

```
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14558
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14558

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14558 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||:::::::             |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::||||||                   :::::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       :::   |||                  |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::  |||  :::       :::       |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::          ::::::   |||   |||:::            ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||           |||   |||          :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::                         :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                ||||||      :::|||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 38
US-09-949-016-14559/c
; Sequence 14559, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14559
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14559

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14559 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||:::::::             |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::||||||                   :::::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||       :::   |||                  |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::  |||  :::       :::       |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::          ::::::   |||   |||:::            ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||           |||   |||          :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::                         :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
```

```
                    ||||||      :::|||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 39
US-09-949-016-14560/c
; Sequence 14560, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14560
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14560

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14560 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::           |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
               |||      :::||||||                 :::::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||       :::   |||                   |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::          ::::::  |||  |||:::                 ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||            |||   |||          :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||       |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-------------------------------------------ArgAsp 147
              |||   |||                                             :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                        :::          :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                 ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 40
US-09-949-016-14561/c
; Sequence 14561, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14561
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14561

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14561 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::           |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
               |||      :::||||||                 :::::::|||      |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||       :::   |||                   |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::          ::::::  |||  |||:::                 ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||            |||   |||          :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
```

```
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu----------------------------------------------ArgAsp 147
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 41
US-09-949-016-14562/c
; Sequence 14562, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14562
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14562

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14562 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
Db     401959 TCAGTGTGCTGTATTCAGGAA----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------------GluLeuCys 101
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu----------------------------------------------ArgAsp 147
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 42
US-09-949-016-14564/c
; Sequence 14564, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14564
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14564

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14564 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
Db     401959 TCAGTGTGCTGTATTCAGGAA----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
```

```
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy          86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101

Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy         102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121

Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy         122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141

Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy         142 AspThrAsnLeu----------------------------------------ArgAsp 147

Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy         148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167

Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy         168 GlyLysProLysHisCysGlnTyrThrTrp 177

Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 43
US-09-949-016-14565/c
; Sequence 14565, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14565
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14565

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14565 (1-818128)

Qy           6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25

Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy          26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45

Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy          46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65

Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy          66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85

Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy          86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101

Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy         102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121

Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy         122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141

Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy         142 AspThrAsnLeu----------------------------------------ArgAsp 147

Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy         148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167

Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy         168 GlyLysProLysHisCysGlnTyrThrTrp 177

Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 44
US-09-949-016-14566/c
; Sequence 14566, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14566
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14566

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14566 (1-818128)

Qy           6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
```

```
                ||||||:::      :::|||   |||   :::   |||::::::              |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                  |||      :::||||||                :::::|||     |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
                  ::::::         |||   |||       :::   |||                |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
                  |||:::   |||   :::         :::         |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
                  |||:::              ::::::   |||   |||:::             ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
                  ::::::      :::      :::|||           |||   |||         :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
                  |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
                  |||   |||                                         :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                  :::                              :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                  ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 45
US-09-949-016-14567/c
; Sequence 14567, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14567
;   LENGTH: 818128
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(818128)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14567

Alignment Scores:
Pred. No.:              1.47e+04        Length:         818128
Score:                  86.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77

Query Match:            7.2%            Indels:         32
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14567 (1-818128)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
                |||||||:::      :::|||   |||   :::   |||::::::              |||
Db     402016 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 401960

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                  |||      :::||||||                :::::|||     |||
Db     401959 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 401918

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
                  ::::::         |||   |||       :::   |||                |||
Db     401917 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 401858

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
                  |||:::   |||   :::         :::         |||      ||||||:::
Db     401857 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 401798

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
                  |||:::              ::::::   |||   |||:::             ||||||
Db     401797 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 401738

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
                  ::::::      :::      :::|||           |||   |||         :::
Db     401737 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 401693

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
                  |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db     401692 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 401636

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
                  |||   |||                                         :::|||
Db     401635 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 401576

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                  :::                              :::::::::|||::::::      |||
Db     401575 ACCCAGGAATTGAACTCAGCTCTGCACCAAGGAGACCTAATAGACATCTACAGAACTCTC 401516

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                  ||||||      :::||||||:::
Db     401515 ---CACCCCAAATCAACAGAATATACATTT 401489

RESULT 46
US-09-774-528-260/c
; Sequence 260, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION:  No. 6743619el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE:  802
;  CURRENT APPLICATION NUMBER: US/09/774,528
;  CURRENT FILING DATE:  2001-01-30
;  NUMBER OF SEQ ID NOS:  441
;  SOFTWARE:  pt_FL_genes Version 2.0
; SEQ ID NO 260
;   LENGTH: 1575
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1152)
US-09-774-528-260

Alignment Scores:
Pred. No.:                1.46          Length:        1575
Score:                    85.50         Matches:       45
Percent Similarity:       39.3%         Conservative:  39
Best Local Similarity:    21.0%         Mismatches:    84
Query Match:              7.2%          Indels:        46
DB:                       3             Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-774-528-260 (1-1575)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
Db       1053 GAAGCGGTTGACATCAGTAAACTGTCTCCTTCTTCTAGCTTCATCTGCTAATTCCAGTGC 994

Qy         22 LeuTyrSerProAspVal--------------------------------IlePheLeu 30
Db        993 TTGTACAAGAACAATATCATCATTAGAGGAGAAAATGGTCAGAGGAGGTGTATCTGGATC 934

Qy         31 GlnGluValIle-----------------------------------------ProPro 36
Db        933 AGGGAAGTTACGCTGAAGTGGATCATAGTGGATGCCATCATAAATAAGCAGAACCCTTTT 874

Qy         37 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 56
Db        873 GGTATATCCTGCATCTTCCCCAAAACGATCAATTCTTACTGTGTGTGTATCCACTACACA 814

Qy         57 GlyTyrPheThrAlaIleMetLeuLysLysSerArg---ValLysLeuLysSerGlnGlu 75
Db        813 ---TATTTCACATTGGTAAAACTTGGACAAAATCGATATCTCTATTGCTCCTCCCCAAGT 757

Qy         76 IleIleProPhe---ProSerThrLysMetMet------------ArgAsnLeuLeuCys 90
Db        756 GTCATCCCTTTTGATCCAGTCACAGTACTCTTGATTTGTTTTTCCCAGTATTGCCTCACT 697

Qy         91 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 110
Db        696 ATAGAAGTCTGGATCGCTTGCTACAATTTGTGCTATGAGGCGTCTCATCTCAGGGGCACA 637

Qy        111 -ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGl 130
Db        636 AGCTGGATTCAAGACTCCTCCTTCGACGACATAGTACACACTAGTAAAGAGGCAAGAGTT 577

Qy        130 uAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluVa 150
Db        576 GTCTGCTGGGACCACGGTTCTGGTAAGCACAGGCAAAGTTTCCCTGACGTAACTAGAAGC 517

Qy        150 lThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPr 170
Db        516 ACCACGT--------------------TTAGTAAATGCAGGTGAACTTCTGGGC----- 483

Qy        170 oLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
Db        482 -------------CTGGTTTGGTCTTCTTCAATGATCAGC 456

RESULT 47
US-10-120-988-260/c
; Sequence 260, Application US/10120988
; Patent No. 6919193
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6919193el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 802CON
;  CURRENT APPLICATION NUMBER: US/10/120,988
;  CURRENT FILING DATE:  2002-04-11
;  PRIOR APPLICATION NUMBER: 09/774,528
;  PRIOR FILING DATE: 2001-01-30
;  NUMBER OF SEQ ID NOS: 441
;  SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 260
;   LENGTH: 1575
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1152)
US-10-120-988-260

Alignment Scores:
Pred. No.:                1.46          Length:        1575
Score:                    85.50         Matches:       45
Percent Similarity:       39.3%         Conservative:  39
Best Local Similarity:    21.0%         Mismatches:    84
Query Match:              7.2%          Indels:        46
DB:                       3             Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-10-120-988-260 (1-1575)

Qy          2 AspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 21
Db       1053 GAAGCGGTTGACATCAGTAAACTGTCTCCTTCTTCTAGCTTCATCTGCTAATTCCAGTGC 994

Qy         22 LeuTyrSerProAspVal--------------------------------IlePheLeu 30
Db        993 TTGTACAAGAACAATATCATCATTAGAGGAGAAAATGGTCAGAGGAGGTGTATCTGGATC 934

Qy         31 GlnGluValIle-----------------------------------------ProPro 36
Db        933 AGGGAAGTTACGCTGAAGTGGATCATAGTGGATGCCATCATAAATAAGCAGAACCCTTTT 874

Qy         37 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 56
Db        873 GGTATATCCTGCATCTTCCCCAAAACGATCAATTCTTACTGTGTGTGTATCCACTACACA 814

Qy         57 GlyTyrPheThrAlaIleMetLeuLysLysSerArg---ValLysLeuLysSerGlnGlu 75
Db        813 ---TATTTCACATTGGTAAAACTTGGACAAAATCGATATCTCTATTGCTCCTCCCCAAGT 757

Qy         76 IleIleProPhe---ProSerThrLysMetMet------------ArgAsnLeuLeuCys 90
Db        756 GTCATCCCTTTTGATCCAGTCACAGTACTCTTGATTTGTTTTTCCCAGTATTGCCTCACT 697

Qy         91 ValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThr 110
Db        696 ATAGAAGTCTGGATCGCTTGCTACAATTTGTGCTATGAGGCGTCTCATCTCAGGGGCACA 637

Qy        111 -ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGl 130
Db        636 AGCTGGATTCAAGACTCCTCCTTCGACGACATAGTACACACTAGTAAAGAGGCAAGAGTT 577

Qy        130 uAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluVa 150
Db        576 GTCTGCTGGGACCACGGTTCTGGTAAGCACAGGCAAAGTTTCCCTGACGTAACTAGAAGC 517

Qy        150 lThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysPr 170
Db        516 ACCACGT--------------------TTAGTAAATGCAGGTGAACTTCTGGGC----- 483

Qy        170 oLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSer 183
Db        482 -------------CTGGTTTGGTCTTCTTCAATGATCAGC 456

RESULT 48
US-09-252-991A-8106
```

```
; Sequence 8106, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 8106
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8106

Alignment Scores:
Pred. No.:                0.561          Length:          756
Score:                    85.00          Matches:         46
Percent Similarity:       34.6%          Conservative:    20
Best Local Similarity:    24.1%          Mismatches:      51
Query Match:              7.1%           Indels:          74
DB:                       3              Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-252-991A-8106 (1-756)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              |||       |||   |||       |||   |||
Db         94 CTGCCGGAGTTGCGCGAAGCGGTGCGCAGCGTC--------------------GGTGCC 132

Qy         26 AspValIlePheLeuGlnGluValIle---------------------------------- 34
              |||::::::|||||||||||||||:::
Db        133 GACCTGGTGTTCCTCCAGGAGGTCCTCGGCTCCCATAGCCTGCACGCCGCGCGCCTGCCG 192

Qy         35 ------------ProProTyrTyrSerTyrLeuLysLys---------------------- 43
                        |||   |||   ||||||
Db        193 TCCTGGCCGCCGGCGCCGCAATACGAGTATCTGGCCGACAGCATGTGGCCGCAGTTCGCC 252

Qy         44 ------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                  |||::::::   |||                       ||||||:::
Db        253 TACGGGCGCAACGCGGTGTAT-----------------CCCGAAGGCCACCATGGCAAT 294

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                ::::|||   |||      :::
Db        295 GCCGTGCTCTCCAAGCATCCGATCCTCGCCCAC-------------------------- 327

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu------ 99
                        |||||||||            |||:::|||:::|||||||||
Db        328 ------------CGCAACCTC---------GACGTCTCGGTGGCGGGCAACGAGGAACGC 366

Qy        100 ----------------------------------LeuCysLeuMetThrSerHisLeu 107
                                               |||      :::      ||||||
Db        367 GGCCTGTTGCATGCGGTGATCGACATCGGCCGTCCGCTGCACGCGGTCTGCGTGCACCTG 426

Qy        108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
                          |||:::|||      ||||||:::::::::   :::
Db        427 GGATTG---------CGCGAGGCGCAGCGCCAGCGCCAGTTGCAACTGCTCTGCGAACTG 477

Qy        128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn--------- 144
                :::          ||||||          |||:::   |||||||||   |||
Db        478 GTGGCGGGCCTGCCGGAGCGCGAGCCGGTGGTGGTCGCCGGCGACTTCAACGACTGGCGC 537

Qy        145 ---LeuArgAspArgGluValThrArgCysGly 154
                    |||:::   :::      ||||||
Db        538 CAGCGCGCTGACCAACTGCTCGCCGGTTGCGGC 570

RESULT 49
US-09-252-991A-8199/c
; Sequence 8199, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 8199
;   LENGTH: 873
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8199

Alignment Scores:
Pred. No.:                0.696          Length:          873
Score:                    85.00          Matches:         46
Percent Similarity:       34.6%          Conservative:    20
Best Local Similarity:    24.1%          Mismatches:      51
Query Match:              7.1%           Indels:          74
DB:                       3              Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-252-991A-8199 (1-873)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              |||       |||   |||       |||   |||
Db        702 CTGCCGGAGTTGCGCGAAGCGGTGCGCAGCGTC--------------------GGTGCC 664

Qy         26 AspValIlePheLeuGlnGluValIle---------------------------------- 34
              |||::::::|||||||||||||||:::
Db        663 GACCTGGTGTTCCTCCAGGAGGTCCTCGGCTCCCATAGCCTGCACGCCGCGCGCCTGCCG 604

Qy         35 ------------ProProTyrTyrSerTyrLeuLysLys---------------------- 43
                        |||   |||   ||||||
Db        603 TCCTGGCCGCCGGCGCCGCAATACGAGTATCTGGCCGACAGCATGTGGCCGCAGTTCGCC 544

Qy         44 ------ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                  |||::::::   |||                       ||||||:::
Db        543 TACGGGCGCAACGCGGTGTAT-----------------CCCGAAGGCCACCATGGCAAT 502

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                ::::|||   |||      :::
Db        501 GCCGTGCTCTCCAAGCATCCGATCCTCGCCCAC-------------------------- 469

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGlu------ 99
                        |||||||||            |||:::|||:::|||||||||
Db        468 ------------CGCAACCTC---------GACGTCTCGGTGGCGGGCAACGAGGAACGC 430

Qy        100 ----------------------------------LeuCysLeuMetThrSerHisLeu 107
                                               |||      :::      ||||||
Db        429 GGCCTGTTGCATGCGGTGATCGACATCGGCCGTCCGCTGCACGCGGTCTGCGTGCACCTG 370

Qy        108 GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLys 127
                          |||:::|||      ||||||:::::::::   :::
Db        369 GGATTG---------CGCGAGGCGCAGCGCCAGCGCCAGTTGCAACTGCTCTGCGAACTG 319

Qy        128 MetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn--------- 144
                :::          ||||||          |||:::   |||||||||   |||
Db        318 GTGGCGGGCCTGCCGGAGCGCGAGCCGGTGGTGGTCGCCGGCGACTTCAACGACTGGCGC 259

Qy        145 ---LeuArgAspArgGluValThrArgCysGly 154
                    |||:::   :::      ||||||
Db        258 CAGCGCGCTGACCAACTGCTCGCCGGTTGCGGC 226

RESULT 50
US-08-735-545-20/c
; Sequence 20, Application US/08735545
; Patent No. 6025131
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Van Dyk, Tina K.
;    APPLICANT:  LaRossa, Robert Alan
;    TITLE OF INVENTION:  A Facile Method for
;    TITLE OF INVENTION:  Identifying Regulated
;    TITLE OF INVENTION:  Promoters
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS
;      ADDRESSEE:  AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 3.1
;      SOFTWARE:  MICROSOFT WORD 2.0C
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/735,545
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA A.
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9989
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  410 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      STRAIN:  dpd3505 lower
US-08-735-545-20

Alignment Scores:
Pred. No.:                0.259          Length:          410
Score:                    84.50          Matches:         32
Percent Similarity:       31.7%          Conservative:    14
Best Local Similarity:    22.1%          Mismatches:      46
Query Match:              7.1%           Indels:          53
DB:                       3              Gaps:            4

US-10-757-745-2_COPY_121_347 (1-227) x US-08-735-545-20 (1-410)

Qy          76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
               |||:::|||   |||   |||          :::::::::|||||||||:::   :::
Db         278 ATTGTGCCG---CCGATGACCGGAAAAGCGATTCATGTGATGTGCGTACATCTGGGCCTG 222

Qy          96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
                                                                      |||
Db         221 ----------------------------------------------------CGTGAGGCG 213

Qy         116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
                  |||      ||||||   |||:::   :::   :::   |||   |||:::
Db         212 CACCGTCAGGCGCAGCTTGCGATGCTCGCCGAATGGGTGAATGAGCTACCGGACGGCGAA 153

Qy         136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
                  |||:::   |||||||||   |||
Db         152 CCGGTATTGGTGGCGGGTGATTTCAAT-------------------------------- 126

Qy         156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175

Db         126 ---------------------------------------------------------- 126

Qy         176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195

               |||       :::   |||       |||   :::   |||
Db         125 GACTGGCGGCAAAAAGCTAATCATCCGTTAAAAGTGCAGGCCGGA------------CTG 78

Qy         196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
                  |||   ||||||   ||||||                                |||
Db          77 GATGAGATTTTTACCCGCGCCCACGGACGCCCGGCGCGCACGTTTCCGGTGCAATTTCCT 18

Qy         216 LeuLeuGlyLeuGlu 220
                  ||||||   |||:::
Db          17 CTACTACGACTGGAC 3

RESULT 51
US-09-449-083-20/c
; Sequence 20, Application US/09449083
; Patent No. 6194159
;  GENERAL INFORMATION:
;    APPLICANT:  Van Dyk, Tina K.
;    APPLICANT:  LaRossa, Robert Alan
;    TITLE OF INVENTION:  A Facile Method for
;    TITLE OF INVENTION:  Identifying Regulated
;    TITLE OF INVENTION:  Promoters
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS
;      ADDRESSEE:  AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 3.1
;      SOFTWARE:  MICROSOFT WORD 2.0C
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/449,083
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA A.
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CR-9989
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  410 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      STRAIN:  dpd3505 lower
US-09-449-083-20

Alignment Scores:
Pred. No.:                0.259          Length:          410
Score:                    84.50          Matches:         32
Percent Similarity:       31.7%          Conservative:    14
Best Local Similarity:    22.1%          Mismatches:      46
Query Match:              7.1%           Indels:          53
DB:                       3              Gaps:            4

US-10-757-745-2_COPY_121_347 (1-227) x US-09-449-083-20 (1-410)

Qy          76 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 95
               |||:::|||   |||   |||          :::::::::|||||||||:::   :::
Db         278 ATTGTGCCG---CCGATGACCGGAAAAGCGATTCATGTGATGTGCGTACATCTGGGCCTG 222

Qy          96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
                                                                      |||
```

```
Db        221 ------------------------------------------------CGTGAGGCG 213

Qy        116 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 135
              |||      ||||||  |||:::  :::  :::  |||  |||:::
Db        212 CACCGTCAGGCGCAGCTTGCGATGCTCGCCGAATGGGTGAATGAGCTACCGGACGGCGAA 153

Qy        136 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCysGlyGly 155
              |||::  |||||||||  |||
Db        152 CCGGTATTGGTGGCGGGTGATTTCAAT------------------------------- 126

Qy        156 LeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyr 175

Db        126 ------------------------------------------------------------ 126

Qy        176 ThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLysLeuArgPhe 195
              |||     :::  |||     |||  :::  |||
Db        125 GACTGGCGGCAAAAAGCTAATCATCCGTTAAAAGTGCAGGCCGGA-----------CTG 78

Qy        196 AspArgIlePhePheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAsp 215
              |||  ||||||  ||||||                          |||
Db         77 GATGAGATTTTTACCCGCGCCCACGGACGCCCGGCGCGCACGTTTCCGGTGCAATTTCCT 18

Qy        216 LeuLeuGlyLeuGlu 220
              ||||||  |||:::
Db         17 CTACTACGACTGGAC 3

RESULT 52
US-09-949-016-16151
; Sequence 16151, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16151
;   LENGTH: 146401
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16151

Alignment Scores:
Pred. No.:              1.73e+03        Length:         146401
Score:                  84.50           Matches:        47
Percent Similarity:     40.7%           Conservative:   34
Best Local Similarity:  23.6%           Mismatches:     83
Query Match:            7.1%            Indels:         35
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-16151 (1-146401)

Qy          1 IleAspGlyLeuAspLeuAsn------------AsnLeuSerGluArgAlaArgGlyVal 16
              |||      |||  ||||||             |||         :::         :::
Db     137807 ATAACAAAATTAACTTTAAATGTATATGGACTAAATGCTCCAATTAAAAGACACAGACTG 137866

Qy         17 CysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProPro 36
              :::::::::             |||  |||      :::||||||
Db     137867 GCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA------------ 137914

Qy         37 TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGlu 56
              ::::::|||      |||  ::::::         |||  |||      :::
Db     137915 ------ACCCATCTCACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAG 137968

Qy         57 GlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIle 76
              |||             ||||||:::  |||  :::         :::
Db     137969 ATCTACCAAGCAAATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAA 138028

Qy         77 IleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSer 96
              |||      ||||||:::  ||||||            ::::::  |||
Db     138029 ACAGACTTTAAACCAACAAAGATCAAAAGAAACAAAGAAGGCCATTACATAATGGTAAAG 138088

Qy         97 GlyAsn------------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 112
              |||:::            ||||||  ::::::      :::      ::|||
Db     138089 GGATCAATTCAACAAGAAGAGCTAACTATCCTAAATATATATGCACCCAATACA------ 138142

Qy        113 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 132
              |||  |||       :::|||  ||||||      :::|||
Db     138143 ---GGAGCACCCAGATTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTA 138193

Qy        133 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeu-------------------- 145
              ||||||  |||:::|||      ||||||  |||
Db     138194 GAATCCCACACATTAATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGA 138250

Qy        146 --------------------ArgAspArgGluValThrArgCysGlyGlyLeuProAsn 158
                                  :::|||  :::
Db     138251 TCAACGAGATAAAAAGTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGCG 138310

Qy        159 AsnIleValAspValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::::::::|||::::::      |||      ||||||      :::|||||:::
Db     138311 GACCTAATAGACATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 138364

RESULT 53
US-09-949-016-12008
; Sequence 12008, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12008
;   LENGTH: 331814
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(331814)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12008

Alignment Scores:
Pred. No.:              5.88e+03        Length:         331814
Score:                  84.50           Matches:        48
Percent Similarity:     40.6%           Conservative:   41
Best Local Similarity:  21.9%           Mismatches:     87
Query Match:            7.1%            Indels:         43
DB:                     3               Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-12008 (1-331814)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::         :::   :::::::::
Db     185345 TTAAATGTAAATGGCCTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 185404

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                    |||   |||      :::||||||                  :::||||||
```

```
Db     185405 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCTATCTC 185446

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61

Db     185447 ATCTGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 185506

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProProSer 81

Db     185507 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 185566

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98

Db     185567 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 185626

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117

Db     185627 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 185677

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137

Db     185678 TTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGCCTTAGACTCCCATACAGTA 185731

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145

Db     185732 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATGAATGAGACAGAAA 185788

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163

Db     185789 GTTAACAAGGATATCCAGAAATTGAATTCAGCTCTGCACCAAGCAGACCTAATACACATC 185848

Qy        164 TrpGluPheLeuGlyLysProLys----------------------------HisCys 173

Db     185849 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTTTTTTCAGCACCACACTGC 185905

Qy        174 GlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192

Db     185906 ACCTATTCCAAAATTGACCACATAGTTGGAAGTAAAGCACTCCTCAGCAAATGTAAA 185962
```

```
RESULT 54
US-09-949-016-17056
; Sequence 17056, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17056
;   LENGTH: 331814
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(331814)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17056

Alignment Scores:
Pred. No.:              5.88e+03        Length:         331814
Score:                  84.50           Matches:        48
Percent Similarity:     40.6%           Conservative:   41
Best Local Similarity:  21.9%           Mismatches:     87
Query Match:            7.1%            Indels:         43
```

```
DB:                     3               Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-17056 (1-331814)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21

Db     185345 TTAAATGTAAATGGCCTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 185404

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41

Db     185405 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCTATCTC 185446

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61

Db     185447 ATCTGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 185506

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProProSer 81

Db     185507 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCA 185566

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98

Db     185567 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 185626

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117

Db     185627 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 185677

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137

Db     185678 TTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGCCTTAGACTCCCATACAGTA 185731

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145

Db     185732 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATGAATGAGACAGAAA 185788

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163

Db     185789 GTTAACAAGGATATCCAGAAATTGAATTCAGCTCTGCACCAAGCAGACCTAATACACATC 185848

Qy        164 TrpGluPheLeuGlyLysProLys----------------------------HisCys 173

Db     185849 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTTTTTTCAGCACCACACTGC 185905

Qy        174 GlnTyrThrTrpAspThrGlnMetAsnSerAsnLeuGlyIleThrAlaAlaCysLys 192

Db     185906 ACCTATTCCAAAATTGACCACATAGTTGGAAGTAAAGCACTCCTCAGCAAATGTAAA 185962
```

```
RESULT 55
US-09-134-000C-2365
; Sequence 2365, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2365
;   LENGTH: 804
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-2365

Alignment Scores:
Pred. No.:              0.822           Length:         804
Score:                  84.00           Matches:        47
Percent Similarity:     37.6%           Conservative:   33
```

```
Best Local Similarity:  22.1%          Mismatches:    66
Query Match:            7.0%           Indels:        67
DB:                     3              Gaps:          10

US-10-757-745-2_COPY_121_347 (1-227) x US-09-134-000C-2365 (1-804)

Qy          9 LeuSerGluArgAlaArgGly---------ValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::|||   ||||||                    ||||||      |||
Db        256 TTATCACAACGTAAACGTGGCGAATTCATTACAATGGATTACTTAGACGATTATCGTGTA 315

Qy         26 AspValIlePhe---------LeuGlnGluValIleProProTyrTyrSerTyrLeuLys 42
              :::|||::::::         |||   |||::::::      ::::::      ||||||
Db        316 AACGTAGTTTATGAAATTCCGTTATCTGAAATCGTGTTTGACTTTTTCGATAAATTGAAA 375

Qy         43 LysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIle 62
                  :::      |||      :::                ||||||   |||:::
Db        376 TCAAGTACAAAAGGCTATGCATCCTTAGATTACGAAATGGCTGGCTATCGTACCAGCCGC 435

Qy         63 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 82
              ::::::|||          :::   |||   ::::::::::::
Db        436 CTAGTGAAAATGGAT---ATTCTATTAAATGCTGAAAAAGTGGATGCTTTAAGC------ 486

Qy         83 LysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu 102
                                  ||||||
Db        487 -----------------TTTATTGTTCAC---------------------------- 498

Qy        103 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg-----------Met 118
                                  |||      |||   ||||||             :::
Db        499 ----------------------CGAGATTTCGCATTTGAGCGTGGTAAAGCGATTGTT 534

Qy        119 AsnGlnLeuLysMetValLeuLysLysMetGln---GluAlaProGluSerAlaThrVal 137
                 :::||||||   ::::::   :::   |||   |||   |||       |||   :::
Db        535 GAGAAACTGAAAAAACTAATTCCACGTCAACAGTTTGAAGTCCCAGTTCAAGCGGCGATT 594

Qy        137 ------------------------------------------------------------ 137

Db        595 GGTCAAAAAATTGTGGCTCGTTCAGATATTAAAGCCTTACGCAAAAACGTACTGGCTAAA 654

Qy        138 IlePheAlaGlyAspThrAsn--------------Leu-ArgAspArgGluValThrAr 152
                 :::   ||||||   :::               ||| |||:::::::      ||
Db        655 TGCTATGGTGGCGATGTTTCTCGTAAACGTAAATTGTTAGAGAAACAAAAAGAAGGGAAG 714

Qy        152 gCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeuGlyLysProLysHi 172
              |   |||            |||   :::|||   |||:::||||||   ||||||
Db        715 AAACGGA---------TGAAACAAATTGGATCCGTGGAAGTTCCTCAAGAAGCCTTTATG 765

Qy        172 sCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
                          |||   |||::::::      |||
Db        766 GCGGTTCTGAAAATGGACGACCAAGATAACGCGAAAT 802
```

RESULT 56
US-09-949-016-16342/c
; Sequence 16342, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16342
; LENGTH: 17644
; TYPE: DNA
; ORGANISM: Human

US-09-949-016-16342

Alignment Scores:
Pred. No.: 83.8 Length: 17644
Score: 84.00 Matches: 40
Percent Similarity: 40.9% Conservative: 21
Best Local Similarity: 26.8% Mismatches: 53
Query Match: 7.0% Indels: 35
DB: 3 Gaps: 8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-16342 (1-17644)

```
Qy         82 ThrLysMetMetArgAsnLeuLeu---------CysValHisValAsnVal--------- 95
              |||:::   |||   :::|||          |||   |||   ::::::
Db       3185 ACACAAGCACAGAGACTCATGCTGGCAAACACATGCACACATACACACCTTCACACACAC 3126

Qy         96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
              :::             |||      |||   ||||||          |||   |||   |||
Db       3125 ACACGCACACGCACATGCACACACACGCCACACCTCTCAGTGCCTAGAGACCATGGAGCC 3066

Qy        116 GluArgMet-AsnGlnLeuLysMetValLeuLysLys------MetGlnGluAlaProGl 133
              :::|||:::              |||:::::::::      :::|||   |||||||
Db       3065 CAGCGCTTGCCCCCCAGCCCATACTGTAATCCAGGAAAACTGCCTTCAAACAGCCCCAGA 3006

Qy        133 uSerAlaThrValIlePheAla------GlyAspThrAsnLeuArgAspArgGluValTh 151
              ||||         :::|||          :::|||||||||   :::   |||
Db       3005 GAGT---------GTTTTCACTCCCTTATCTTTGTCTAATCTCAGGGGCAAAAGAGTGGA 2955

Qy        151 rArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------- 167
                       |||   :::            :::|||:::   :::
Db       2954 AGAGGGCCTTTCTCTTGATTCCTAGAGGTGGTTTATTTGGAAACTAGTGTGAATATGGGG 2895

Qy        168 ----------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAs 184
                        |||   |||   |||
Db       2894 GGCTCCCTGAGGGAGTCCTGCTCAC--------------------------------AC 2868

Qy        184 nLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
              |||          ||||||   :::|||           |||:::||||||   |||
Db       2867 ACTTGCAGAGCCAGCTGCAGACAGGCTG---------GGCATTTGGTTCAGGGTGGCTGG 2817

Qy        204 aGluGluGlyHisIleIleProArg 212
               :::   |||          ||||||
Db       2816 GGACAGCGGGCAAAGTTTCCCCAGG 2792
```

RESULT 57
US-09-949-016-11805/c
; Sequence 11805, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11805
; LENGTH: 17645
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11805

Alignment Scores:
Pred. No.: 83.8 Length: 17645
Score: 84.00 Matches: 40
Percent Similarity: 40.9% Conservative: 21

```
Best Local Similarity:  26.8%          Mismatches:    53
Query Match:            7.0%           Indels:        35
DB:                     3              Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-11805 (1-17645)

Qy          82 ThrLysMetMetArgAsnLeuLeu---------CysValHisValAsnVal--------- 95
               |||:::     |||   :::|||          |||   |||   ::::::
Db        3186 ACACAAGCACAGAGACTCATGCTGGCAAACACATGCACACATACACACCTTCACACACAC 3127

Qy          96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
               :::          |||      |||   ||||||          |||   |||   |||
Db        3126 ACACGCACACGCACATGCACACACACGCCACACCTCTCAGTGCCTAGAGACCATGGAGCC 3067

Qy          116 GluArgMet-AsnGlnLeuLysMetValLeuLysLys------MetGlnGluAlaProGl 133
                :::|||:::             |||:::::::::      :::|||   ||||||||
Db        3066 CAGCGCTTGCCCCCCAGCCCATACTGTAATCCAGGAAAACTGCCTTCAAACAGCCCCAGA 3007

Qy          133 uSerAlaThrValIlePheAla------GlyAspThrAsnLeuArgAspArgGluValTh 151
                ||||         :::|||               :::|||||||||   :::   |||
Db        3006 GAGT---------GTTTTCACTCCCTTATCTTTGTCTAATCTCAGGGGCAAAAGAGTGGA 2956

Qy          151 rArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------- 167
                    |||   :::          :::|||:::   :::
Db        2955 AGAGGGCCTTTCTCTTGATTCCTAGAGGTGGTTTATTTGGAAACTAGTGTGAATATGGGG 2896

Qy          168 ----------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAs 184
                         |||   |||   |||
Db        2895 GGCTCCCTGAGGGAGTCCTGCTCAC------------------------------AC 2869

Qy          184 nLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                ||| |||||| :::|||          |||:::||||||   |||
Db        2868 ACTTGCAGAGCCAGCTGCAGACAGGCTG---------GGCATTTGGTTCAGGGTGGCTGG 2818

Qy          204 aGluGluGlyHisIleIleProArg 212
                :::   |||         ||||||
Db        2817 GGACAGCGGGCAAAGTTTCCCCAGG 2793

RESULT 58
US-09-949-016-14188/c
; Sequence 14188, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14188
;   LENGTH: 17645
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14188

Alignment Scores:
Pred. No.:              83.8           Length:        17645
Score:                  84.00          Matches:       40
Percent Similarity:     40.9%          Conservative:  21
Best Local Similarity:  26.8%          Mismatches:    53
Query Match:            7.0%           Indels:        35
DB:                     3              Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14188 (1-17645)

Qy          82 ThrLysMetMetArgAsnLeuLeu---------CysValHisValAsnVal--------- 95
               |||:::     |||   :::|||          |||   |||   ::::::
Db        3185 ACACAAGCACAGAGACTCATGCTGGCAAACACATGCACACATACACACCTTCACACACAC 3126

Qy          96 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 115
               :::          |||      |||   ||||||          |||   |||   |||
Db        3125 ACACGCACACGCACATGCACACACACGCCACACCTCTCAGTGCCTAGAGACCATGGAGCC 3066

Qy          116 GluArgMet-AsnGlnLeuLysMetValLeuLysLys------MetGlnGluAlaProGl 133
                :::|||:::             |||:::::::::      :::|||   ||||||||
Db        3065 CAGCGCTTGCCCCCCAGCCCATACTGTAATCCAGGAAAACTGCCTTCAAACAGCCCCAGA 3006

Qy          133 uSerAlaThrValIlePheAla------GlyAspThrAsnLeuArgAspArgGluValTh 151
                ||||         :::|||               :::|||||||||   :::   |||
Db        3005 GAGT---------GTTTTCACTCCCTTATCTTTGTCTAATCTCAGGGGCAAAAGAGTGGA 2955

Qy          151 rArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu---------- 167
                    |||   :::          :::|||:::   :::
Db        2954 AGAGGGCCTTTCTCTTGATTCCTAGAGGTGGTTTATTTGGAAACTAGTGTGAATATGGGG 2895

Qy          168 ----------GlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAs 184
                         |||   |||   |||
Db        2894 GGCTCCCTGAGGGAGTCCTGCTCAC------------------------------AC 2868

Qy          184 nLeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhePheArgAlaAlaAl 204
                ||| |||||| :::|||          |||:::||||||   |||
Db        2867 ACTTGCAGAGCCAGCTGCAGACAGGCTG---------GGCATTTGGTTCAGGGTGGCTGG 2817

Qy          204 aGluGluGlyHisIleIleProArg 212
                :::   |||         ||||||
Db        2816 GGACAGCGGGCAAAGTTTCCCCAGG 2792

RESULT 59
US-09-245-281-44/c
; Sequence 44, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
;  APPLICANT: Bertin, John
;  TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
;  TITLE OF INVENTION: AND USES THEREOF
;  FILE REFERENCE: 07334/118001
;  CURRENT APPLICATION NUMBER: US/09/245,281
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: US 09/207,359
;  EARLIER FILING DATE: 1998-12-08
;  EARLIER APPLICATION NUMBER: US 09/099,041
;  EARLIER FILING DATE: 1998-06-17
;  EARLIER APPLICATION NUMBER: US 09/019,942
;  EARLIER FILING DATE: 1998-02-06
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
;   LENGTH: 32042
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-245-281-44

Alignment Scores:
Pred. No.:              205            Length:        32042
Score:                  84.00          Matches:       45
Percent Similarity:     42.6%          Conservative:  36
Best Local Similarity:  23.7%          Mismatches:    77
Query Match:            7.0%           Indels:        32
DB:                     3              Gaps:          8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-245-281-44 (1-32042)

Qy           6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
               ||||||:::      :::|||   |||   :::   |||:::::              |||
Db       16126 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 16070

Qy          26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                  |||      :::||||||                   ::::::|||      |||
```

```
Db      16069 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 16028

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db      16027 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 15968

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db      15967 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 15908

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
Db      15907 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 15848

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db      15847 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 15803

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db      15802 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 15746

Qy        142 AspThrAsnLeu----------------------------------------ArgAsp 147
Db      15745 GACTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGAT 15686

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
Db      15685 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 15626

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
Db      15625 ---CACCCCAAATCAACAGAATATACATTT 15599
```

RESULT 60
US-09-340-620A-63/c
; Sequence 63, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 32042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-620A-63

Alignment Scores:

Pred. No.:	205	Length:	32042
Score:	84.00	Matches:	45
Percent Similarity:	42.6%	Conservative:	36
Best Local Similarity:	23.7%	Mismatches:	77
Query Match:	7.0%	Indels:	32
DB:	3	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-340-620A-63 (1-32042)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
Db      16126 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 16070

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
Db      16069 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 16028

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db      16027 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 15968

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db      15967 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 15908

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
Db      15907 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 15848

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db      15847 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 15803

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db      15802 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 15746

Qy        142 AspThrAsnLeu----------------------------------------ArgAsp 147
Db      15745 GACTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGAT 15686

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
Db      15685 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 15626

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
Db      15625 ---CACCCCAAATCAACAGAATATACATTT 15599
```

RESULT 61
US-09-728-721-63/c
; Sequence 63, Application US/09728721
; Patent No. 6869775
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 32042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-728-721-63

Alignment Scores:

Pred. No.:	205	Length:	32042
Score:	84.00	Matches:	45
Percent Similarity:	42.6%	Conservative:	36
Best Local Similarity:	23.7%	Mismatches:	77
Query Match:	7.0%	Indels:	32
DB:	3	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-728-721-63 (1-32042)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::            |||
Db      16126 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 16070

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
              |||      :::||||||                 ::::::|||      |||
Db      16069 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 16028

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||      :::   |||                  |||
Db      16027 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 15968

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db      15967 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 15908

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::           ::::::   |||   |||:::            ||||||
Db      15907 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 15848

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||   |||          :::
Db      15847 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 15803

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db      15802 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 15746

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db      15745 GACTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGAT 15686

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::                    ::::::::::|||::::::      |||
Db      15685 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 15626

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                 ||||||      :::|||||||:::
Db      15625 ---CACCCCAAATCAACAGAATATACATTT 15599
```

RESULT 62
US-09-949-002-810/c
; Sequence 810, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 810
; LENGTH: 39553
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-810

Alignment Scores:

Pred. No.:	281	Length:	39553
Score:	84.00	Matches:	45
Percent Similarity:	42.6%	Conservative:	36
Best Local Similarity:	23.7%	Mismatches:	77
Query Match:	7.0%	Indels:	32
DB:	3	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-002-810 (1-39553)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
```

```
              ||||||:::      :::|||   |||   :::   |||::::::            |||
Db      20762 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 20706

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
              |||      :::||||||                 ::::::|||      |||
Db      20705 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 20664

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::         |||   |||      :::   |||                  |||
Db      20663 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 20604

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db      20603 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 20544

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::           ::::::   |||   |||::            ||||||
Db      20543 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 20484

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||   |||          :::
Db      20483 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 20439

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||         :::|||   |||:::|||      |||
Db      20438 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 20382

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                           :::|||
Db      20381 GACTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGAT 20322

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                 :::                    ::::::::::|||::::::      |||
Db      20321 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 20262

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                 ||||||      :::|||||||:::
Db      20261 ---CACCCCAAATCAACAGAATATACATTT 20235
```

RESULT 63
US-09-949-002-673/c
; Sequence 673, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 673
; LENGTH: 58111
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-673

Alignment Scores:

Pred. No.:	500	Length:	58111
Score:	84.00	Matches:	45
Percent Similarity:	42.6%	Conservative:	36
Best Local Similarity:	23.7%	Mismatches:	77
Query Match:	7.0%	Indels:	32
DB:	3	Gaps:	8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-002-673 (1-58111)

```
Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||   :::   |||::::::            |||
```

```
Db      39365 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 39309

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
Db      39308 TCAGTGTGCTGTATTCAGGAA----------------ACCCATCTCACGTGCAGAGAC 39267

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db      39266 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 39207

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db      39206 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 39147

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn----------GluLeuCys 101
Db      39146 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 39087

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db      39086 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 39042

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db      39041 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 38985

Qy        142 AspThrAsnLeu-------------------------------------------ArgAsp 147
Db      38984 GACTTTAACACCCCACTGTCAACATTAGACAGATCAATGAGACAGAAAGTCAACAAGGAT 38925

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
Db      38924 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 38865

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
Db      38864 ---CACCCCAAATCAACAGAATATACATTT 38838
```

RESULT 64
US-09-949-016-13335/c
; Sequence 13335, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13335
; LENGTH: 87567
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13335

Alignment Scores:

Pred. No.:	923	Length:	87567
Score:	84.00	Matches:	49
Percent Similarity:	47.2%	Conservative:	43
Best Local Similarity:	25.1%	Mismatches:	68
Query Match:	7.0%	Indels:	35
DB:	3	Gaps:	9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-13335 (1-87567)

```
Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
Db      41392 TTAAATGTAAATGGGCTAAATGCTCCAATTAAAAGACACAGACTGGCAAATTGGATAAAG 41333

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
Db      41332 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA----------------ACTCATCTC 41291

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
Db      41290 ACATGCAGAGACACACATAGGCTCAAAATAAAGGGATGGAGGAAGATCTACCAAGCAAAT 41231

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
Db      41230 GGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCC 41171

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
Db      41170 ACAAAGATCAAAAGAGACAAAGTAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 41111

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
Db      41110 GAAGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 41060

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
Db      41059 TTC------ATAAAGCAAGTCCTTAGAGACCTACAAAGAGACTTAGACTCCCACACAATA 41006

Qy        138 IlePheAlaGlyAsp------------------------ThrAsnLeuArgAspArgGl 149
Db      41005 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 40949

Qy        149 uValThrArgCysGly-GlyLeu--------------ProAsn-AsnIleValAspVal 163
Db      40948 GTTAACAAGGATATCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATC 40889

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 176
Db      40888 TACAGAACTCTC---CACCCCAAATCAACAGAATATACA 40853
```

RESULT 65
US-09-949-016-17367/c
; Sequence 17367, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17367
; LENGTH: 102520
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17367

Alignment Scores:

Pred. No.:	1.17e+03	Length:	102520
Score:	84.00	Matches:	45
Percent Similarity:	42.6%	Conservative:	36
Best Local Similarity:	23.7%	Mismatches:	77
Query Match:	7.0%	Indels:	32
DB:	3	Gaps:	8

```
US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-17367 (1-102520)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
Db      57383 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 57327

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
Db      57326 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 57285

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db      57284 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 57225

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db      57224 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 57165

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------------GluLeuCys 101
Db      57164 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 57105

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db      57104 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 57060

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db      57059 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 57003

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
Db      57002 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 56943

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
Db      56942 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 56883

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
Db      56882 ---CACCCCAAATCAACAGAATATACATTT 56856

RESULT 66
US-09-949-016-12448/c
; Sequence 12448, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12448
;   LENGTH: 102526
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12448

Alignment Scores:
Pred. No.:              1.17e+03        Length:         102526
Score:                  84.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
Best Local Similarity:  23.7%           Mismatches:     77
Query Match:            7.0%            Indels:         32

DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-12448 (1-102526)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
Db      57390 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 57334

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
Db      57333 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 57292

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
Db      57291 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 57232

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
Db      57231 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 57172

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------------GluLeuCys 101
Db      57171 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 57112

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
Db      57111 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 57067

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
Db      57066 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 57010

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
Db      57009 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 56950

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
Db      56949 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 56890

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
Db      56889 ---CACCCCAAATCAACAGAATATACATTT 56863

RESULT 67
US-09-949-016-14353/c
; Sequence 14353, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14353
;   LENGTH: 175236
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14353

Alignment Scores:
Pred. No.:              2.61e+03        Length:         175236
Score:                  84.00           Matches:        45
Percent Similarity:     42.6%           Conservative:   36
```

```
Best Local Similarity:  23.7%          Mismatches:     77
Query Match:            7.0%           Indels:         32
DB:                     3              Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-14353 (1-175236)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||  :::   |||:::::              |||
Db       9306 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 9250

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::||||||                  :::::::|||      |||
Db       9249 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 9208

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||       :::   |||                  |||
Db       9207 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 9148

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db       9147 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 9088

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::          ::::::   |||   |||:::            ||||||
Db       9087 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 9028

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||  |||         :::
Db       9027 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 8983

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||        :::|||   |||:::|||      |||
Db       8982 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 8926

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db       8925 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 8866

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::                          :::::::::|||:::::      |||
Db       8865 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 8806

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                    ||||||      :::||||||:::
Db       8805 ---CACCCCAAATCAACAGAATATACATTT 8779

RESULT 68
US-09-949-016-13527/c
; Sequence 13527, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13527
;   LENGTH: 206433
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13527

Alignment Scores:
Pred. No.:              3.34e+03       Length:        206433

Score:                  84.00          Matches:        45
Percent Similarity:     42.6%          Conservative:   36
Best Local Similarity:  23.7%          Mismatches:     77
Query Match:            7.0%           Indels:         32
DB:                     3              Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-13527 (1-206433)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::      :::|||   |||  :::   |||:::::              |||
Db     190472 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 190416

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::|||||||                  :::::|||      |||
Db     190415 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 190374

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||       :::   |||                  |||
Db     190373 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 190314

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::         :::         |||      ||||||:::
Db     190313 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 190254

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::          ::::::   |||   |||:::            ||||||
Db     190253 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACT 190194

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::      :::      :::|||          |||  |||         :::
Db     190193 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 190149

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||        :::|||   |||:::|||      |||
Db     190148 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 190092

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db     190091 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 190032

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::                          :::::::::|||:::::      |||
Db     190031 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 189972

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                    ||||||      :::||||||:::
Db     189971 ---CACCCCAAATCAACAGAGTATACATTT 189945

RESULT 69
US-09-949-016-12417/c
; Sequence 12417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12417
;   LENGTH: 254778
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-12417
```

```
Alignment Scores:
Pred. No.:                4.57e+03        Length:         254778
Score:                    84.00           Matches:        45
Percent Similarity:       42.6%           Conservative:   36
Best Local Similarity:    23.7%           Mismatches:     77
Query Match:              7.0%            Indels:         32
DB:                       3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-12417 (1-254778)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::    :::|||  |||   :::   |||::::::           |||
Db     190817 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 190761

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::||||||                  ::::::|||      |||
Db     190760 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 190719

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||       :::   |||                |||
Db     190718 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCCAATGGAAAACAAAAA 190659

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::           :::          |||      |||||:::
Db     190658 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 190599

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::           ::::::   |||   |||:::           ||||||
Db     190598 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAAGAGCTAACT 190539

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::     :::      :::|||          |||  |||      :::
Db     190538 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 190494

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||          :::|||   |||:::|||      |||
Db     190493 AAGCAAGTCCTCAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 190437

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db     190436 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 190377

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::                    ::::::::::|||::::::      |||
Db     190376 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 190317

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                 ||||||      :::|||||||:::
Db     190316 ---CACCCCAAATCAACAGAGTATACATTT 190290

RESULT 70
US-09-949-016-15974/c
; Sequence 15974, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15974
;   LENGTH: 294836
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(294836)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15974

Alignment Scores:
Pred. No.:                5.69e+03        Length:         294836
Score:                    84.00           Matches:        45
Percent Similarity:       42.6%           Conservative:   36
Best Local Similarity:    23.7%           Mismatches:     77
Query Match:              7.0%            Indels:         32
DB:                       3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-15974 (1-294836)

Qy          6 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 25
              ||||||:::    :::|||  |||   :::   |||::::::           |||
Db     269676 CTAAATTCTGCAATTAAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCA 269620

Qy         26 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 45
                |||      :::||||||                  ::::::|||      |||
Db     269619 TCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTCACGTGCAGAGAC 269578

Qy         46 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 65
              ::::::          |||   |||       :::   |||                |||
Db     269577 ACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAATGGAAAACAAAAA 269518

Qy         66 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 85
              |||:::   |||   :::           :::          |||      |||||:::
Db     269517 AAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAA 269458

Qy         86 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----------GluLeuCys 101
              |||:::           ::::::   |||   |||:::           ||||||
Db     269457 AGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAAGAGGAGCTAACT 269398

Qy        102 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 121
              ::::::     :::      :::|||          |||  |||      :::
Db     269397 ATCCTAAATATTTATGCACCCAATACA---------GGAGCACCCAGATTC------ATA 269353

Qy        122 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 141
              |||   ||||||      :::|||          :::|||   |||:::|||      |||
Db     269352 AAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACATTAATAATG---GGA 269296

Qy        142 AspThrAsnLeu-----------------------------------------ArgAsp 147
              |||   |||                                          :::|||
Db     269295 GACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAAGTCAACAAGGAT 269236

Qy        148 ArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspValTrpGluPheLeu 167
                :::                    ::::::::::|||::::::      |||
Db     269235 ACCCAGGAATTGAACTCAGCTCTGCACCAAGCAGACCTAATAGACATCTACAGAACTCTC 269176

Qy        168 GlyLysProLysHisCysGlnTyrThrTrp 177
                 ||||||      :::|||||||:::
Db     269175 ---CACCCCAAATCAACAGAATATACATTT 269149

RESULT 71
US-09-252-991A-14967
; Sequence 14967, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 14967
```

```
;   LENGTH: 840
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14967

Alignment Scores:
Pred. No.:                1.01          Length:          840
Score:                    83.50         Matches:         51
Percent Similarity:       34.7%         Conservative:    41
Best Local Similarity:    19.2%         Mismatches:      120
Query Match:              7.0%          Indels:          53
DB:                       3             Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-252-991A-14967 (1-840)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||          ::::::|||   |||       |||          :::   :::   :::
Db         34 ATCGTCTCTTTCAATATCAACGGTCTGCGCGCCCGGCCGCACCAGTTGGCGGCGCTGATC 93

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGlu------ValIleProProTyrTyr 38
                 :::   ||||||||||||   |||||||||      |||              :::
Db         94 GAGCGGCACCAGCCCGACGTGATCGGGCTGCAGGAAACCAAGGTCGCCGACGACCAGTTT 153

Qy         39 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 58
                    :::       :::       |||   :::            ::::::|||:::
Db        154 CCCCAGGCCGAGATCGAGGCGCTCGGCTACCACGTCCACTACCACGGGCAGAAAGGCCAC 213

Qy         59 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 78
              :::         :::|||   :::   :::      |||:::                |||
Db        214 TACGGCGTCGCCCTGCTCTCCCGC---CAGGCGCCCCTGGAACTGCATCGCGGCTTCCCG 270

Qy         79 PheProSerThrLysMetMetArgAsnLeuLeuCys---ValHisValAsnValSerGly 97
                    :::       |||       :::                  :::      |||
Db        271 AGCGACGGCGAGGAATCCCAGCGGCGCTTCATCTGGGGCACCTTCAGCGACGCCCAGGGC 330

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeu--------------------GluSerThr 110
              |||   :::   :::|||      :::                              |||
Db        331 AACCCGATCACCGTGATGAACGGTTACTTCCCGCAAGGCGAGAACCGCGAGCATCCGACC 390

Qy        111 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 130
              :::       |||   :::|||                  |||  :::
Db        391 AAGTTCCCGGCCAAGCAGCGCTTCTACGCCGATCTCCAGTCGCTGCTGGAGAACAGCTTC 450

Qy        131 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--------------- 145
              ||||||   |||   |||:::      ||||||   |||:::
Db        451 AAGCCCGAGCAAGCCCTGGTGGTGATG---GGCGACATCAATATCTCGCCGGAGGACTGC 507

Qy        146 ----------------------------ArgAspArgGluValThrArgCysGlyGly 155
                                           |||      :::      :::|||
Db        508 GACATCGGCATTGGCGAGGAAAACCGCAAGCGCTGGCTGAAGACCGGCAAGTGCAGCTTC 567

Qy        156 LeuProAsn--------------------------------AsnIleValAspValTrp 164
              ||||||                                  :::||||||   :::
Db        568 CTGCCGGAAGAGCGCGAATGGCTGGCTACGTTGAAGAACTGGGGCCTGGTCGACAGCTTC 627

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
                  |||             :::   :::::::::|||            :::
Db        628 CGGGCCCTGAATCCGGAGGTGAACGACCGCTTCAGCTGGTTCGACTACCGCAGCCGTGGC 687

Qy        185 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe--------------- 199
                                  ||||||   |||   |||
Db        688 TTCGAGGACAACCCCAAGCGCGGCCTGCGCATCGATGTGATCCTGGCCAGCCGGGTGCTG 747

Qy        200 ---------PheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeu 216
                    |||:::   |||      :::                        ||||||
Db        748 CAGTCGCGCTTCAAGGACGCCGGCATCGAC--------------------TACGACCTG 786

Qy        217 LeuGlyLeuGluLys 221
                |||:::||||||
Db        787 CGCGGCATGGAGAAG 801

RESULT 72
US-09-252-991A-14783/c
```

```
; Sequence 14783, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 14783
;   LENGTH: 1149
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14783

Alignment Scores:
Pred. No.:                1.62          Length:          1149
Score:                    83.50         Matches:         51
Percent Similarity:       34.7%         Conservative:    41
Best Local Similarity:    19.2%         Mismatches:      120
Query Match:              7.0%          Indels:          53
DB:                       3             Gaps:            9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-252-991A-14783 (1-1149)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||          ::::::|||   |||       |||          :::   :::   :::
Db        810 ATCGTCTCTTTCAATATCAACGGTCTGCGCGCCCGGCCGCACCAGTTGGCGGCGCTGATC 751

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGlu------ValIleProProTyrTyr 38
                 :::   ||||||||||||   |||||||||      |||              :::
Db        750 GAGCGGCACCAGCCCGACGTGATCGGGCTGCAGGAAACCAAGGTCGCCGACGACCAGTTT 691

Qy         39 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 58
                    :::       :::       |||   :::            ::::::|||:::
Db        690 CCCCAGGCCGAGATCGAGGCGCTCGGCTACCACGTCCACTACCACGGGCAGAAAGGCCAC 631

Qy         59 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 78
              :::         :::|||   :::   :::      |||:::                |||
Db        630 TACGGCGTCGCCCTGCTCTCCCGC---CAGGCGCCCCTGGAACTGCATCGCGGCTTCCCG 574

Qy         79 PheProSerThrLysMetMetArgAsnLeuLeuCys---ValHisValAsnValSerGly 97
                    :::       |||       :::                  :::      |||
Db        573 AGCGACGGCGAGGAATCCCAGCGGCGCTTCATCTGGGGCACCTTCAGCGACGCCCAGGGC 514

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeu--------------------GluSerThr 110
              |||   :::   :::|||      :::                              |||
Db        513 AACCCGATCACCGTGATGAACGGTTACTTCCCGCAAGGCGAGAACCGCGAGCATCCGACC 454

Qy        111 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 130
              :::       |||   :::|||                  |||  :::
Db        453 AAGTTCCCGGCCAAGCAGCGCTTCTACGCCGATCTCCAGTCGCTGCTGGAGAACAGCTTC 394

Qy        131 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--------------- 145
              ||||||   |||   |||:::      ||||||   |||:::
Db        393 AAGCCCGAGCAAGCCCTGGTGGTGATG---GGCGACATCAATATCTCGCCGGAGGACTGC 337

Qy        146 ----------------------------ArgAspArgGluValThrArgCysGlyGly 155
                                           |||      :::      :::|||
Db        336 GACATCGGCATTGGCGAGGAAAACCGCAAGCGCTGGCTGAAGACCGGCAAGTGCAGCTTC 277

Qy        156 LeuProAsn--------------------------------AsnIleValAspValTrp 164
              ||||||                                  :::||||||   :::
Db        276 CTGCCGGAAGAGCGCGAATGGCTGGCTACGTTGAAGAACTGGGGCCTGGTCGACAGCTTC 217

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
                  |||             :::   :::::::::|||            :::
Db        216 CGGGCCCTGAATCCGGAGGTGAACGACCGCTTCAGCTGGTTCGACTACCGCAGCCGTGGC 157
```

```
Qy        185 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe--------------- 199
                                          ||||||  |||  |||
Db        156 TTCGAGGACAACCCCAAGCGCGGCCTGCGCATCGATGTGATCCTGGCCAGCCGGGTGCTG 97

Qy        200 ---------PheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeu 216
                      |||:::   |||      :::                          ||||||
Db         96 CAGTCGCGCTTCAAGGACGCCGGCATCGAC--------------------TACGACCTG 58

Qy        217 LeuGlyLeuGluLys 221
                 |||:::||||||
Db         57 CGCGGCATGGAGAAG 43

RESULT 73
US-09-252-991A-14833
; Sequence 14833, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 14833
;   LENGTH: 1461
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14833

Alignment Scores:
Pred. No.:              2.32          Length:         1461
Score:                  83.50         Matches:        51
Percent Similarity:     34.7%         Conservative:   41
Best Local Similarity:  19.2%         Mismatches:     120
Query Match:            7.0%          Indels:         53
DB:                     3             Gaps:           9

US-10-757-745-2_COPY_121_347 (1-227) x US-09-252-991A-14833 (1-1461)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              |||        :::::::|||   |||      |||       :::   :::   :::
Db         77 ATCGTCTCTTTCAATATCAACGGTCTGCGCGCCCGGCCGCACCAGTTGGCGGCGCTGATC 136

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGlu------ValIleProProTyrTyr 38
                  :::   |||||||||||   |||||||||      |||          :::
Db        137 GAGCGGCACCAGCCCGACGTGATCGGGCTGCAGGAAACCAAGGTCGCCGACGACCAGTTT 196

Qy         39 SerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 58
                    :::     :::      |||   :::          ::::::|||:::
Db        197 CCCCAGGCCGAGATCGAGGCGCTCGGCTACCACGTCCACTACCACGGGCAGAAAGGCCAC 256

Qy         59 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIlePro 78
              :::         :::|||   :::   :::      |||:::               |||
Db        257 TACGGCGTCGCCCTGCTCTCCCGC---CAGGCGCCCCTGGAACTGCATCGCGGCTTCCCG 313

Qy         79 PheProSerThrLysMetMetArgAsnLeuLeuCys---ValHisValAsnValSerGly 97
                    :::      |||      :::          :::      |||
Db        314 AGCGACGGCGAGGAATCCCAGCGGCGCTTCATCTGGGGCACCTTCAGCGACGCCCAGGGC 373

Qy         98 AsnGluLeuCysLeuMetThrSerHisLeu--------------------GluSerThr 110
              |||  :::  :::|||      :::                          |||
Db        374 AACCCGATCACCGTGATGAACGGTTACTTCCCGCAAGGCGAGAACCGCGAGCATCCGACC 433

Qy        111 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 130
              :::      |||   :::|||                    |||   :::
Db        434 AAGTTCCCGGCCAAGCAGCGCTTCTACGCCGATCTCCAGTCGCTGCTGGAGAACAGCTTC 493

Qy        131 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeu--------------- 145
                 ||||||   |||   |||:::      ||||||   |||:::
Db        494 AAGCCCGAGCAAGCCCTGGTGGTGATG---GGCGACATCAATATCTCGCCGGAGGACTGC 550

Qy        146 ------------------------------ArgAspArgGluValThrArgCysGlyGly 155
                                                |||      :::      :::|||
Db        551 GACATCGGCATTGGCGAGGAAAACCGCAAGCGCTGGCTGAAGACCGGCAAGTGCAGCTTC 610

Qy        156 LeuProAsn--------------------------------AsnIleValAspValTrp 164
                 ||||||                                :::||||||   :::
Db        611 CTGCCGGAAGAGCGCGAATGGCTGGCTACGTTGAAGAACTGGGGCCTGGTCGACAGCTTC 670

Qy        165 GluPheLeuGlyLysProLysHisCysGlnTyrThrTrpAspThrGlnMetAsnSerAsn 184
                 |||            :::   ::::::::::|||             :::
Db        671 CGGGCCCTGAATCCGGAGGTGAACGACCGCTTCAGCTGGTTCGACTACCGCAGCCGTGGC 730

Qy        185 LeuGlyIleThrAlaAlaCysLysLeuArgPheAspArgIlePhe--------------- 199
                                     ||||||  |||  |||
Db        731 TTCGAGGACAACCCCAAGCGCGGCCTGCGCATCGATGTGATCCTGGCCAGCCGGGTGCTG 790

Qy        200 ---------PheArgAlaAlaAlaGluGluGlyHisIleIleProArgSerLeuAspLeu 216
                      |||:::   |||      :::                          ||||||
Db        791 CAGTCGCGCTTCAAGGACGCCGGCATCGAC--------------------TACGACCTG 829

Qy        217 LeuGlyLeuGluLys 221
                 |||:::||||||
Db        830 CGCGGCATGGAGAAG 844

RESULT 74
US-09-949-016-12574
; Sequence 12574, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12574
;   LENGTH: 72843
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(72843)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12574

Alignment Scores:
Pred. No.:              810           Length:         72843
Score:                  83.50         Matches:        44
Percent Similarity:     41.8%         Conservative:   37
Best Local Similarity:  22.7%         Mismatches:     80
Query Match:            7.0%          Indels:         33
DB:                     3             Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-12574 (1-72843)

Qy          1 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 20
              ::::::||||||         |||  |||   :::      :::   ::::::::
Db      18413 GTAAATGGACTA---------AATGCTTCAATCAAAAGACACAGACTGGCAAATTGGATA 18463

Qy         21 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 40
                        |||   |||      :::||||||                  ::::::
```

```
Db      18464 AAGAGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCAT 18505

Qy         41 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 60
              |||      |||   ::::::          |||   |||       :::   |||
Db      18506 CTCACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCA 18565

Qy         61 AlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPhePro 80
                    ||||||:::   |||   :::         :::          |||
Db      18566 AATGGAAAACAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAA 18625

Qy         81 SerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn------ 98
                 ||||||:::   |||:::          ::::::   |||   |||:::
Db      18626 CCAACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAA 18685

Qy         99 ------GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 116
                   ||||||   ::::::      :::      :::      |||      |||
Db      18686 CAGGAAGAGCTAACTATCCTAAATATATATGCACCCAACAATACAGGA------GCACCC 18739

Qy        117 ArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThr 136
              |||         :::|||   ||||||      :::|||         :::|||   |||
Db      18740 AGATTC------ATAAAGCAAGTCCTGAGTGACCTACAAAGAGACTTAGACTCCCACACA 18793

Qy        137 ValIlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              :::|||      ||||||   |||
Db      18794 ATAATAATT---GGAGACTTTAACACCCCACTGTCAATATTAGACAGATCAACGAGACAA 18850

Qy        146 ---------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAsp 162
                    :::|||   :::                        :::::::::|||
Db      18851 AAAGTTAACAAAGATAGCCAGGAATTGAACTCAGCTCTGCACCAAGCGGACCTAATAGAC 18910

Qy        163 ValTrpGluPheLeuGlyLysProLysHisCysGlnTyrThr 176
              ::::::      |||      ||||||      :::||||||
Db      18911 ATCTACAGAACTCTC---CACCCCAAATCAACAGAATATACA 18949

RESULT 75
US-09-949-016-11757/c
; Sequence 11757, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11757
;   LENGTH: 131631
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(131631)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11757

Alignment Scores:
Pred. No.:              1.96e+03        Length:         131631
Score:                  83.50           Matches:        44
Percent Similarity:     42.3%           Conservative:   38
Best Local Similarity:  22.7%           Mismatches:     79
Query Match:            7.0%            Indels:         33
DB:                     3               Gaps:           8

US-10-757-745-2_COPY_121_347 (1-227) x US-09-949-016-11757 (1-131631)

Qy          4 LeuAspLeuAsnAsnLeuSerGlu------ArgAlaArgGlyValCysSerTyrLeuAla 21
              |||::::::|||   |||:::         :::          :::   |||::::::
Db     116252 TTAAATGTAAATGGACTAAATTCTCCAATTAAAAGACACAGACTGGCAAGTTGGATAAAG 116193

Qy         22 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 41
                    |||   |||      :::||||||                    ::::::|||
Db     116192 AGTCAAGACCCATCAGTGTGCTGTATTCAGGAA-----------------ACCCATCTC 116151

Qy         42 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 61
                 |||   ::::::          |||   |||       :::   |||
Db     116150 ACGTGCAGAGACACACATAGGCTCAAAATAAAAGGATGGAGGAAGATCTACCAAGCAAAT 116091

Qy         62 IleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 81
                    ||||||:::   |||   :::         :::          |||
Db     116090 GGAAAACAAAAAAAAGGCAGGGGTTGCAATCCTAGTCTCTGATAAAACAGACTTTAAGCCA 116031

Qy         82 ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn--------- 98
                 ||||||:::   |||:::          ::::::   |||   |||:::
Db     116030 ACAAAGATCAAAAGAGACAAAGAAGGCCATTACATAATGGTAAAGGGATCAATTCAACAA 115971

Qy         99 ---GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 117
                   ||||||   ::::::      :::      :::|||              |||   |||
Db     115970 GAGGAGCTAACTATCCTAAATATATATGCACCCAATACA---------GGAGCACCCAGA 115920

Qy        118 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 137
                    :::|||   ||||||      :::|||         :::|||   |||:::
Db     115919 TTC------ATAAAGAAAGTCCTGAGTGACCTACAAAGAGACCTAGACTCCCACACATTA 115866

Qy        138 IlePheAlaGlyAspThrAsnLeu---------------------------------- 145
              |||      ||||||   |||
Db     115865 ATAATG---GGAGACTTTAACACCCCACTGTCAACATTAGACAGATCAACGAGACAGAAA 115809

Qy        146 ------ArgAspArgGluValThrArgCysGlyGlyLeuProAsnAsnIleValAspVal 163
                 :::|||   :::                        :::::::::|||:::
Db     115808 GTCAACAAGGATACCCAGGAATTGAACTCAGCTCTGCACCAAGGGGACCTAATAGACATC 115749

Qy        164 TrpGluPheLeuGlyLysProLysHisCysGlnTyrThrTrp 177
              :::      |||      ||||||      :::||||||:::
Db     115748 TACAGAACTCTC---CACCCCAAATCAACAGAATATACATTT 115710

Search completed: March 10, 2006, 22:26:14
Job time : 2750.15 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 20:06:44 ; Search time 10.2351 Seconds
(without alignments)
1563.065 Million cell updates/sec

Title: US-10-757-745-2_COPY_145_153
Perfect score: 45
Sequence: 1 PDVIFLQEV 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170331_4111/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss08 -USER=US10757745_@CGN_1_1_417_@runat_09032006_170331_4111 -NCPU=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	45	100.0	1079	3	US-09-118-554-63	Sequence 63, Appl
c	2	45	100.0	1079	3	US-09-118-627-63	Sequence 63, Appl
c	3	45	100.0	1079	3	US-09-602-877A-63	Sequence 63, Appl
	4	45	100.0	1920	3	US-09-697-863A-1	Sequence 1, Appli
	5	45	100.0	1948	3	US-09-620-312D-889	Sequence 889, App
	6	44	97.8	1312	3	US-09-697-863A-3	Sequence 3, Appli
	7	40	88.9	133559	3	US-09-949-016-15845	Sequence 15845, A
	8	37	82.2	913	3	US-09-270-767-13729	Sequence 13729, A
c	9	36	80.0	523	3	US-09-270-767-31503	Sequence 31503, A
c	10	36	80.0	768	3	US-09-489-039A-1842	Sequence 1842, Ap
	11	36	80.0	864	3	US-09-902-540-5281	Sequence 5281, Ap
	12	36	80.0	1068	3	US-09-170-496D-129	Sequence 129, App
	13	36	80.0	1068	3	US-09-170-496D-231	Sequence 231, App
	14	36	80.0	1104	3	US-09-902-540-3201	Sequence 3201, Ap
	15	36	80.0	1161	2	US-08-153-848-31	Sequence 31, Appl
	16	36	80.0	1161	6	PCT-US93-11153-31	Sequence 31, Appl
c	17	36	80.0	1369	3	US-09-270-767-15178	Sequence 15178, A
	18	36	80.0	2254	2	US-08-153-848-27	Sequence 27, Appl
	19	36	80.0	2254	3	US-09-299-843A-27	Sequence 27, Appl
	20	36	80.0	2254	3	US-09-088-337B-27	Sequence 27, Appl
	21	36	80.0	2254	6	PCT-US93-11153-27	Sequence 27, Appl
	22	36	80.0	3100	3	US-09-016-434-1442	Sequence 1442, Ap
	23	36	80.0	3100	3	US-09-917-254-17	Sequence 17, Appl
	24	36	80.0	3119	3	US-09-299-843A-31	Sequence 31, Appl
	25	36	80.0	3119	3	US-09-088-337B-31	Sequence 31, Appl
	26	36	80.0	15133	3	US-09-949-016-15001	Sequence 15001, A
	27	36	80.0	19455	3	US-09-902-540-1147	Sequence 1147, Ap
	28	36	80.0	30783	3	US-09-902-540-1258	Sequence 1258, Ap
c	29	36	80.0	145928	3	US-09-949-016-15444	Sequence 15444, A
c	30	35	77.8	819	3	US-09-328-352-3246	Sequence 3246, Ap
c	31	35	77.8	11879	3	US-09-949-016-14211	Sequence 14211, A
c	32	35	77.8	131978	3	US-09-949-016-13117	Sequence 13117, A
c	33	35	77.8	189560	3	US-09-949-016-17202	Sequence 17202, A
	34	35	77.8	1830121	3	US-09-557-884-1	Sequence 1, Appli
c	35	35	77.8	1830121	3	US-09-557-884-1	Sequence 1, Appli
	36	35	77.8	1830121	3	US-09-643-990A-1	Sequence 1, Appli
c	37	35	77.8	1830121	3	US-09-643-990A-1	Sequence 1, Appli
	38	35	77.8	1830121	3	US-10-158-865-1	Sequence 1, Appli
c	39	35	77.8	1830121	3	US-10-158-865-1	Sequence 1, Appli
c	40	34	75.6	485	3	US-09-621-976-1908	Sequence 1908, Ap
	41	34	75.6	582	3	US-09-252-991A-5517	Sequence 5517, Ap
	42	34	75.6	756	3	US-09-252-991A-8106	Sequence 8106, Ap
c	43	34	75.6	756	3	US-09-107-532A-2040	Sequence 2040, Ap
	44	34	75.6	828	3	US-09-540-236-1914	Sequence 1914, Ap
	45	34	75.6	864	3	US-09-902-540-5749	Sequence 5749, Ap
c	46	34	75.6	873	3	US-09-252-991A-8199	Sequence 8199, Ap
	47	34	75.6	951	3	US-09-902-540-3685	Sequence 3685, Ap
	48	34	75.6	963	3	US-09-252-991A-8063	Sequence 8063, Ap
	49	34	75.6	1164	3	US-09-252-991A-5541	Sequence 5541, Ap
c	50	34	75.6	1182	3	US-09-489-039A-82	Sequence 82, Appl
c	51	34	75.6	1263	3	US-09-252-991A-5682	Sequence 5682, Ap
c	52	34	75.6	1399	3	US-09-991-181-27	Sequence 27, Appl
c	53	34	75.6	1399	3	US-09-990-444-27	Sequence 27, Appl
c	54	34	75.6	1399	3	US-09-997-333-27	Sequence 27, Appl
c	55	34	75.6	1399	3	US-09-992-598-27	Sequence 27, Appl
c	56	34	75.6	2263	3	US-09-632-538C-16	Sequence 16, Appl
c	57	34	75.6	2277	3	US-09-632-538C-31	Sequence 31, Appl
c	58	34	75.6	2364	2	US-08-981-690-1	Sequence 1, Appli
c	59	34	75.6	16924	3	US-09-902-540-1178	Sequence 1178, Ap
	60	34	75.6	36470	3	US-08-311-731A-123	Sequence 123, App
	61	34	75.6	41171	3	US-08-311-731A-122	Sequence 122, App
c	62	34	75.6	54484	3	US-09-902-540-1272	Sequence 1272, Ap
c	63	34	75.6	116966	3	US-09-949-016-17557	Sequence 17557, A
	64	34	75.6	228896	3	US-09-949-016-17127	Sequence 17127, A
	65	34	75.6	269223	3	US-09-596-002-41	Sequence 41, Appl
c	66	34	75.6	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	67	34	75.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	68	33	73.3	247	3	US-09-513-999C-28064	Sequence 28064, A
c	69	33	73.3	507	3	US-09-252-991A-5110	Sequence 5110, Ap
c	70	33	73.3	601	3	US-09-949-016-182845	Sequence 182845,
c	71	33	73.3	601	3	US-09-949-002-3663	Sequence 3663, Ap
c	72	33	73.3	601	3	US-09-949-002-3664	Sequence 3664, Ap
c	73	33	73.3	601	3	US-09-949-002-3665	Sequence 3665, Ap
c	74	33	73.3	601	3	US-09-949-002-5508	Sequence 5508, Ap
c	75	33	73.3	601	3	US-09-949-002-5509	Sequence 5509, Ap

ALIGNMENTS

RESULT 1
US-09-118-554-63/c
; Sequence 63, Application US/09118554A

; Patent No. 6365348
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND
; TITLE OF INVENTION: METHODS FOR THEIR USE
; FILE REFERENCE: 210121.450C1
; CURRENT APPLICATION NUMBER: US/09/118,554A
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 08/988,255
; EARLIER FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-118-554-63

Alignment Scores:

Pred. No.:	0.696	Length:	1079
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-118-554-63 (1-1079)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024
```

RESULT 2
US-09-118-627-63/c
; Sequence 63, Application US/09118627A
; Patent No. 6379951
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER
; TITLE OF INVENTION: AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446C1
; CURRENT APPLICATION NUMBER: US/09/118,627A
; CURRENT FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-118-627-63

Alignment Scores:

Pred. No.:	0.696	Length:	1079
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-118-627-63 (1-1079)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024
```

RESULT 3
US-09-602-877A-63/c
; Sequence 63, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PastSEQ for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-63

Alignment Scores:

Pred. No.:	0.696	Length:	1079
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-602-877A-63 (1-1079)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024
```

RESULT 4
US-09-697-863A-1
; Sequence 1, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/09/697,863A
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1627)..(1627)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1108)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-09-697-863A-1

Alignment Scores:

Pred. No.:	1.39	Length:	1920
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-697-863A-1 (1-1920)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        452 CCAGATGTGATATTTCTACAGGAAGTT 478

RESULT 5
US-09-620-312D-889
; Sequence 889, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-09-620-312D-889

Alignment Scores:
Pred. No.:              1.42          Length:         1948
Score:                  45.00         Matches:        9
Percent Similarity:     100.0%        Conservative:   0
Best Local Similarity:  100.0%        Mismatches:     0
Query Match:            100.0%        Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-620-312D-889 (1-1948)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        471 CCAGATGTGATATTTCTACAGGAAGTT 497

RESULT 6
US-09-697-863A-3
; Sequence 3, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/09/697,863A
;  CURRENT FILING DATE:  2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (122)..(1234)
;   OTHER INFORMATION:
US-09-697-863A-3

Alignment Scores:
Pred. No.:              1.49          Length:         1312
Score:                  44.00         Matches:        8
Percent Similarity:     100.0%        Conservative:   1
Best Local Similarity:  88.9%         Mismatches:     0
Query Match:            97.8%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-697-863A-3 (1-1312)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||
Db        584 CCAGATGTGGTATTTCTACAGGAAGTT 610

RESULT 7
US-09-949-016-15845
; Sequence 15845, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15845
;   LENGTH: 133559
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(133559)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15845

Alignment Scores:
Pred. No.:              3.21e+03      Length:         133559
Score:                  40.00         Matches:        7
Percent Similarity:     100.0%        Conservative:   2
Best Local Similarity:  77.8%         Mismatches:     0
Query Match:            88.9%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-016-15845 (1-133559)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||:::|||:::
Db      65973 CCAGATGTTATTTTTCTCAGGGAAATT 65999

RESULT 8
US-09-270-767-13729
; Sequence 13729, Application US/09270767
```

; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13729
; LENGTH: 913
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-13729

Alignment Scores:

Pred. No.:	38.8	Length:	913
Score:	37.00	Matches:	8
Percent Similarity:	88.9%	Conservative:	0
Best Local Similarity:	88.9%	Mismatches:	1
Query Match:	82.2%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-270-767-13729 (1-913)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||   ||||||||||||
Db        818 CCCGATGTGATTTGTCTGCAAGAGGTA 844
```

RESULT 9
US-09-270-767-31503/c
; Sequence 31503, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31503
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-31503

Alignment Scores:

Pred. No.:	33.6	Length:	523
Score:	36.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-270-767-31503 (1-523)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||:::   |||
Db        367 CCCGATGTCATCTTCCTGGAGGCGGTC 341
```

RESULT 10
US-09-489-039A-1842/c
; Sequence 1842, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1842
; LENGTH: 768
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1842

Alignment Scores:

Pred. No.:	53.4	Length:	768
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	80.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-489-039A-1842 (1-768)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||:::|||||||||:::
Db        546 CCAGACGTTCTCTTCCTGCAGCAG 523
```

RESULT 11
US-09-902-540-5281
; Sequence 5281, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5281
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-5281

Alignment Scores:

Pred. No.:	61.5	Length:	864
Score:	36.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-902-540-5281 (1-864)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||:::   ||||||||||||
Db        199 CCGGACGTCGTCGCCTTGCAGGAGGTG 225
```

RESULT 12
US-09-170-496D-129
; Sequence 129, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 129
;   LENGTH: 1068
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-170-496D-129

Alignment Scores:
Pred. No.:                79.4          Length:         1068
Score:                    36.00         Matches:        6
Percent Similarity:       87.5%         Conservative:   1
Best Local Similarity:    75.0%         Mismatches:     1
Query Match:              80.0%         Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-170-496D-129 (1-1068)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::||||||||||||
Db        645 CCAGACGCTGTTTTCCTGCAAGAA 668

RESULT 13
US-09-170-496D-231
; Sequence 231, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
;  APPLICANT: Behan, Dominic P.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Liaw, Chen W.
;  TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0040
;  CURRENT APPLICATION NUMBER: US/09/170,496D
;  CURRENT FILING DATE:  1998-10-13
;  NUMBER OF SEQ ID NOS: 294
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 231
;   LENGTH: 1068
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-170-496D-231

Alignment Scores:
Pred. No.:                79.4          Length:         1068
Score:                    36.00         Matches:        6
Percent Similarity:       87.5%         Conservative:   1
Best Local Similarity:    75.0%         Mismatches:     1
Query Match:              80.0%         Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-170-496D-231 (1-1068)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::||||||||||||
Db        645 CCAGACGCTGTTTTCCTGCAAGAA 668

RESULT 14
US-09-902-540-3201
; Sequence 3201, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 3201
;   LENGTH: 1104
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-3201

Alignment Scores:
Pred. No.:                82.6          Length:         1104
Score:                    36.00         Matches:        7
Percent Similarity:       88.9%         Conservative:   1
Best Local Similarity:    77.8%         Mismatches:     1
Query Match:              80.0%         Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-902-540-3201 (1-1104)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   ||||||||||||
Db        241 CCGGATGTCGTCTGCCTGCAAGAGGTG 267

RESULT 15
US-08-153-848-31
; Sequence 31, Application US/08153848
; Patent No. 5759804
;  GENERAL INFORMATION:
;    APPLICANT:  Godiska, Ronald
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Schweikart, Vicki L.
;    TITLE OF INVENTION:  No. 5759804el Seven Transmembrane Receptors
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/153,848
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/977,452
;      FILING DATE:  17-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5759804and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  31794
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1161 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  7..80
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  94..1158
US-08-153-848-31

Alignment Scores:
Pred. No.:                87.8          Length:         1161
```

```
Score:                    36.00          Matches:       6
Percent Similarity:       87.5%          Conservative:  1
Best Local Similarity:    75.0%          Mismatches:    1
Query Match:              80.0%          Indels:        0
DB:                       2              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-153-848-31 (1-1161)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        738 CCAGACGCTGTTTTCCTGCAAGAA 761

RESULT 16
PCT-US93-11153-31
; Sequence 31, Application PC/TUS9311153
;  GENERAL INFORMATION:
;    APPLICANT:  Godiska, Ronald
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Schweikart, Vicki L.
;    TITLE OF INVENTION:  Novel Seven Transmembrane Receptors
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/11153
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/977,452
;      FILING DATE:  17-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Noland, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  31794
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1161 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  7..80
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  94..1158
PCT-US93-11153-31

Alignment Scores:
Pred. No.:                87.8           Length:        1161
Score:                    36.00          Matches:       6
Percent Similarity:       87.5%          Conservative:  1
Best Local Similarity:    75.0%          Mismatches:    1
Query Match:              80.0%          Indels:        0
DB:                       6              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x PCT-US93-11153-31 (1-1161)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        738 CCAGACGCTGTTTTCCTGCAAGAA 761

RESULT 17
US-09-270-767-15178/c
; Sequence 15178, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15178
;   LENGTH: 1369
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-15178

Alignment Scores:
Pred. No.:                107            Length:        1369
Score:                    36.00          Matches:       7
Percent Similarity:       88.9%          Conservative:  1
Best Local Similarity:    77.8%          Mismatches:    1
Query Match:              80.0%          Indels:        0
DB:                       3              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-270-767-15178 (1-1369)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||:::   |||
Db        367 CCCGATGTCATCTTCCTGGAGGCGGTC 341

RESULT 18
US-08-153-848-27
; Sequence 27, Application US/08153848
; Patent No. 5759804
; GENERAL INFORMATION:
;    APPLICANT:  Godiska, Ronald
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Schweikart, Vicki L.
;    TITLE OF INVENTION:  No. 5759804el Seven Transmembrane Receptors
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/153,848
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/977,452
;      FILING DATE:  17-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5759804and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  31794
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
```

```
;       TELEFAX:  (312) 474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  27:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2254 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  593..1657
US-08-153-848-27

Alignment Scores:
Pred. No.:                195           Length:        2254
Score:                    36.00         Matches:       6
Percent Similarity:       87.5%         Conservative:  1
Best Local Similarity:    75.0%         Mismatches:    1
Query Match:              80.0%         Indels:        0
DB:                       2             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-153-848-27 (1-2254)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db       1238 CCAGACGCTGTTTTCCTGCAAGAA 1261

RESULT 19
US-09-299-843A-27
; Sequence 27, Application US/09299843A
; Patent No. 6107475
;  GENERAL INFORMATION:
;    APPLICANT:  Godiska, Ronald
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Schweikart, Vicki L.
;    TITLE OF INVENTION:  No. 6107475el Seven Transmembrane Receptors
;    NUMBER OF SEQUENCES:  66
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/299,843A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/088,337
;      FILING DATE:  01-JUN-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/153,848
;      FILING DATE:  17-NOV-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/977,452
;      FILING DATE:  17-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jill E. Uhl
;      REGISTRATION NUMBER:  43,213
;      REFERENCE/DOCKET NUMBER:  27866/32059B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2254 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  593..1657
US-09-299-843A-27

Alignment Scores:
Pred. No.:                195           Length:        2254
Score:                    36.00         Matches:       6
Percent Similarity:       87.5%         Conservative:  1
Best Local Similarity:    75.0%         Mismatches:    1
Query Match:              80.0%         Indels:        0
DB:                       3             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-299-843A-27 (1-2254)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db       1238 CCAGACGCTGTTTTCCTGCAAGAA 1261

RESULT 20
US-09-088-337B-27
; Sequence 27, Application US/09088337B
; Patent No. 6348574
;   GENERAL INFORMATION:
;       APPLICANT: Godiska, Ronald
;                  Gray, Patrick W.
;                  Schweikart, Vicki L.
;       TITLE OF INVENTION: No. 6348574el Seven Transmembrane Receptors
;       NUMBER OF SEQUENCES: 66
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
;                       Borun
;            STREET: 6300 Sears Tower, 233 South Wacker Drive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: USA
;            ZIP: 60606
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/088,337B
;            FILING DATE: 01-Jun-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/153,848
;            FILING DATE: 17-NOV-1993
;            APPLICATION NUMBER: US 07/977,452
;            FILING DATE: 17-NOV-1992
;       ATTORNEY/AGENT INFORMATION:
;            NAME: No. 6348574and, Greta E.
;            REGISTRATION NUMBER: 35,302
;            REFERENCE/DOCKET NUMBER: 31794
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (312) 474-6300
;            TELEFAX: (312) 474-0448
;            TELEX: 25-3856
;  INFORMATION FOR SEQ ID NO: 27:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2254 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       FEATURE:
```

```
;             NAME/KEY:  CDS
;             LOCATION:  593..1657
;       SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-088-337B-27

Alignment Scores:
Pred. No.:                195             Length:         2254
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-088-337B-27 (1-2254)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db       1238 CCAGACGCTGTTTTCCTGCAAGAA 1261

RESULT 21
PCT-US93-11153-27
; Sequence 27, Application PC/TUS9311153
;  GENERAL INFORMATION:
;    APPLICANT:  Godiska, Ronald
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Schweikart, Vicki L.
;    TITLE OF INVENTION:  Novel Seven Transmembrane Receptors
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/11153
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/977,452
;      FILING DATE:  17-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Noland, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  31794
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2254 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  593..1657
PCT-US93-11153-27

Alignment Scores:
Pred. No.:                195             Length:         2254
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x PCT-US93-11153-27 (1-2254)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db       1238 CCAGACGCTGTTTTCCTGCAAGAA 1261

RESULT 22
US-09-016-434-1442
; Sequence 1442, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1442:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3100 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g665580
US-09-016-434-1442

Alignment Scores:
Pred. No.:                286             Length:         3100
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-016-434-1442 (1-3100)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        732 CCAGACGCTGTTTTCCTGCAAGAA 755

RESULT 23
US-09-917-254-17
```

```
; Sequence 17, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
;  APPLICANT: Mutter, George
;  APPLICANT:  Baak, Jan
;  TITLE OF INVENTION: Prognostic Classification of Breast Cancer
;  FILE REFERENCE: B0801/7224(JRV)
;  CURRENT APPLICATION NUMBER: US/09/917,254
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: US 60/222,093
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
;   LENGTH: 3100
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-917-254-17

Alignment Scores:
Pred. No.:                286             Length:         3100
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-917-254-17 (1-3100)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        732 CCAGACGCTGTTTTCCTGCAAGAA 755

RESULT 24
US-09-299-843A-31
; Sequence 31, Application US/09299843A
; Patent No. 6107475
;  GENERAL INFORMATION:
;    APPLICANT:  Godiska, Ronald
;    APPLICANT:  Gray, Patrick W.
;    APPLICANT:  Schweikart, Vicki L.
;    TITLE OF INVENTION:  No. 6107475el Seven Transmembrane Receptors
;    NUMBER OF SEQUENCES:  66
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/299,843A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 09/088,337
;      FILING DATE:  01-JUN-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/153,848
;      FILING DATE:  17-NOV-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/977,452
;      FILING DATE:  17-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jill E. Uhl
;      REGISTRATION NUMBER:  43,213
;      REFERENCE/DOCKET NUMBER:  27866/32059B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3119 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  7..80
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  94..1158
US-09-299-843A-31

Alignment Scores:
Pred. No.:                288             Length:         3119
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-299-843A-31 (1-3119)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        738 CCAGACGCTGTTTTCCTGCAAGAA 761

RESULT 25
US-09-088-337B-31
; Sequence 31, Application US/09088337B
; Patent No. 6348574
;   GENERAL INFORMATION:
;        APPLICANT: Godiska, Ronald
;                   Gray, Patrick W.
;                   Schweikart, Vicki L.
;        TITLE OF INVENTION: No. 6348574el Seven Transmembrane Receptors
;        NUMBER OF SEQUENCES: 66
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
;                        Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: USA
;             ZIP: 60606
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/088,337B
;             FILING DATE: 01-Jun-1998
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/153,848
;             FILING DATE: 17-NOV-1993
;             APPLICATION NUMBER: US 07/977,452
;             FILING DATE: 17-NOV-1992
;        ATTORNEY/AGENT INFORMATION:
;             NAME: No. 6348574and, Greta E.
;             REGISTRATION NUMBER: 35,302
;             REFERENCE/DOCKET NUMBER: 31794
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (312) 474-6300
;             TELEFAX: (312) 474-0448
;             TELEX: 25-3856
```

```
;   INFORMATION FOR SEQ ID NO: 31:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 3119 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  exon
;             LOCATION:  7..80
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  94..1158
;        SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-088-337B-31

Alignment Scores:
Pred. No.:              288             Length:         3119
Score:                  36.00           Matches:        6
Percent Similarity:     87.5%           Conservative:   1
Best Local Similarity:  75.0%           Mismatches:     1
Query Match:            80.0%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-088-337B-31 (1-3119)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        738 CCAGACGCTGTTTTCCTGCAAGAA 761

RESULT 26
US-09-949-016-15001
; Sequence 15001, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15001
;   LENGTH: 15133
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(15133)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15001

Alignment Scores:
Pred. No.:              1.93e+03        Length:         15133
Score:                  36.00           Matches:        7
Percent Similarity:     87.5%           Conservative:   0
Best Local Similarity:  87.5%           Mismatches:     1
Query Match:            80.0%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-016-15001 (1-15133)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||||||||||   |||
Db      15076 CCTGATGTCATATTCTTGATGGAG 15099

RESULT 27
US-09-902-540-1147
; Sequence 1147, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1147
;   LENGTH: 19455
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(19455)
;   OTHER INFORMATION: unsure at all n locations
US-09-902-540-1147

Alignment Scores:
Pred. No.:              2.61e+03        Length:         19455
Score:                  36.00           Matches:        7
Percent Similarity:     88.9%           Conservative:   1
Best Local Similarity:  77.8%           Mismatches:     1
Query Match:            80.0%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-902-540-1147 (1-19455)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   ||||||||||||
Db        294 CCGGATGTCGTCTGCCTGCAAGAGGTG 320

RESULT 28
US-09-902-540-1258
; Sequence 1258, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1258
;   LENGTH: 30783
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-1258

Alignment Scores:
Pred. No.:              4.53e+03        Length:         30783
Score:                  36.00           Matches:        7
Percent Similarity:     88.9%           Conservative:   1
Best Local Similarity:  77.8%           Mismatches:     1
Query Match:            80.0%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-902-540-1258 (1-30783)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
```

```
                  |||||||||:::   ||||||||||||
Db      23061 CCGGACGTCGTCGCCTTGCAGGAGGTG 23087
```

RESULT 29
US-09-949-016-15444/c
; Sequence 15444, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15444
; LENGTH: 145928
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(145928)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15444

Alignment Scores:

Pred. No.:	2.93e+04	Length:	145928
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	80.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-016-15444 (1-145928)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||::::::
Db     111424 CCAGATCTTCTGTTTTTGCAACAAATC 111398
```

RESULT 30
US-09-328-352-3246/c
; Sequence 3246, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3246
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3246

Alignment Scores:

Pred. No.:	97.8	Length:	819
Score:	35.00	Matches:	6
Percent Similarity:	88.9%	Conservative:	2
Best Local Similarity:	66.7%	Mismatches:	1
Query Match:	77.8%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-328-352-3246 (1-819)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||   |||
Db        676 CCGGATCTTCTGTTCCTGCAATCAGTA 650
```

RESULT 31
US-09-949-016-14211/c
; Sequence 14211, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14211
; LENGTH: 11879
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14211

Alignment Scores:

Pred. No.:	2.44e+03	Length:	11879
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	77.8%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-016-14211 (1-11879)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||:::||||||:::|||
Db      10400 CCTGATGTCCTTTTTCTAAGGGAG 10377
```

RESULT 32
US-09-949-016-13117/c
; Sequence 13117, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13117
; LENGTH: 131978
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13117

Alignment Scores:

Pred. No.:	4.39e+04	Length:	131978
Score:	35.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4

```
Best Local Similarity:  55.6%          Mismatches:     0
Query Match:            77.8%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-016-13117 (1-131978)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::::::::::|||||||||||
Db      71803 CCTGAAATTTTATATCTGCAGGAGGTC 71777

RESULT 33
US-09-949-016-17202/c
; Sequence 17202, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17202
;   LENGTH: 189560
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(189560)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17202

Alignment Scores:
Pred. No.:              6.75e+04       Length:         189560
Score:                  35.00          Matches:        7
Percent Similarity:     87.5%          Conservative:   0
Best Local Similarity:  87.5%          Mismatches:     1
Query Match:            77.8%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-016-17202 (1-189560)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||   ||||||||||||||||||
Db     185422 CCTTCTGTCATCTTTCTTCAGGAA 185399

RESULT 34
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
;   GENERAL INFORMATION:
;        APPLICANT: Fleischmann et al.
;        TITLE OF INVENTION: The Nucleotide sequence of
;                            the Haemophilus influenzae Rd Genome, Fragments
;                            Thereof, and Uses Thereof
;        NUMBER OF SEQUENCES: 1
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3 1/2 inch diskette
;             COMPUTER: Dell Pentium
;             OPERATING SYSTEM: MS DOS v6.22
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/557,884
;             FILING DATE: 25-Apr-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/476,102
;             FILING DATE: JUN-5-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Michelle S. Marks
;             REGISTRATION NUMBER: 41,971
;             REFERENCE/DOCKET NUMBER: PB186P3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-309-8504
;             TELEFAX: 301-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1830121 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Alignment Scores:
Pred. No.:              8.61e+05       Length:         1830121
Score:                  35.00          Matches:        6
Percent Similarity:     87.5%          Conservative:   1
Best Local Similarity:  75.0%          Mismatches:     1
Query Match:            77.8%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-557-884-1 (1-1830121)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||:::||||||   |||
Db    1553499 CCAGATGTCGTATTTCTCGATGAA 1553522

RESULT 35
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
;   GENERAL INFORMATION:
;        APPLICANT: Fleischmann et al.
;        TITLE OF INVENTION: The Nucleotide sequence of
;                            the Haemophilus influenzae Rd Genome, Fragments
;                            Thereof, and Uses Thereof
;        NUMBER OF SEQUENCES: 1
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3 1/2 inch diskette
;             COMPUTER: Dell Pentium
;             OPERATING SYSTEM: MS DOS v6.22
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/557,884
;             FILING DATE: 25-Apr-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/476,102
;             FILING DATE: JUN-5-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Michelle S. Marks
;             REGISTRATION NUMBER: 41,971
;             REFERENCE/DOCKET NUMBER: PB186P3
;        TELECOMMUNICATION INFORMATION:
```

```
;            TELEPHONE: 301-309-8504
;            TELEFAX: 301-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 1830121 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Alignment Scores:
Pred. No.:              8.61e+05       Length:         1830121
Score:                  35.00          Matches:        7
Percent Similarity:     88.9%          Conservative:   1
Best Local Similarity:  77.8%          Mismatches:     1
Query Match:            77.8%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-557-884-1 (1-1830121)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||   |||||||||:::
Db      43794 CCAGATGTGATTGGCTTACAGGAAATT 43768

RESULT 36
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
;   GENERAL INFORMATION:
;        APPLICANT: Robert D. Fleischmann
;                   Mark D. Adams
;                   Owen White
;                   Hamilton O. Smith
;                   J. Craig Venter
;        TITLE OF INVENTION: The Nucleotide sequence of
;                            the Haemophilus influenzae Rd Genome, Fragments
;                            Thereof, and Uses Thereof
;        NUMBER OF SEQUENCES: 1
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville,
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3 1/2 inch diskette
;             COMPUTER: Dell Pentium
;             OPERATING SYSTEM: MS DOS v6.22
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/643,990A
;             FILING DATE: 23-Aug-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/487,429
;             FILING DATE: 1995-06-07
;             APPLICATION NUMBER: 08/426,787
;             FILING DATE: 1995-04-21
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB186P1C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-610-5790
;             TELEFAX: 310-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1830121 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Alignment Scores:
Pred. No.:              8.61e+05       Length:         1830121
Score:                  35.00          Matches:        6
Percent Similarity:     87.5%          Conservative:   1
Best Local Similarity:  75.0%          Mismatches:     1
Query Match:            77.8%          Indels:         0
DB:                     3              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-643-990A-1 (1-1830121)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||:::||||||   |||
Db    1553499 CCAGATGTCGTATTTCTCGATGAA 1553522

RESULT 37
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
;   GENERAL INFORMATION:
;        APPLICANT: Robert D. Fleischmann
;                   Mark D. Adams
;                   Owen White
;                   Hamilton O. Smith
;                   J. Craig Venter
;        TITLE OF INVENTION: The Nucleotide sequence of
;                            the Haemophilus influenzae Rd Genome, Fragments
;                            Thereof, and Uses Thereof
;        NUMBER OF SEQUENCES: 1
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville,
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3 1/2 inch diskette
;             COMPUTER: Dell Pentium
;             OPERATING SYSTEM: MS DOS v6.22
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/643,990A
;             FILING DATE: 23-Aug-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/487,429
;             FILING DATE: 1995-06-07
;             APPLICATION NUMBER: 08/426,787
;             FILING DATE: 1995-04-21
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB186P1C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-610-5790
;             TELEFAX: 310-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1830121 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Alignment Scores:
Pred. No.:              8.61e+05       Length:         1830121
Score:                  35.00          Matches:        7
Percent Similarity:     88.9%          Conservative:   1
Best Local Similarity:  77.8%          Mismatches:     1
```

```
Query Match:            77.8%          Indels:        0
DB:                     3              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-643-990A-1 (1-1830121)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||   |||||||||:::
Db      43794 CCAGATGTGATTGGCTTACAGGAAATT 43768

RESULT 38
US-10-158-865-1
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
;  TITLE OF INVENTION:  Thereof, and Uses Thereof
;  FILE REFERENCE: PB186P2C1D1
;  CURRENT APPLICATION NUMBER: US/10/158,865
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US 09/557,884
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US 08/476,102
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/426,787
;  PRIOR FILING DATE: 1995-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1830121
;   TYPE: DNA
;   ORGANISM: Haemophilus influenzae
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4747)..(4747)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9921)..(9921)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10150)..(10150)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (29298)..(29298)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36543)..(36543)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36551)..(36551)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36636)..(36636)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (40808)..(40810)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44416)..(44416)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44905)..(44905)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (44975)..(44975)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45593)..(45593)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45732)..(45732)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (47036)..(47036)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51334)..(51334)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51602)..(51602)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51786)..(51786)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51805)..(51805)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (55369)..(55369)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65309)..(65309)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (65313)..(65313)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (80024)..(80024)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (100091)..(100091)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (102696)..(102696)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (105121)..(105121)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (107248)..(107248)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (117136)..(117136)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (119750)..(119750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
```

; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature

Alignment Scores:

Pred. No.:	8.61e+05	Length:	1830121
Score:	35.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	77.8%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-158-865-1 (1-1830121)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||:::||||||   |||
Db    1553499 CCAGATGTCGTATTTCTCGATGAA 1553522
```

RESULT 39
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fra
; Patent No. 6846651
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature

Alignment Scores:

Pred. No.:	8.61e+05	Length:	1830121
Score:	35.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	77.8%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-158-865-1 (1-1830121)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||   |||||||||:::
Db      43794 CCAGATGTGATTGGCTTACAGGAAATT 43768

RESULT 40
US-09-621-976-1908/c
; Sequence 1908, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 1908
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 65..316
;   NAME/KEY: sig_peptide
;   LOCATION: 65..256
;   OTHER INFORMATION: Von Heijne matrix
;   OTHER INFORMATION: score 4
;   OTHER INFORMATION: seq LTPVIPALXEAEA/GD
US-09-621-976-1908

Alignment Scores:
Pred. No.:              88.2          Length:         485
Score:                  34.00         Matches:        7
Percent Similarity:     87.5%         Conservative:   0
Best Local Similarity:  87.5%         Mismatches:     1
Query Match:            75.6%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-621-976-1908 (1-485)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||   |||||||||
Db        139 CCTGATGTAATACACCTTCAAGAG 116

RESULT 41
US-09-252-991A-5517
; Sequence 5517, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 5517
;   LENGTH: 582
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5517

Alignment Scores:
Pred. No.:              110           Length:         582
Score:                  34.00         Matches:        5
Percent Similarity:     88.9%         Conservative:   3
Best Local Similarity:  55.6%         Mismatches:     1
Query Match:            75.6%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-252-991A-5517 (1-582)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::   |||||||||:::
Db        374 CCGGACCTGGTCCTGCTCCAGGAGATC 400

RESULT 42
US-09-252-991A-8106
; Sequence 8106, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 8106
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8106

Alignment Scores:
Pred. No.:              151           Length:         756
Score:                  34.00         Matches:        6
Percent Similarity:     100.0%        Conservative:   2
Best Local Similarity:  75.0%         Mismatches:     0
Query Match:            75.6%         Indels:         0
DB:                     3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-252-991A-8106 (1-756)

Qy          2 AspValIlePheLeuGlnGluVal 9
              |||:::::|||||||||||||||
Db        133 GACCTGGTGTTCCTCCAGGAGGTC 156

RESULT 43
US-09-107-532A-2040/c
; Sequence 2040, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
```

```
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;    INFORMATION FOR SEQ ID NO: 2040:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 756 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: circular
;        MOLECULE TYPE: DNA (genomic)
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION: (B) LOCATION 1...756
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2040:
US-09-107-532A-2040

Alignment Scores:
Pred. No.:                151             Length:         756
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   3
Best Local Similarity:    62.5%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-107-532A-2040 (1-756)

Qy          2 AspValIlePheLeuGlnGluVal 9
              |||::::::|||||||||||||:::
Db        402 GACATTCTTTTCCTTCAAGAAATC 379

RESULT 44
US-09-540-236-1914
; Sequence 1914, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
;  TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.2005-001
;  CURRENT APPLICATION NUMBER: US/09/540,236
;  CURRENT FILING DATE: 2000-04-04
;  NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1914
;   LENGTH: 828
;   TYPE: DNA
;   ORGANISM: M.catarrhalis
US-09-540-236-1914

Alignment Scores:
Pred. No.:                168             Length:         828
Score:                    34.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
Best Local Similarity:    66.7%           Mismatches:     1
Query Match:              75.6%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-540-236-1914 (1-828)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::   |||||||||||
Db        184 CCTGATATTTTATGTCTGCAGGAGGTG 210

RESULT 45
US-09-902-540-5749
; Sequence 5749, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5749
;   LENGTH: 864
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-5749

Alignment Scores:
Pred. No.:                177             Length:         864
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    75.0%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-902-540-5749 (1-864)

Qy          2 AspValIlePheLeuGlnGluVal 9
              |||::::::|||||||||||||||||
Db        202 GACCTGGTGTTCCTGCAGGAGGTG 225

RESULT 46
US-09-252-991A-8199/c
; Sequence 8199, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 8199
;   LENGTH: 873
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8199

Alignment Scores:
Pred. No.:                179             Length:         873
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    75.0%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-252-991A-8199 (1-873)

Qy          2 AspValIlePheLeuGlnGluVal 9
              |||::::::|||||||||||||||||
Db        663 GACCTGGTGTTCCTCCAGGAGGTC 640

RESULT 47
US-09-902-540-3685
```

; Sequence 3685, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 3685
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-3685

Alignment Scores:

Pred. No.:	198	Length:	951
Score:	34.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	75.6%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-902-540-3685 (1-951)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||:::||||||   |||
Db        457 CCGGACGTGCTCTTCCTGGACGAG 480
```

RESULT 48
US-09-252-991A-8063
; Sequence 8063, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8063
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8063

Alignment Scores:

Pred. No.:	201	Length:	963
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-252-991A-8063 (1-963)

```
Qy          2 AspValIlePheLeuGlnGluVal 9
              |||::::::|||||||||||||||||
Db        926 GACCTGGTGTTCCTCCAGGAGGTC 949
```

RESULT 49
US-09-252-991A-5541
; Sequence 5541, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5541
; LENGTH: 1164
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5541

Alignment Scores:

Pred. No.:	253	Length:	1164
Score:	34.00	Matches:	5
Percent Similarity:	88.9%	Conservative:	3
Best Local Similarity:	55.6%	Mismatches:	1
Query Match:	75.6%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-252-991A-5541 (1-1164)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::   |||||||||:::
Db        487 CCGGACCTGGTCCTGCTCCAGGAGATC 513
```

RESULT 50
US-09-489-039A-82/c
; Sequence 82, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 82
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-82

Alignment Scores:

Pred. No.:	258	Length:	1182
Score:	34.00	Matches:	6
Percent Similarity:	88.9%	Conservative:	2
Best Local Similarity:	66.7%	Mismatches:	1
Query Match:	75.6%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-489-039A-82 (1-1182)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||   |||:::|||
Db         54 CCTGACGTCCTCTTTAAGCAGGACGTG 28
```

RESULT 51
US-09-252-991A-5682/c
; Sequence 5682, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5682
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5682

Alignment Scores:

Pred. No.:	279	Length:	1263
Score:	34.00	Matches:	5
Percent Similarity:	88.9%	Conservative:	3
Best Local Similarity:	55.6%	Mismatches:	1
Query Match:	75.6%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-252-991A-5682 (1-1263)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::   |||||||||:::
Db        960 CCGGACCTGGTCCTGCTCCAGGAGATC 934
```

RESULT 52
US-09-991-181-27/c
; Sequence 27, Application US/09991181
; Patent No. 6913919
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi,Avi J.
; APPLICANT: Baker,Kevin P.
; APPLICANT: Botstein,David
; APPLICANT: Desnoyers,Luc
; APPLICANT: Eaton,Dan L.
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Fong,Sherman
; APPLICANT: Gerber,Hanspeter
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,J.Christopher
; APPLICANT: Gurney,Austin L.
; APPLICANT: Kljavin,Ivar J.
; APPLICANT: Napier,Mary A.
; APPLICANT: Pan,James
; APPLICANT: Paoni,Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C53
; CURRENT APPLICATION NUMBER: US/09/991,181
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	316	Length:	1399
Score:	34.00	Matches:	7
Percent Similarity:	87.5%	Conservative:	0
Best Local Similarity:	87.5%	Mismatches:	1
Query Match:	75.6%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-991-181-27 (1-1399)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||   |||||||||
Db       1053 CCTGATGTAATACACCTTCAAGAG 1030
```

RESULT 53
US-09-990-444-27/c
; Sequence 27, Application US/09990444
; Patent No. 6930170
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi,Avi J.
; APPLICANT: Baker,Kevin P.
; APPLICANT: Botstein,David
; APPLICANT: Desnoyers,Luc
; APPLICANT: Eaton,Dan L.
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Fong,Sherman
; APPLICANT: Gerber,Hanspeter
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,J.Christopher
; APPLICANT: Gurney,Austin L.
; APPLICANT: Kljavin,Ivar J.
; APPLICANT: Napier,Mary A.
; APPLICANT: Pan,James
; APPLICANT: Paoni,Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C19
; CURRENT APPLICATION NUMBER: US/09/990,444
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519

; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	316	Length:	1399
Score:	34.00	Matches:	7
Percent Similarity:	87.5%	Conservative:	0
Best Local Similarity:	87.5%	Mismatches:	1
Query Match:	75.6%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-990-444-27 (1-1399)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||   |||||||||
Db       1053 CCTGATGTAATACACCTTCAAGAG 1030
```

RESULT 54
US-09-997-333-27/c
; Sequence 27, Application US/09997333
; Patent No. 6953836
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi,Avi J.
; APPLICANT: Baker,Kevin P.
; APPLICANT: Botstein,David
; APPLICANT: Desnoyers,Luc
; APPLICANT: Eaton,Dan L.
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Fong,Sherman
; APPLICANT: Gerber,Hanspeter
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,J.Christopher
; APPLICANT: Gurney,Austin L.
; APPLICANT: Kljavin,Ivar J.
; APPLICANT: Napier,Mary A.
; APPLICANT: Pan,James
; APPLICANT: Paoni,Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C27
; CURRENT APPLICATION NUMBER: US/09/997,333
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	316	Length:	1399
Score:	34.00	Matches:	7
Percent Similarity:	87.5%	Conservative:	0
Best Local Similarity:	87.5%	Mismatches:	1
Query Match:	75.6%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-997-333-27 (1-1399)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||   |||||||||
Db       1053 CCTGATGTAATACACCTTCAAGAG 1030
```

RESULT 55
US-09-992-598-27/c
; Sequence 27, Application US/09992598
; Patent No. 6956108
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi,Avi J.
; APPLICANT: Baker,Kevin P.
; APPLICANT: Botstein,David
; APPLICANT: Desnoyers,Luc
; APPLICANT: Eaton,Dan L.
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Fong,Sherman
; APPLICANT: Gerber,Hanspeter
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,J.Christopher
; APPLICANT: Gurney,Austin L.
; APPLICANT: Kljavin,Ivar J.
; APPLICANT: Napier,Mary A.
; APPLICANT: Pan,James
; APPLICANT: Paoni,Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C20
; CURRENT APPLICATION NUMBER: US/09/992,598
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633

```
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091978
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/091982
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/092182
;  PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.:                316             Length:         1399
Score:                    34.00           Matches:        7
Percent Similarity:       87.5%           Conservative:   0
Best Local Similarity:    87.5%           Mismatches:     1
Query Match:              75.6%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-992-598-27 (1-1399)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||   |||||||||
Db       1053 CCTGATGTAATACACCTTCAAGAG 1030

RESULT 56
US-09-632-538C-16/c
; Sequence 16, Application US/09632538C
; Patent No. 6440674
; GENERAL INFORMATION:
;  APPLICANT: Misra, Santosh et al.
;  TITLE OF INVENTION: PLANT PROMOTER DERIVED FROM LUMINAL BINDING PROTEIN GENE AND METH
;  TITLE OF INVENTION:  ITS USE
;  FILE REFERENCE: 54359
;  CURRENT APPLICATION NUMBER: US/09/632,538C
;  CURRENT FILING DATE:  2000-08-04
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
;   LENGTH: 2263
;   TYPE: DNA
;   ORGANISM: Pseudotsuga menziesii
US-09-632-538C-16

Alignment Scores:
Pred. No.:                563             Length:         2263
Score:                    34.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-632-538C-16 (1-2263)

Qy          2 AspValIlePheLeuGlnGlu 8
              |||||||||||||||||||
Db        832 GATGTAATTTTCTTGCAAGAA 812

RESULT 57
US-09-632-538C-31/c
; Sequence 31, Application US/09632538C
; Patent No. 6440674
; GENERAL INFORMATION:
;  APPLICANT: Misra, Santosh et al.
;  TITLE OF INVENTION: PLANT PROMOTER DERIVED FROM LUMINAL BINDING PROTEIN GENE AND METH
;  TITLE OF INVENTION:  ITS USE
;  FILE REFERENCE: 54359
;  CURRENT APPLICATION NUMBER: US/09/632,538C
;  CURRENT FILING DATE:  2000-08-04
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
;   LENGTH: 2277
;   TYPE: DNA
;   ORGANISM: Pseudotsuga menziesii
US-09-632-538C-31

Alignment Scores:
Pred. No.:                567             Length:         2277
Score:                    34.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-632-538C-31 (1-2277)

Qy          2 AspValIlePheLeuGlnGlu 8
              |||||||||||||||||||
Db        838 GATGTAATTTTCTTGCAAGAA 818

RESULT 58
US-08-981-690-1/c
; Sequence 1, Application US/08981690
; Patent No. 5976844
; GENERAL INFORMATION:
;    APPLICANT:  KAESLER, Bruno; SAHM, Hermann; STAHMANN, Klaus-
;    APPLICANT:  Peter; SCHMIDT, Georg; BOEDDECKER, Theo;
;    APPLICANT:  SEULBERGER, Harald
;    TITLE OF INVENTION:  PREPARATION OF RIBOFLAVIN USING
;    TITLE OF INVENTION:  MICROORGANISMS HAVING AN ALTERED
;    TITLE OF INVENTION:  ISOCITRATE LYASE ACTIVITY
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Keil & Weinkauf
;      STREET:  1101 Connecticut Avenue
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.2 Mb storage
;      COMPUTER:  IBM AT-compatible, 80486 processor
;      OPERATING SYSTEM:  MS-DOS version 6.1
;      SOFTWARE:  WordPerfect version 8.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/981,690
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/590,454
;      FILING DATE:  23-JAN-1996
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2364 base pairs
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..550
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  551..2233
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  2234..2364
US-08-981-690-1

Alignment Scores:
Pred. No.:                593             Length:         2364
Score:                    34.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
Best Local Similarity:    66.7%           Mismatches:     1
Query Match:              75.6%           Indels:         0
DB:                       2               Gaps:           0
```

```
US-10-757-745-2_COPY_145_153 (1-9) x US-08-981-690-1 (1-2364)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||::::::   ||||||
Db       2301 CCCGATGTTATATATATTCCAGAAGTC 2275

RESULT 59
US-09-902-540-1178/c
; Sequence 1178, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1178
;   LENGTH: 16924
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(16924)
;   OTHER INFORMATION: unsure at all n locations
US-09-902-540-1178

Alignment Scores:
Pred. No.:              6.34e+03        Length:         16924
Score:                  34.00           Matches:        6
Percent Similarity:     87.5%           Conservative:   1
Best Local Similarity:  75.0%           Mismatches:     1
Query Match:            75.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-902-540-1178 (1-16924)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||:::||||||   |||
Db      10280 CCGGACGTGCTCTTCCTGGACGAG 10257

RESULT 60
US-08-311-731A-123
; Sequence 123, Application US/08311731A
; Patent No. 6583266
;  GENERAL INFORMATION:
;    APPLICANT:  SMITH, DOUGLAS
;    APPLICANT:  MAO, JEN-I
;    TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES
;    TITLE OF INVENTION:  RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
;    TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;    NUMBER OF SEQUENCES:  411
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WOLF, GREENFIELD & SACKS, P.C.
;      STREET:  600 ATLANTIC AVENUE
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,731A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GATES, EDWARD R.
;      REGISTRATION NUMBER:  31,616
;      REFERENCE/DOCKET NUMBER:  C0044/7125
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/720-3500
;      TELEFAX:  617/720-2441
;  INFORMATION FOR SEQ ID NO:  123:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36470 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  MYCOBACTERIUM LEPRAE
US-08-311-731A-123

Alignment Scores:
Pred. No.:              1.6e+04         Length:         36470
Score:                  34.00           Matches:        6
Percent Similarity:     87.5%           Conservative:   1
Best Local Similarity:  75.0%           Mismatches:     1
Query Match:            75.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-311-731A-123 (1-36470)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||:::||||||   |||
Db      19155 CCCGACGTGCTCTTCTTATCCGAA 19178

RESULT 61
US-08-311-731A-122
; Sequence 122, Application US/08311731A
; Patent No. 6583266
;  GENERAL INFORMATION:
;    APPLICANT:  SMITH, DOUGLAS
;    APPLICANT:  MAO, JEN-I
;    TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES
;    TITLE OF INVENTION:  RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
;    TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;    NUMBER OF SEQUENCES:  411
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WOLF, GREENFIELD & SACKS, P.C.
;      STREET:  600 ATLANTIC AVENUE
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,731A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GATES, EDWARD R.
;      REGISTRATION NUMBER:  31,616
;      REFERENCE/DOCKET NUMBER:  C0044/7125
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/720-3500
;      TELEFAX:  617/720-2441
;  INFORMATION FOR SEQ ID NO:  122:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  41171 base pairs
;      TYPE:  nucleic acid
```

```
;       STRANDEDNESS:  double
;       TOPOLOGY:  circular
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  MYCOBACTERIUM LEPRAE
US-08-311-731A-122

Alignment Scores:
Pred. No.:              1.85e+04        Length:         41171
Score:                  34.00           Matches:        6
Percent Similarity:     87.5%           Conservative:   1
Best Local Similarity:  75.0%           Mismatches:     1
Query Match:            75.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-08-311-731A-122 (1-41171)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||:::||||||   |||
Db       2661 CCGGACGTACTGTTTCTCGACGAG 2684

RESULT 62
US-09-902-540-1272/c
; Sequence 1272, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1272
;   LENGTH: 54484
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-1272

Alignment Scores:
Pred. No.:              2.59e+04        Length:         54484
Score:                  34.00           Matches:        6
Percent Similarity:     100.0%          Conservative:   2
Best Local Similarity:  75.0%           Mismatches:     0
Query Match:            75.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-902-540-1272 (1-54484)

Qy          2 AspValIlePheLeuGlnGluVal 9
              |||::::::|||||||||||||||
Db      12287 GACCTGGTGTTCCTGCAGGAGGTG 12264

RESULT 63
US-09-949-016-17557/c
; Sequence 17557, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17557
;   LENGTH: 116966
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17557

Alignment Scores:
Pred. No.:              6.4e+04         Length:         116966
Score:                  34.00           Matches:        6
Percent Similarity:     88.9%           Conservative:   2
Best Local Similarity:  66.7%           Mismatches:     1
Query Match:            75.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-016-17557 (1-116966)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||   ::::::|||||||||||||
Db      94482 CCTGACTTTGTTTGGCTACAGGAGGTA 94456

RESULT 64
US-09-949-016-17127
; Sequence 17127, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17127
;   LENGTH: 228896
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(228896)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17127

Alignment Scores:
Pred. No.:              1.41e+05        Length:         228896
Score:                  34.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   4
Best Local Similarity:  55.6%           Mismatches:     0
Query Match:            75.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-016-17127 (1-228896)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||::::::||||||:::|||:::
Db      33561 CCAGATCTTGTGTTTCTCAGAGAGCTC 33587

RESULT 65
US-09-596-002-41
; Sequence 41, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
```

```
;  APPLICANT:      Lagace, Robert, E.
;  APPLICANT:  Patterson, Chandra
;  APPLICANT:  Berg, Kim, L.
;  TITLE OF INVENTION:  NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
;  FILE REFERENCE:  PM-0008-4 US
;  CURRENT APPLICATION NUMBER: US/09/596,002
;  CURRENT FILING DATE:  2000-06-16
;  PRIOR APPLICATION NUMBER:  60/140,121
;  PRIOR FILING DATE:  1999-06-18
;  NUMBER OF SEQ ID NOS:      41
;  SOFTWARE:       PERL Program
; SEQ ID NO 41
;   LENGTH: 269223
;   TYPE: DNA
;   ORGANISM: Moraxella catarrhalis
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte template ID No. 6632636 41
;  PUBLICATION INFORMATION:
US-09-596-002-41

Alignment Scores:
Pred. No.:              1.71e+05        Length:         269223
Score:                  34.00           Matches:        6
Percent Similarity:     88.9%           Conservative:   2
Best Local Similarity:  66.7%           Mismatches:     1
Query Match:            75.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-596-002-41 (1-269223)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::   ||||||||||||
Db     221174 CCTGATATTTTATGTCTGCAGGAGGTG 221200

RESULT 66
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.:              2.23e+06        Length:         4403765
Score:                  34.00           Matches:        6
Percent Similarity:     87.5%           Conservative:   1
Best Local Similarity:  75.0%           Mismatches:     1
Query Match:            75.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-103-840A-2 (1-4403765)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||:::||||||   |||
Db    1494304 CCCGACGTGCTGTTCCTGTCCGAG 1494281

RESULT 67
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.:              2.23e+06        Length:         4411529
Score:                  34.00           Matches:        6
Percent Similarity:     87.5%           Conservative:   1
Best Local Similarity:  75.0%           Mismatches:     1
Query Match:            75.6%           Indels:         0
DB:                     3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-103-840A-1 (1-4411529)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||:::||||||   |||
Db    1493106 CCCGACGTGCTGTTCCTGTCCGAG 1493083

RESULT 68
US-09-513-999C-28064/c
; Sequence 28064, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 28064
;   LENGTH: 247
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 202
;   OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-28064

Alignment Scores:
Pred. No.:              66.4            Length:         247
Score:                  33.00           Matches:        5
Percent Similarity:     100.0%          Conservative:   3
Best Local Similarity:  62.5%           Mismatches:     0
Query Match:            73.3%           Indels:         0
DB:                     3               Gaps:           0
```

US-10-757-745-2_COPY_145_153 (1-9) x US-09-513-999C-28064 (1-247)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||:::::::::|||||||||||||
Db         59 CCTGAACTTGTCTTTCTTCAGGAG 36
```

RESULT 69
US-09-252-991A-5110/c
; Sequence 5110, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5110
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5110

Alignment Scores:

Pred. No.:	158	Length:	507
Score:	33.00	Matches:	5
Percent Similarity:	87.5%	Conservative:	2
Best Local Similarity:	62.5%	Mismatches:	1
Query Match:	73.3%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-252-991A-5110 (1-507)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||::::::||||||   |||
Db        417 CCGGACATCCTGTTCCTCGACGAG 394
```

RESULT 70
US-09-949-016-182845/c
; Sequence 182845, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 182845
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-182845

Alignment Scores:

Pred. No.:	194	Length:	601
Score:	33.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	73.3%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-016-182845 (1-601)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||:::|||   ||||||
Db        438 CCAGATGTCTTGTTTTCACAAGAG 415
```

RESULT 71
US-09-949-002-3663/c
; Sequence 3663, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3663
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-3663

Alignment Scores:

Pred. No.:	194	Length:	601
Score:	33.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	3
Best Local Similarity:	62.5%	Mismatches:	0
Query Match:	73.3%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-002-3663 (1-601)

```
Qy          2 AspValIlePheLeuGlnGluVal 9
              |||:::|||||||||||||::::::
Db        140 GATATTATATTCTTACAACAAATT 117
```

RESULT 72
US-09-949-002-3664/c
; Sequence 3664, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3664
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-3664

Alignment Scores:

Pred. No.:	194	Length:	601
Score:	33.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	3
Best Local Similarity:	62.5%	Mismatches:	0
Query Match:	73.3%	Indels:	0

```
DB:                    3             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-002-3664 (1-601)

Qy          2 AspValIlePheLeuGlnGluVal 9
              |||:::|||||||||||||:::::::
Db         85 GATATTATATTCTTACAACAAATT 62

RESULT 73
US-09-949-002-3665/c
; Sequence 3665, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3665
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-3665

Alignment Scores:
Pred. No.:              194           Length:        601
Score:                  33.00         Matches:       5
Percent Similarity:     100.0%        Conservative:  3
Best Local Similarity:  62.5%         Mismatches:    0
Query Match:            73.3%         Indels:        0
DB:                     3             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-002-3665 (1-601)

Qy          2 AspValIlePheLeuGlnGluVal 9
              |||:::|||||||||||||:::::::
Db         61 GATATTATATTCTTACAACAAATT 38

RESULT 74
US-09-949-002-5508/c
; Sequence 5508, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5508
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-5508

Alignment Scores:
Pred. No.:              194           Length:        601
Score:                  33.00         Matches:       5
Percent Similarity:     100.0%        Conservative:  3
Best Local Similarity:  62.5%         Mismatches:    0
Query Match:            73.3%         Indels:        0
DB:                     3             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-002-5508 (1-601)

Qy          2 AspValIlePheLeuGlnGluVal 9
              |||:::|||||||||||||:::::::
Db        140 GATATTATATTCTTACAACAAATT 117

RESULT 75
US-09-949-002-5509/c
; Sequence 5509, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5509
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-5509

Alignment Scores:
Pred. No.:              194           Length:        601
Score:                  33.00         Matches:       5
Percent Similarity:     100.0%        Conservative:  3
Best Local Similarity:  62.5%         Mismatches:    0
Query Match:            73.3%         Indels:        0
DB:                     3             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-949-002-5509 (1-601)

Qy          2 AspValIlePheLeuGlnGluVal 9
              |||:::|||||||||||||:::::::
Db         85 GATATTATATTCTTACAACAAATT 62


Search completed: March 10, 2006, 21:44:42
Job time : 266.235 secs
```

THIS PAGE BLANK (USPTO)

Date not found

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 14:38:30 ; Search time 34.8261 Seconds
(without alignments)
1722.336 Million cell updates/sec

Title: US-10-757-745-2_COPY_145_153
Perfect score: 45
Sequence: 1 PDVIFLQEV 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170325_3993/app_query.fasta_1
-DB=N_Geneseq -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=75
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1589_@runat_09032006_170325_3993 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	45	100.0	176	6	ABK27684	Abk27684 Human col
c	2	45	100.0	176	6	ABV95959	Abv95959 Human pan
c	3	45	100.0	179	6	ABL36461	Abl36461 Human col
	4	45	100.0	401	11	ADT95844	Adt95844 Colon can
	5	45	100.0	401	11	ADX42326	Adx42326 Human cDN
	6	45	100.0	444	4	AAI15311	Aai15311 Probe #52
	7	45	100.0	483	2	AAX40590	Aax40590 Human sec
	8	45	100.0	553	11	ADT95464	Adt95464 Colon can
	9	45	100.0	553	11	ADX41946	Adx41946 Human cDN
	10	45	100.0	625	11	ADT95512	Adt95512 Colon can
	11	45	100.0	625	11	ADX41994	Adx41994 Human cDN
	12	45	100.0	633	11	ADT95565	Adt95565 Colon can
	13	45	100.0	633	11	ADX42047	Adx42047 Human cDN
	14	45	100.0	644	11	ADT95551	Adt95551 Colon can
	15	45	100.0	644	11	ADX42033	Adx42033 Human cDN
	16	45	100.0	674	11	ADT95548	Adt95548 Colon can
	17	45	100.0	674	11	ADX42030	Adx42030 Human cDN
	18	45	100.0	691	4	AAI24484	Aai24484 Probe #14
	19	45	100.0	752	4	AAH08073	Aah08073 Human cDN
c	20	45	100.0	1079	2	AAX84209	Aax84209 DNA encod
c	21	45	100.0	1079	3	AAC79438	Aac79438 cDNA sequ
c	22	45	100.0	1079	6	ABK28982	Abk28982 Human bre
	23	45	100.0	1088	5	AAS86254	Aas86254 DNA encod
	24	45	100.0	1296	3	AAC98160	Aac98160 Human col
	25	45	100.0	1898	4	AAH15146	Aah15146 Human cDN
	26	45	100.0	1920	3	AAZ47118	Aaz47118 Human CD4
	27	45	100.0	1936	10	ADD19013	Add19013 Human dis
	28	45	100.0	1936	13	ADP25361	Adp25361 PRO polyp
	29	45	100.0	1940	14	ADX06356	Adx06356 Cyclin-de
	30	45	100.0	1948	4	AAI58997	Aai58997 Human pol
	31	45	100.0	1948	5	ADQ99219	Adq99219 DNA encod
	32	45	100.0	1948	9	ADB48979	Adb48979 Novel hum
	33	45	100.0	1958	4	AAI60783	Aai60783 Human pol
	34	45	100.0	2499	2	AAX28153	Aax28153 Topoisome
	35	45	100.0	3152	9	ADA10970	Ada10970 Human cDN
	36	44	97.8	602	10	ADD34376	Add34376 Mouse mit
	37	44	97.8	1312	3	AAZ47119	Aaz47119 Mouse CD4
	38	40	88.9	625	11	ADT95029	Adt95029 Colon can
	39	40	88.9	625	11	ADX41511	Adx41511 Human cDN
	40	39	86.7	110000	14	AEB39175_05	Continuation (6 of
	41	39	86.7	110000	14	AEB42401_05	Continuation (6 of
	42	39	86.7	243335	14	AEB42735	Aeb42735 L. pneumo
	43	39	86.7	295644	14	AEB35721	Aeb35721 L. pneumo
	44	39	86.7	298667	14	AEB39173	Aeb39173 L. pneumo
c	45	38	84.4	244	6	ABN16687	Abn16687 Human ORF
	46	38	84.4	912	5	AAH65711	Aah65711 C glutami
	47	38	84.4	921	8	ACA00116	Aca00116 C. glutam
	48	38	84.4	1838	13	ADX50393	Adx50393 Plant ful
	49	38	84.4	349980	5	AAH68526	Aah68526 C glutami
c	50	37	82.2	516	10	ADE81339	Ade81339 Arabidops
c	51	37	82.2	753	12	ADN72674	Adn72674 Thale cre
c	52	37	82.2	766	3	AAC44301	Aac44301 Arabidops
c	53	37	82.2	768	3	AAZ98318	Aaz98318 A. thalia
c	54	37	82.2	768	13	ADU92189	Adu92189 Thale cre
	55	37	82.2	774	13	ADT47013	Adt47013 Bacterial
	56	37	82.2	777	13	ADS56055	Ads56055 Bacterial
c	57	37	82.2	833	3	AAC40379	Aac40379 Arabidops
	58	37	82.2	852	13	ADT48666	Adt48666 Bacterial
c	59	37	82.2	936	3	AAC40378	Aac40378 Arabidops
	60	37	82.2	1350	4	ABL08315	Abl08315 Drosophil
	61	37	82.2	39949	12	ADF31997	Adf31997 Full leng
	62	37	82.2	48200	12	ADF31998	Adf31998 Cosmid 2A
	63	37	82.2	208765	12	ADQ97430	Adq97430 Mouse can
	64	36	80.0	201	13	ADQ45678	Adq45678 Myocardia
	65	36	80.0	201	13	ADQ45683	Adq45683 Myocardia
	66	36	80.0	201	13	ADQ45674	Adq45674 Myocardia
	67	36	80.0	201	13	ADQ45687	Adq45687 Myocardia
	68	36	80.0	245	12	ADQ06782	Adq06782 Soybean t
	69	36	80.0	254	12	ADQ06781	Adq06781 Soybean t
	70	36	80.0	346	4	AAI87230	Aai87230 Human pol
c	71	36	80.0	396	4	AAI87801	Aai87801 Human pol
	72	36	80.0	717	8	ACA48082	Aca48082 Prokaryot
c	73	36	80.0	768	11	ACH96047	Ach96047 Klebsiell
	74	36	80.0	864	14	ACL68818	Acl68818 M. xanthu
	75	36	80.0	1040	14	ADV92147	Adv92147 Mouse CX3

ALIGNMENTS

```
RESULT 1
ABK27684
ID   ABK27684 standard; cDNA; 176 BP.
XX
AC   ABK27684;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human colon cancer expressed sequence tag, Seq ID no 121.
XX
KW   Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
KW   expressed sequence tag.
XX
OS   Homo sapiens.
XX
PN   WO200196390-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US018577.
XX
PR   09-JUN-2000; 2000US-0210821P.
PR   18-DEC-2000; 2000US-0256571P.
PR   10-MAY-2001; 2001US-0290240P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Hepler WT,  Clapper JD,  Wang A,  Secrist H;
XX
DR   WPI; 2002-139708/18.
XX
PT   Novel isolated polynucleotide encoding a polypeptide comprising a portion
PT   of colon tumor protein, useful for detection, diagnosis and therapy of
PT   human colon cancer.
XX
PS   Claim 1; Page 182; 220pp; English.
XX
CC   The invention relates to an isolated polynucleotide (I) encoding a
CC   polypeptide (II) comprising at least a portion of a colon tumour protein.
CC   (I), (II) and antibody (III) to (II) are useful for determining the
CC   presence of a cancer in a patient. (I), (II) or antigen presenting cells
CC   expressing (I) is useful for stimulating and/or expanding T cells
CC   specific for a tumour protein, by contacting T cells with (I), (II) or
CC   antigen-presenting cells that express (I), under conditions and for a
CC   time sufficient to permit the stimulation and/or expansion of T cells.
CC   (I), (II), or antigen presenting cells that express (II) are useful for
CC   treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
CC   isolated from a patient with (I), (II) or antigen presenting cells that
CC   express (II), such that T cells proliferate, and administering to the
CC   patient an effective amount of the proliferated T cells, thus inhibiting
CC   the development of a cancer in the patient. (I) or (II) is useful in
CC   vaccines and pharmaceutical compositions for prevention and treatment of
CC   colon malignancies and for the diagnosis and monitoring of such cancers.
CC   (I), (II) or (III) is useful for detection, diagnosis and/or therapy of
CC   human colon cancer. (I) is useful as a probe or primer for nucleic acid
CC   hybridisation, and in the design and preparation of ribozyme molecules
CC   for inhibiting expression of (II) in tumour cells. ABK27564-ABK27807
CC   represent novel human colon cancer coding sequences and primers of the
CC   invention
XX
SQ   Sequence 176 BP; 68 A; 27 C; 31 G; 50 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0.398          Length:         176
Score:                  45.00          Matches:        9
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
Query Match:            100.0%         Indels:         0
DB:                     6              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x ABK27684 (1-176)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          6 CCAGATGTGATATTTCTACAGGAAGTT 32

RESULT 2
ABV95959/c
ID   ABV95959 standard; cDNA; 176 BP.
XX
AC   ABV95959;
XX
DT   14-JAN-2003  (first entry)
XX
DE   Human pancreatic cancer expressed cDNA SEQ ID NO 1367.
XX
KW   Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW   cytostatic; tumour; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200260317-A2.
XX
PD   08-AUG-2002.
XX
PF   30-JAN-2002; 2002WO-US002781.
XX
PR   30-JAN-2001; 2001US-0265305P.
PR   31-JAN-2001; 2001US-0265682P.
PR   09-FEB-2001; 2001US-0267568P.
PR   21-MAR-2001; 2001US-0278651P.
PR   28-APR-2001; 2001US-0287112P.
PR   16-MAY-2001; 2001US-0291631P.
PR   12-JUL-2001; 2001US-0305484P.
PR   20-AUG-2001; 2001US-0313999P.
PR   27-NOV-2001; 2001US-0333626P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Benson DR,  Kalos MD,  Lodes MJ,  Persing DH,  Hepler WT,  Jiang Y;
XX
DR   WPI; 2002-627435/67.
XX
PT   New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT   diagnosing, preventing and/or treating cancer, particularly pancreatic
PT   cancer.
XX
PS   Claim 1; SEQ ID NO 1367; 300pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (I) comprising: (a)
CC   any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
CC   complements of (a); (c) sequences consisting of at least 20 contiguous
CC   residues of (a); (d) sequences that hybridize to (a), under moderately
CC   stringent conditions; (e) sequences having at least 75% or 90% identity
CC   to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
CC   ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer
CC   in a patient and compositions comprising polypeptides, polynucleotides,
CC   antibodies, fusion proteins, T cell populations and antigen presenting
CC   cells expressing the polypeptide are useful in treating pancreatic cancer
CC   and stimulating an immune response. The polynucleotides can be used as
CC   probes or primers for nucleic acid hybridisation, in the design and
CC   preparation of ribozyme molecules for inhibiting expression of the tumour
CC   polypeptides and proteins in the tumour cells, in vaccines and for gene
CC   therapy. Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 176 BP; 50 A; 31 C; 27 G; 68 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0.398          Length:         176
Score:                  45.00          Matches:        9
Percent Similarity:     100.0%         Conservative:   0
Best Local Similarity:  100.0%         Mismatches:     0
```

```
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x ABV95959 (1-176)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145

RESULT 3
ABL36461/c
ID   ABL36461 standard; cDNA; 179 BP.
XX
AC   ABL36461;
XX
DT   08-APR-2002  (first entry)
XX
DE   Human colon tumour antigen polynucleotide SEQ ID NO:50.
XX
KW   Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW   colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200196388-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US018557.
XX
PR   09-JUN-2000; 2000US-0210899P.
PR   20-FEB-2001; 2001US-0270216P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Harlocker SL,  Secrist H;
XX
DR   WPI; 2002-114514/15.
XX
PT   Novel isolated colon tumor polynucleotide differentially expressed in
PT   colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT   useful for inhibiting development of cancer in patient.
XX
PS   Claim 1; SEQ ID NO 50; 105pp; English.
XX
CC   ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC   which were isolated from human colon tumour and colon metastatic tumour
CC   cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC   production. (I) can be used for stimulating and/or expanding T cells
CC   specific for a tumour protein on contact with the T cells. They are also
CC   useful for inhibiting the development of cancer in a patient. (I) can be
CC   used as probes or primers for nucleic acid hybridisation, for preparing
CC   mutant species primers, or primers for use in genetic constructions. (I)
CC   can be used in the diagnosis of a colon tumour
XX
SQ   Sequence 179 BP; 51 A; 31 C; 27 G; 68 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              0.406           Length:         179
Score:                  45.00           Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x ABL36461 (1-179)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145

RESULT 4
ADT95844
ID   ADT95844 standard; cDNA; 401 BP.
XX
AC   ADT95844;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1363.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1363; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 401 BP; 132 A; 67 C; 89 G; 113 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.04            Length:         401
Score:                  45.00           Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     11              Gaps:           0
```

US-10-757-745-2_COPY_145_153 (1-9) x ADT95844 (1-401)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db         55 CCAGATGTGATATTTCTACAGGAAGTT 81
```

RESULT 5
ADX42326
ID ADX42326 standard; cDNA; 401 BP.
XX
AC ADX42326;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 1363.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Panger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1363; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 401 BP; 132 A; 67 C; 89 G; 113 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.04	Length:	401
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	11	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADX42326 (1-401)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db         55 CCAGATGTGATATTTCTACAGGAAGTT 81
```

RESULT 6
AAI15311
ID AAI15311 standard; DNA; 444 BP.
XX
AC AAI15311;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #5244 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 5244; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 444 BP; 138 A; 60 C; 80 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.17	Length:	444
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AAI15311 (1-444)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||
Db        260 CCAGATGTGATATTTCTACAGGAAGTT 286
```

RESULT 7
AAX40590
ID AAX40590 standard; cDNA; 483 BP.
XX
AC AAX40590;

```
XX
DT   18-JUN-1999  (first entry)
XX
DE   Human secreted protein 5' EST SEQ ID No: 190.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS   Homo sapiens.
XX
PN   WO9906550-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB001232.
XX
PR   01-AUG-1997;   97US-00905144.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Lacroix B;
XX
DR   WPI; 1999-153780/13.
DR   P-PSDB; AAY11868.
XX
PT   New isolated prostate-derived nucleic acids - used to develop products
PT   which may have cytokine, immune regulatory, haematopoiesis regulating,
PT   anti-inflammatory or tumour inhibition activity.
XX
PS   Claim 1; Page 298; 675pp; English.
XX
CC   AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
CC   human secreted proteins expressed in prostate, and encode the proteins
CC   given in AAY11716 to AAY11993 respectively. The proteins given represent
CC   the signal peptide and an N-terminal fragment of a secreted protein. The
CC   nucleic acid sequences can be used for producing secreted human gene
CC   products. They can also be used to develop products for diagnosis and
CC   therapy. The proteins obtained may have cytokine activity, cell
CC   proliferation and differentiation activity, haematopoiesis regulating
CC   activity, tissue growth regulating activity, reproductive hormone
CC   regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC   thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, tumour inhibition activity or other activities. The products
CC   can be used in forensic, gene therapy and chromosome mapping procedures.
CC   The sequences can also be used for obtaining corresponding promoter
CC   sequences. The nucleic acids encoding the signal peptides can be used for
CC   directing extracellular secretion of a polypeptide or the insertion of a
CC   polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ   Sequence 483 BP; 123 A; 111 C; 139 G; 110 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.29          Length:          483
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       2             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AAX40590 (1-483)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459

RESULT 8
ADT95464
ID   ADT95464 standard; cDNA; 553 BP.
XX
AC   ADT95464;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #983.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 983; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:                1.51          Length:          553
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       11            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ADT95464 (1-553)
```

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         78 CCAGATGTGATATTTCTACAGGAAGTT 104

RESULT 9
ADX41946
ID   ADX41946 standard; cDNA; 553 BP.
XX
AC   ADX41946;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 983.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 983; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              1.51          Length:          553
Score:                  45.00         Matches:         9
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
DB:                     11            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ADX41946 (1-553)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         78 CCAGATGTGATATTTCTACAGGAAGTT 104

RESULT 10
ADT95512
ID   ADT95512 standard; cDNA; 625 BP.
XX
AC   ADT95512;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1031.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1031; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              1.74          Length:          625
Score:                  45.00         Matches:         9
Percent Similarity:     100.0%        Conservative:    0
Best Local Similarity:  100.0%        Mismatches:      0
Query Match:            100.0%        Indels:          0
```

```
DB:                    11            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ADT95512 (1-625)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         78 CCAGATGTGATATTTCTACAGGAAGTT 104

RESULT 11
ADX41994
ID   ADX41994 standard; cDNA; 625 BP.
XX
AC   ADX41994;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1031.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1031; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              1.74          Length:        625
Score:                  45.00         Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            100.0%        Indels:        0
DB:                     11            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ADX41994 (1-625)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         78 CCAGATGTGATATTTCTACAGGAAGTT 104

RESULT 12
ADT95565
ID   ADT95565 standard; cDNA; 633 BP.
XX
AC   ADT95565;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1084.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1084; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              1.77          Length:        633
Score:                  45.00         Matches:       9
```

```
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            100.0%          Indels:        0
DB:                     11              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ADT95565 (1-633)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         70 CCAGATGTGATATTTCTACAGGAAGTT 96

RESULT 13
ADX42047
ID   ADX42047 standard; cDNA; 633 BP.
XX
AC   ADX42047;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1084.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1084; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              1.77            Length:        633
Score:                  45.00           Matches:       9
Percent Similarity:     100.0%          Conservative:  0
Best Local Similarity:  100.0%          Mismatches:    0
Query Match:            100.0%          Indels:        0
DB:                     11              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ADX42047 (1-633)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         70 CCAGATGTGATATTTCTACAGGAAGTT 96

RESULT 14
ADT95551
ID   ADT95551 standard; cDNA; 644 BP.
XX
AC   ADT95551;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1070.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1070; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:                1.81          Length:          644
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       11            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ADT95551 (1-644)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         68 CCAGATGTGATATTTCTACAGGAAGTT 94

RESULT 15
ADX42033
ID   ADX42033 standard; cDNA; 644 BP.
XX
AC   ADX42033;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1070.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1070; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.81          Length:          644
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       11            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ADX42033 (1-644)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         68 CCAGATGTGATATTTCTACAGGAAGTT 94

RESULT 16
ADT95548
ID   ADT95548 standard; cDNA; 674 BP.
XX
AC   ADT95548;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1067.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1067; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
```

```
XX
SQ   Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.:              1.9             Length:         674
Score:                  45.00           Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x ADT95548 (1-674)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         54 CCAGATGTGATATTTCTACAGGAAGTT 80

RESULT 17
ADX42030
ID   ADX42030 standard; cDNA; 674 BP.
XX
AC   ADX42030;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1067.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1067; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.:              1.9             Length:         674
Score:                  45.00           Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     11              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x ADX42030 (1-674)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         54 CCAGATGTGATATTTCTACAGGAAGTT 80

RESULT 18
AAI24484
ID   AAI24484 standard; DNA; 691 BP.
XX
AC   AAI24484;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #14417 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 14417; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 691 BP; 230 A; 81 C; 112 G; 268 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.96            Length:         691
Score:                  45.00           Matches:        9
Percent Similarity:     100.0%          Conservative:   0
Best Local Similarity:  100.0%          Mismatches:     0
Query Match:            100.0%          Indels:         0
DB:                     4               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AAI24484 (1-691)
```

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        122 CCAGATGTGATATTTCTACAGGAAGTT 148
```

RESULT 19
AAH08073
ID AAH08073 standard; cDNA; 752 BP.
XX
AC AAH08073;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:4908.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 1; SEQ ID NO 4908; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 752 BP; 225 A; 148 C; 200 G; 176 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	2.16	Length:	752
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AAH08073 (1-752)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459
```

RESULT 20
AAX84209/c
ID AAX84209 standard; cDNA; 1079 BP.
XX
AC AAX84209;
XX
DT 08-SEP-1999 (first entry)
XX
DE DNA encoding human breast tumour protein immunogenic fragment.
XX
KW Breast tumour protein; immunogenic fragment; vaccine; detection;
KW breast cancer development; therapy; ss.
XX
OS Homo sapiens.
XX
PN WO9933869-A2.
XX
PD 08-JUL-1999.
XX
PF 22-DEC-1998; 98WO-US027416.
XX
PR 24-DEC-1997; 97US-00998253.
PR 24-DEC-1997; 97US-00998255.
PR 17-JUL-1998; 98US-00118554.
PR 17-JUL-1998; 98US-00118627.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J;
XX
DR WPI; 1999-405486/34.
XX
PT New breast tumor protein genes used, in vaccines for immunotherapy, or
PT for diagnosis of breast cancer.
XX
PS Claim 3; Page 60; 70pp; English.
XX
CC This sequence encodes a human breast tumour protein immunogenic fragment
CC of the invention. The polypeptides or nucleic acids encoding them are
CC useful in vaccines and pharmaceutical compositions for manufacture of
CC medicaments for inhibiting the development of breast cancer in a patient.
CC They can also be used to treat breast cancer. Antibodies against these
CC polypeptides can be used to detect and monitor progression of breast
CC cancer in patients. Primers and probes derived from the polynucleotides
CC encoding the breast proteins are useful for detection of breast cancer.
CC Peripheral blood cells from a patient incubated in the presence of at
CC least one polypeptide, such that T cells proliferate, are useful in
CC manufacture of a medicament for treating breast cancer in a patient.
CC Antigen presenting cells incubated in the presence of at least one
CC polypeptide are also useful for treating breast cancer
XX
SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.3	Length:	1079
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AAX84209 (1-1079)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024

RESULT 21
AAC79438/c
ID   AAC79438 standard; cDNA; 1079 BP.
XX
AC   AAC79438;
XX
DT   07-FEB-2001  (first entry)
XX
DE   cDNA sequence of human breast tumour clone 1015D11.
XX
KW   Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;
KW   vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061756-A2.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009688.
XX
PR   09-APR-1999;   99US-00288950.
PR   02-JUL-1999;   99US-00346327.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J,  Dillon DC;
XX
DR   WPI; 2000-638568/61.
XX
PT   A novel isolated polypeptide comprising an immunogenic portion of a
PT   breast cancer protein useful in the detection and treatment of breast
PT   cancer.
XX
PS   Claim 4; Page 77; 95pp; English.
XX
CC   The present sequence was isolated from a breast tumour cDNA library. It
CC   is provided in a specification relating to compounds for immunotherapy
CC   and diagnosis of breast cancer. Breast tumour antigens and the
CC   polynucleotides that encode them may be used in the production of a
CC   pharmaceutical composition to be used in the treatment of breast cancer.
CC   Proliferated T cells and incubated antigen presenting cells are also
CC   required. The polypeptides and polynucleotides may also be used to
CC   produce a vaccine
XX
SQ   Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.3          Length:          1079
Score:                    45.00        Matches:         9
Percent Similarity:       100.0%       Conservative:    0
Best Local Similarity:    100.0%       Mismatches:      0
Query Match:              100.0%       Indels:          0
DB:                       3            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AAC79438 (1-1079)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024

RESULT 22
ABK28982/c
ID   ABK28982 standard; cDNA; 1079 BP.
XX
AC   ABK28982;
XX
DT   23-APR-2002  (first entry)
XX
DE   Human breast tumour polypeptide cDNA clone #11.
XX
KW   Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
KW   immunostimulant.
XX
OS   Homo sapiens.
XX
PN   WO200198339-A2.
XX
PD   27-DEC-2001.
XX
PF   12-JUN-2001; 2001WO-US019032.
XX
PR   22-JUN-2000; 2000US-00602877.
PR   12-OCT-2000; 2000US-00687507.
PR   06-FEB-2001; 2001US-00778381.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J,  Dillon DC,  Retter MW,  Harlocker SL;
XX
DR   WPI; 2002-147792/19.
XX
PT   Polynucleotides encoding breast tumor polypeptides, useful for treating
PT   breast cancer or stimulating an immune response.
XX
PS   Claim 1; Page 131; 150pp; English.
XX
CC   The invention relates to polynucleotides encoding breast tumour
CC   polypeptides. The sequences are useful for treating cancer, preferably
CC   breast cancer, in a patient or for stimulating an immune response. The
CC   polynucleotides and polypeptides are also useful in the diagnosis and
CC   monitoring of breast cancer. A method for detecting the presence of a
CC   cancer in a patient, comprises obtaining a biological sample from the
CC   patient, contacting the biological sample with a binding agent that binds
CC   to a breast tumour polypeptide, detecting in the sample an amount of
CC   polypeptide that binds to the binding agent, and comparing the amount of
CC   polypeptide to a predetermined cut-off value, therefore determining the
CC   presence of a cancer in the patient. Sequences ABK28920-ABK29025
CC   represent cDNA clones encoding human breast tumour polypeptides of the
CC   invention
XX
SQ   Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.3          Length:          1079
Score:                    45.00        Matches:         9
Percent Similarity:       100.0%       Conservative:    0
Best Local Similarity:    100.0%       Mismatches:      0
Query Match:              100.0%       Indels:          0
DB:                       6            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ABK28982 (1-1079)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024

RESULT 23
AAS86254
ID   AAS86254 standard; cDNA; 1088 BP.
XX
AC   AAS86254;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #22058.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
```

XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG22067.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 22058; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1088 BP; 345 A; 182 C; 236 G; 325 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.33	Length:	1088
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AAS86254 (1-1088)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         37 CCAGATGTGATATTTCTACAGGAAGTT 63
```

RESULT 24
AAC98160
ID AAC98160 standard; cDNA; 1296 BP.
XX
AC AAC98160;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human colon cancer antigen nucleotide sequence SEQ ID NO:170.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200055351-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005883.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587534/55.
DR P-PSDB; AAB53403.
XX
PT Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer.
XX
PS Claim 1; Page 597; 2104pp; English.
XX
CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins may
CC also be used to prevent diseases such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, wounds, renal disorders, infectious diseases, and
CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 1296 BP; 376 A; 250 C; 326 G; 333 T; 0 U; 11 Other;

Alignment Scores:

Pred. No.:	4.08	Length:	1296
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AAC98160 (1-1296)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        456 CCAGATGTGATATTTCTACAGGAAGTT 482
```

RESULT 25
AAH15146
ID AAH15146 standard; cDNA; 1898 BP.
XX
AC AAH15146;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:13209.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 13209; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 1898 BP; 590 A; 327 C; 426 G; 555 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.37	Length:	1898
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AAH15146 (1-1898)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        433 CCAGATGTGATATTTCTACAGGAAGTT 459
```

RESULT 26
AAZ47118
ID AAZ47118 standard; cDNA; 1920 BP.
XX
AC AAZ47118;
XX
DT 15-MAR-2000 (first entry)
XX
DE Human CD40 receptor associated protein gene.
XX
KW Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW human; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW graft versus host disease; autoimmune disease; ds.
XX
OS Homo sapiens.
XX
PN WO9955859-A2.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1999; 99WO-EP003025.
XX
PR 29-APR-1998; 98EP-00201392.
XX
PA (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Pype SMC, Remacle JEFJG, Huylebroeck DFE;
XX
DR WPI; 2000-062029/05.
DR P-PSDB; AAY56019.
XX
PT Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT diseases and for improvement of anti-tumor treatments.
XX
PS Claim 9; Page 37-39; 48pp; English.
XX
CC This sequence represents the gene encoding human CD40 receptor associated
CC protein (CRAP). CRAP is a functional protein capable of interacting with
CC the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC I, where the protein has no homology to TNF receptor associated factor
CC (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC diseases, and for the improvement of anti-tumour diseases. Diseases which
CC may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC systemic lupus erythematosus, graft rejection, graft versus host disease,
CC allergy, and autoimmune disease. The proteins can be used to sensitize
CC tumour cells to anti-tumour treatments and to screen for compounds which
CC interfere with the interaction of the proteins with other protein
CC components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ Sequence 1920 BP; 599 A; 327 C; 435 G; 557 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	6.46	Length:	1920
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AAZ47118 (1-1920)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        452 CCAGATGTGATATTTCTACAGGAAGTT 478
```

RESULT 27
ADD19013
ID ADD19013 standard; DNA; 1936 BP.
XX
AC ADD19013;

XX
DT 15-JAN-2004 (first entry)
XX
DE Human disease related protein DNA sequence SeqID502.
XX
KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW antiarteriosclerotic; vulnerary; gene therapy;
KW hypoxia-regulated condition; tumourigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003018621-A2.
XX
PD 06-MAR-2003.
XX
PF 23-AUG-2002; 2002WO-GB003892.
XX
PR 23-AUG-2001; 2001GB-00020558.
PR 05-OCT-2001; 2001GB-00024037.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
DR WPI; 2003-290046/28.
DR P-PSDB; ADD19012.
XX
PT New substantially purified polypeptide, useful for diagnosing or treating
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT wound healing.
XX
PS Claim 27; SEQ ID NO 502; 424pp; English.
XX
CC This invention relates to novel human genes and gene product which are
CC implicated in certain disease states. Compounds which modulate the
CC proteins of the invention may have cytostatic, antiinflammatory,
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The
CC sequences of the invention may be useful for gene therapy. The invention
CC may be useful for diagnosing or treating a hypoxia-regulated condition,
CC such as tumourigenesis, angiogenesis, apoptosis, inflammation,
CC erythropoiesis, or the biological response to hypoxia conditions
CC including processes such as glycolysis, gluconeogenesis, glucose
CC transportation, catecholamine synthesis, iron transport or nitric oxide
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC inflammatory conditions or wound healing. The present sequence is that of
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.52	Length:	1936
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADD19013 (1-1936)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        449 CCAGATGTGATATTTCTACAGGAAGTT 475
```

RESULT 28
ADP25361
ID ADP25361 standard; cDNA; 1936 BP.
XX
AC ADP25361;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide encoding cDNA SEQ ID NO:475.
XX
KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-419628/39.
DR P-PSDB; ADP25362.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 1; SEQ ID NO 475; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence encodes a PRO protein of
CC the invention.
XX
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.52	Length:	1936
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0

```
DB:                    13            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ADP25361 (1-1936)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        449 CCAGATGTGATATTTCTACAGGAAGTT 475

RESULT 29
ADX06356
ID   ADX06356 standard; DNA; 1940 BP.
XX
AC   ADX06356;
XX
DT   21-APR-2005  (first entry)
XX
DE   Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 921.
XX
KW   cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2005012875-A2.
XX
PD   10-FEB-2005.
XX
PF   29-JUL-2004; 2004WO-US024424.
XX
PR   29-JUL-2003; 2003US-0490890P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Li M,  Rupnow BA,  Webster KR,  Jackson DG,  Wong TW;
XX
DR   WPI; 2005-163068/17.
DR   P-PSDB; ADX06357.
XX
PT   Biomarkers useful for predicting or determining the response of a mammal
PT   to a cancer treatment comprising administration of a modulator of cyclin-
PT   dependent kinase activity.
XX
PS   Claim 5; SEQ ID NO 921; 141pp; English.
XX
CC   This invention describes a novel method of predicting or determining
CC   whether a mammal will respond or is responding to an anti-cancer agent
CC   that modulates cyclin-dependent kinase (cdk) activity. The method
CC   comprises measuring the level of one or more biomarkers selected from
CC   2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC   NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC   invention is utilized in a kit for determining or predicting whether
CC   patient would be susceptible or resistant to treatment by an agent
CC   modulating cdk activity. The invention also describes a method for
CC   utilizing individualized genetic profiles for treating diseases and
CC   disorders based on patient's response and molecular level, specialized
CC   microarrays comprising the biomarkers described, antibodies directed
CC   against the biomarkers and a cell culture model to identify biomarkers.
CC   The cdk modulator is preferably N-5-[[5-(1,1-Dimethylethyl)-2-
CC   oxazolyl]methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC   tartaric acid salt. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC   sequence encodes a biomarker used in the method of the invention.
XX
SQ   Sequence 1940 BP; 613 A; 330 C; 439 G; 558 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.54          Length:        1940
Score:                  45.00         Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            100.0%        Indels:        0
DB:                     14            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ADX06356 (1-1940)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        458 CCAGATGTGATATTTCTACAGGAAGTT 484

RESULT 30
AAI58997
ID   AAI58997 standard; cDNA; 1948 BP.
XX
AC   AAI58997;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 1200.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J,  Zhao QA;
PI   Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39841.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders such
PT   as central nervous system injuries.
XX
PS   Claim 1; SEQ ID NO 1200; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC   encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders. Note: The sequence data for this patent did not form
CC   part of the printed specification
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:                6.57          Length:          1948
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       4             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AAI58997 (1-1948)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        471 CCAGATGTGATATTTCTACAGGAAGTT 497

RESULT 31
ADQ99219
ID   ADQ99219 standard; cDNA; 1948 BP.
XX
AC   ADQ99219;
XX
DT   23-SEP-2004  (first entry)
XX
DE   DNA encoding human GPCR-like protein seqid 889.
XX
KW   ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
KW   antidiabetic; GPCR-like protein; ophthalmic disorder;
KW   neurological disorder; immunological disorder; nephritic disorder;
KW   hormonal dysfunction; cancer; atherosclerosis; diabetes;
KW   molecular weight marker; food supplement; human; ss.
XX
OS   Homo sapiens.
XX
PN   US6569662-B1.
XX
PD   27-MAY-2003.
XX
PF   19-JUL-2000; 2000US-00620312.
XX
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Zhou P,  Drmanac RT;
XX
DR   WPI; 2001-442255/47.
XX
PT   New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT   useful for treating diseases of ophthalmic, neurological, immunological
PT   and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT   and diabetes.
XX
PS   Example 2; SEQ ID NO 889; 92pp; English.
XX
CC   The invention describes an isolated polynucleotide (I) comprising a fully
CC   defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC   1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC   given in the specification, its translated or protein coding portion, its
CC   extracellular portion or its active domain. The GPCR-like polypeptides
CC   and polynucleotides are useful for the treatment of diseases of
CC   ophthalmic, neurological, immunological and nephritic systems. They may
CC   also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC   diabetes. The antibodies are useful for detecting or quantitating the
CC   polypeptide in tissue. The polypeptides can also be used as molecular
CC   weight markers and as a food supplement. This sequence represents a human
CC   polynucleotide of the invention.
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.57          Length:          1948
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ADQ99219 (1-1948)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        471 CCAGATGTGATATTTCTACAGGAAGTT 497

RESULT 32
ADB48979
ID   ADB48979 standard; cDNA; 1948 BP.
XX
AC   ADB48979;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Novel human cDNA SEQ ID NO 889.
XX
KW   ss; cancer; neurodegenerative disease; human.
XX
OS   Homo sapiens.
XX
PN   US2003104529-A1.
XX
PD   05-JUN-2003.
XX
PF   04-JAN-2002; 2002US-00037270.
XX
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   19-JUL-2000; 2000US-00620312.
XX
PA   (ZHOU/) ZHOU P.
PA   (TANG/) TANG Y T.
PA   (LIUC/) LIU C.
PA   (ASUN/) ASUNDI V.
PA   (DRMA/) DRMANAC R T.
XX
PI   Zhou P,  Tang YT,  Liu C,  Asundi V,  Drmanac RT;
XX
DR   WPI; 2003-678194/64.
XX
PT   New polynucleotide, useful for treating diseases e.g., cancer or
PT   neurodegenerative diseases.
XX
PS   Claim 1; SEQ ID NO 889; 99pp; English.
XX
CC   The invention relates to a polynucleotide comprising a sequence given in
CC   the specification, or its mature protein-coding portion, or its
CC   complement. The polynucleotide is useful for treating diseases e.g.,
CC   cancer or neurodegenerative diseases and many others listed in the
CC   specification. The present sequence represents a novel human cDNA. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=20030104529.
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.57          Length:          1948
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       9             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ADB48979 (1-1948)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        471 CCAGATGTGATATTTCTACAGGAAGTT 497
```

```
RESULT 33
AAI60783
ID   AAI60783 standard; cDNA; 1958 BP.
XX
AC   AAI60783;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4772.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J,  Zhao QA;
PI   Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41627.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders such
PT   as central nervous system injuries.
XX
PS   Claim 1; SEQ ID NO 4772; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC   encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders. Note: The sequence data for this patent did not form
CC   part of the printed specification
XX
SQ   Sequence 1958 BP; 613 A; 339 C; 442 G; 564 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.61          Length:          1958
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       4             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AAI60783 (1-1958)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        483 CCAGATGTGATATTTCTACAGGAAGTT 509

RESULT 34
AAX28153
ID   AAX28153 standard; DNA; 2499 BP.
XX
AC   AAX28153;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Topoisomerase II binding protein 1 coding sequence.
XX
KW   Topoisomerase II binding protein; TopBP; anticancer agent; ds.
XX
OS   Homo sapiens.
XX
PN   JP11075856-A.
XX
PD   23-MAR-1999.
XX
PF   17-SEP-1997;   97JP-00251544.
XX
PR   17-SEP-1997;   97JP-00251544.
XX
PA   (TSUR/) TSURUO T.
PA   (CHUS ) CHUGAI PHARM CO LTD.
XX
DR   WPI; 1999-257704/22.
DR   P-PSDB; AAY03182.
XX
PT   New Topoisomerase II- binding protein - useful as an anticancer agent.
XX
PS   Disclosure; Page 18-19; 28pp; Japanese.
XX
CC   This sequence encodes the topoisomerase II binding protein (TopBP) of the
CC   invention. The TopBP protein is useful as an anticancer agent. TopBP can
CC   be used as the target molecule for anticancer agent
XX
SQ   Sequence 2499 BP; 720 A; 505 C; 587 G; 687 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                8.78          Length:          2499
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       2             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AAX28153 (1-2499)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        914 CCAGATGTGATATTTCTACAGGAAGTT 940

RESULT 35
ADA10970
ID   ADA10970 standard; cDNA; 3152 BP.
XX
AC   ADA10970;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human cDNA differentially expressed in colon cancer #60.
XX
KW   ss; differential expression; colon cancer; cancer; human.
XX
```

OS Homo sapiens.
XX
PN US2002160382-A1.
XX
PD 31-OCT-2002.
XX
PF 11-OCT-2001; 2001US-00981353.
XX
PR 11-OCT-2000; 2000US-0239841P.
XX
PA (LASE/) LASEK A W.
PA (JONE/) JONES D A.
XX
PI Lasek AW, Jones DA;
XX
DR WPI; 2003-265756/26.
XX
PT New combination comprising cDNAs that are differentially expressed in
PT colon disorder, useful for diagnosing, treating, staging or monitoring
PT treatment for colon cancers.
XX
PS Claim 1; SEQ ID NO 88; 231pp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in colon disorder. The methods and compositions
CC of the present invention are useful for diagnosing, treating, staging or
CC monitoring treatment for colon cancer. They are also useful in high
CC throughput methods for using cDNAs to detect differential expression of
CC nucleic acids in a sample, screening molecules or compounds to identify a
CC ligand which specifically binds a cDNA and using a protein to screen
CC molecules or compounds to identify at least one ligand which specifically
CC binds the protein. The present sequence represents a human cDNA
CC differentially expressed in colon cancer.
XX
SQ Sequence 3152 BP; 875 A; 643 C; 707 G; 926 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	11.5	Length:	3152
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADA10970 (1-3152)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1684 CCAGATGTGATATTTCTACAGGAAGTT 1710
```

RESULT 36
ADD34376
ID ADD34376 standard; DNA; 602 BP.
XX
AC ADD34376;
XX
DT 15-JAN-2004 (first entry)
XX
DE Mouse mitochondrial DNA sequence SEQ ID NO:2154.
XX
KW ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW mitochondrial disease; oxidative phosphorylation dysfunction;
KW oxidative stress; apoptosis; aging.
XX
OS Mus musculus.
XX
PN WO2003020220-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027886.
XX
PR 30-AUG-2001; 2001US-0316323P.
PR 31-AUG-2001; 2001CA-02356540.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Wallace DC, Levy S, Kerstann K, Procaccio V;
XX
DR WPI; 2003-300821/29.
XX
PT Array containing probes for genes involved in mitochondrial biology,
PT useful for determining mitochondrial biology gene expression profiles for
PT use in diagnosing pathologies and identifying biochemical pathways.
XX
PS Claim 2; SEQ ID NO 2154; 201pp; English.
XX
CC The invention relates to a novel array comprising at least two isolated
CC nucleotide molecules, each molecule having a sequence capable of uniquely
CC hybridising to a nucleic acid molecule which is an expression product of
CC a gene involved in mitochondrial biology. The array comprises two or more
CC isolated nucleic acid molecules or spots, each molecule having a sequence
CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC of the invention is useful for determining an expression profile of a
CC mouse or human sample containing nucleic acid, by contacting the array
CC with the sample under conditions allowing selective hybridisation, and
CC measuring hybridisation of nucleic acid in the sample to the array to
CC produce an expression profile. The array is also useful for determining
CC an expression profile of a first labelled sample containing nucleic acid
CC relative to a second, differently labelled sample containing nucleic
CC acid. The second sample is a reference or a standard. An array is useful
CC for determining an expression profile diagnostic of an energy-metabolism-
CC related physiological condition. An array of the invention is useful for
CC determining mitochondrial biology gene expression profiles of organisms,
CC such as human, mice and closely related species, tissue and organs of
CC such organisms, which are useful for determining expression profiles
CC diagnostic of energy metabolism-related physiological conditions,
CC diagnosing such physiological conditions, identifying biochemical
CC pathways, genes, and mutations involved in such physiological conditions,
CC identifying therapeutic agents useful for preventing and/or treating such
CC physiological conditions, evaluating and/or monitoring the efficacy of
CC such therapies, and creating and identifying animal models of human
CC energy metabolism-related physiological conditions. An array is also
CC useful for defining expression signatures or profiles for mitochondrial
CC diseases, as well as distinguishing clinical disorders that result from
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC apoptosis and aging. An array of the invention contains probes of genes
CC not previously recognised to participate in mitochondrial biology. The
CC sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC clones used to make the probes of the invention. Some sequences are not
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC 1906, 2408 and 2643.
XX
SQ Sequence 602 BP; 194 A; 119 C; 134 G; 155 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.81	Length:	602
Score:	44.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	88.9%	Mismatches:	0
Query Match:	97.8%	Indels:	0
DB:	10	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADD34376 (1-602)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||
Db        228 CCAGATGTGGTATTTCTACAGGAAGTT 254
```

RESULT 37
AAZ47119
ID AAZ47119 standard; cDNA; 1312 BP.
XX
AC AAZ47119;
XX
DT 15-MAR-2000 (first entry)

```
XX
DE   Mouse CD40 receptor associated protein gene.
XX
KW   Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW   immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW   mouse; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW   tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW   TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW   diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW   arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW   graft versus host disease; autoimmune disease; ds.
XX
OS   Mus musculus.
XX
PN   WO9955859-A2.
XX
PD   04-NOV-1999.
XX
PF   28-APR-1999;   99WO-EP003025.
XX
PR   29-APR-1998;   98EP-00201392.
XX
PA   (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI   Pype SMC,  Remacle JEFJG,  Huylebroeck DFE;
XX
DR   WPI; 2000-062029/05.
DR   P-PSDB; AAY56020.
XX
PT   Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT   diseases and for improvement of anti-tumor treatments.
XX
PS   Claim 10; Page 41-43; 48pp; English.
XX
CC   This sequence represents the gene encoding mouse CD40 receptor associated
CC   protein (CRAP). CRAP is a functional protein capable of interacting with
CC   the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC   necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC   I, where the protein has no homology to TNF receptor associated factor
CC   (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC   the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC   related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC   diseases, and for the improvement of anti-tumour diseases. Diseases which
CC   may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC   systemic lupus erythematosus, graft rejection, graft versus host disease,
CC   allergy, and autoimmune disease. The proteins can be used to sensitize
CC   tumour cells to anti-tumour treatments and to screen for compounds which
CC   interfere with the interaction of the proteins with other protein
CC   components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ   Sequence 1312 BP; 359 A; 279 C; 352 G; 322 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.97          Length:          1312
Score:                    44.00         Matches:         8
Percent Similarity:       100.0%        Conservative:    1
Best Local Similarity:    88.9%         Mismatches:      0
Query Match:              97.8%         Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AAZ47119 (1-1312)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||
Db        584 CCAGATGTGGTATTTCTACAGGAAGTT 610

RESULT 38
ADT95029
ID   ADT95029 standard; cDNA; 625 BP.
XX
AC   ADT95029;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #548.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 548; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:                23.5          Length:          625
Score:                    40.00         Matches:         8
Percent Similarity:       88.9%         Conservative:    0
Best Local Similarity:    88.9%         Mismatches:      1
Query Match:              88.9%         Indels:          0
DB:                       11            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ADT95029 (1-625)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||   |||||||||||||||
```

```
Db         62 CCAGATGTGNTATTTCTACAGGAAGTT 88

RESULT 39
ADX41511
ID   ADX41511 standard; cDNA; 625 BP.
XX
AC   ADX41511;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 548.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 548; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              23.5          Length:         625
Score:                  40.00         Matches:        8
Percent Similarity:     88.9%         Conservative:   0
Best Local Similarity:  88.9%         Mismatches:     1
Query Match:            88.9%         Indels:         0
DB:                     11            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x ADX41511 (1-625)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||   |||||||||||||||
Db         62 CCAGATGTGNTATTTCTACAGGAAGTT 88

RESULT 40
AEB39175_05
Continuation (6 of 35) of AEB39175 from base 500001 (L. pneumophila DNA SEQ ID NO 3507.
WP Sequence split into 35 fragments  LOCUS AEB39175 Accession Aeb39175
WP      Fragment Name       Begin      End
WP      AEB39175_00             1    110000
WP      AEB39175_01        100001    210000
WP      AEB39175_02        200001    310000
WP      AEB39175_03        300001    410000
WP      AEB39175_04        400001    510000
WP      AEB39175_05        500001    610000
WP      AEB39175_06        600001    710000
WP      AEB39175_07        700001    810000
WP      AEB39175_08        800001    910000
WP      AEB39175_09        900001   1010000
WP      AEB39175_10       1000001   1110000
WP      AEB39175_11       1100001   1210000
WP      AEB39175_12       1200001   1310000
WP      AEB39175_13       1300001   1410000
WP      AEB39175_14       1400001   1510000
WP      AEB39175_15       1500001   1610000
WP      AEB39175_16       1600001   1710000
WP      AEB39175_17       1700001   1810000
WP      AEB39175_18       1800001   1910000
WP      AEB39175_19       1900001   2010000
WP      AEB39175_20       2000001   2110000
WP      AEB39175_21       2100001   2210000
WP      AEB39175_22       2200001   2310000
WP      AEB39175_23       2300001   2410000
WP      AEB39175_24       2400001   2510000
WP      AEB39175_25       2500001   2610000
WP      AEB39175_26       2600001   2710000
WP      AEB39175_27       2700001   2810000
WP      AEB39175_28       2800001   2910000
WP      AEB39175_29       2900001   3010000
WP      AEB39175_30       3000001   3110000
WP      AEB39175_31       3100001   3210000
WP      AEB39175_32       3200001   3310000
WP      AEB39175_33       3300001   3410000
WP      AEB39175_34       3400001   3503610

Alignment Scores:
Pred. No.:              1.65e+04      Length:         110000
Score:                  39.00         Matches:        7
Percent Similarity:     88.9%         Conservative:   1
Best Local Similarity:  77.8%         Mismatches:     1
Query Match:            86.7%         Indels:         0
DB:                     14            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AEB39175_05 (1-110000)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||   :::|||||||||||||||
Db      83038 CCGGACTTTGTTTTTTTACAAGAAGTG 83064

RESULT 41
AEB42401_05
Continuation (6 of 34) of AEB42401 from base 500001 (L. pneumophila DNA SEQ ID NO 6733.
WP Sequence split into 34 fragments  LOCUS AEB42401 Accession Aeb42401
WP      Fragment Name       Begin      End
WP      AEB42401_00             1    110000
WP      AEB42401_01        100001    210000
WP      AEB42401_02        200001    310000
WP      AEB42401_03        300001    410000
WP      AEB42401_04        400001    510000
WP      AEB42401_05        500001    610000
WP      AEB42401_06        600001    710000
WP      AEB42401_07        700001    810000
WP      AEB42401_08        800001    910000
WP      AEB42401_09        900001   1010000
WP      AEB42401_10       1000001   1110000
WP      AEB42401_11       1100001   1210000
WP      AEB42401_12       1200001   1310000
WP      AEB42401_13       1300001   1410000
WP      AEB42401_14       1400001   1510000
WP      AEB42401_15       1500001   1610000
WP      AEB42401_16       1600001   1710000
```

```
WP      AEB42401_17       1700001   1810000
WP      AEB42401_18       1800001   1910000
WP      AEB42401_19       1900001   2010000
WP      AEB42401_20       2000001   2110000
WP      AEB42401_21       2100001   2210000
WP      AEB42401_22       2200001   2310000
WP      AEB42401_23       2300001   2410000
WP      AEB42401_24       2400001   2510000
WP      AEB42401_25       2500001   2610000
WP      AEB42401_26       2600001   2710000
WP      AEB42401_27       2700001   2810000
WP      AEB42401_28       2800001   2910000
WP      AEB42401_29       2900001   3010000
WP      AEB42401_30       3000001   3110000
WP      AEB42401_31       3100001   3210000
WP      AEB42401_32       3200001   3310000
WP      AEB42401_33       3300001   3345687

Alignment Scores:
Pred. No.:              1.65e+04        Length:         110000
Score:                  39.00           Matches:        7
Percent Similarity:     88.9%           Conservative:   1
Best Local Similarity:  77.8%           Mismatches:     1
Query Match:            86.7%           Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AEB42401_05 (1-110000)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||   :::|||||||||||||||
Db      59104 CCGGACTTTGTTTTTTTACAAGAAGTG 59130

RESULT 42
AEB42735
ID   AEB42735 standard; DNA; 243335 BP.
XX
AC   AEB42735;
XX
DT   08-SEP-2005  (first entry)
XX
DE   L. pneumophila DNA SEQ ID NO 7067.
XX
KW   detection; infection; Antibacterial; Vaccine; ds; gene.
XX
OS   Legionella pneumophila.
XX
PN   WO2005049642-A2.
XX
PD   02-JUN-2005.
XX
PF   23-SEP-2004; 2004WO-IB003578.
XX
PR   21-NOV-2003; 2003FR-00013687.
XX
PA   (INSP ) INST PASTEUR.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA   (UYLY-) UNIV LYON 1 BERNARD CLAUDE.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Buchrieser C,  Tichit M,  Etienne J,  Ma L,  Cazalet C,  Glaser P;
PI   Rusniok C,  Bouchier C,  Zidane N,  Magnier A,  Kunst F,  Vandenesch F;
PI   Jarraud S;
XX
DR   WPI; 2005-388305/40.
XX
PT   New genome of Legionella pneumophila Paris strain and derived
PT   polypeptides, useful for detection or identification of the strain and
PT   for treatment and prevention of infections.
XX
PS   Example 9; SEQ ID NO 7067; 660pp; English.
XX
CC   The invention relates to an isolated or purified nucleotide sequences (I)
CC   from Legionella pneumophila Paris strain. (I), and their related
CC   sequences or fragments, are useful as primers and probes for detection
CC   and amplification, including differentiation between the Paris and
CC   Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC   (hybrid) polypeptides (II). (II) are also useful for preparation of
CC   specific antibodies (Ab), also used for detection/identification of
CC   Legionella, and some (I), specifically those involved in synthesis of
CC   surface proteins, are targets for identification of inhibitors. (II), or
CC   vectors that contain (I), are useful as vaccines and immunogenic
CC   compositions, for treatment and prevention of infections by L.
CC   pneumophila. The present squence represents a L. pneumophila DNA.
XX
SQ   Sequence 243335 BP; 73649 A; 42176 C; 51283 G; 76227 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.16e+04        Length:         243335
Score:                  39.00           Matches:        7
Percent Similarity:     88.9%           Conservative:   1
Best Local Similarity:  77.8%           Mismatches:     1
Query Match:            86.7%           Indels:         0
DB:                     14              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AEB42735 (1-243335)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||   :::|||||||||||||||
Db     147039 CCGGACTTTGTTTTTTTACAAGAAGTG 147065

RESULT 43
AEB35721
ID   AEB35721 standard; DNA; 295644 BP.
XX
AC   AEB35721;
XX
DT   08-SEP-2005  (first entry)
XX
DE   L. pneumophila DNA SEQ ID NO 53.
XX
KW   detection; infection; Antibacterial; Vaccine; ds; gene.
XX
OS   Legionella pneumophila.
XX
PN   WO2005049642-A2.
XX
PD   02-JUN-2005.
XX
PF   23-SEP-2004; 2004WO-IB003578.
XX
PR   21-NOV-2003; 2003FR-00013687.
XX
PA   (INSP ) INST PASTEUR.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA   (UYLY-) UNIV LYON 1 BERNARD CLAUDE.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Buchrieser C,  Tichit M,  Etienne J,  Ma L,  Cazalet C,  Glaser P;
PI   Rusniok C,  Bouchier C,  Zidane N,  Magnier A,  Kunst F,  Vandenesch F;
PI   Jarraud S;
XX
DR   WPI; 2005-388305/40.
XX
PT   New genome of Legionella pneumophila Paris strain and derived
PT   polypeptides, useful for detection or identification of the strain and
PT   for treatment and prevention of infections.
XX
PS   Claim 1; SEQ ID NO 53; 660pp; English.
XX
CC   The invention relates to an isolated or purified nucleotide sequences (I)
CC   from Legionella pneumophila Paris strain. (I), and their related
CC   sequences or fragments, are useful as primers and probes for detection
CC   and amplification, including differentiation between the Paris and
CC   Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC   (hybrid) polypeptides (II). (II) are also useful for preparation of
CC   specific antibodies (Ab), also used for detection/identification of
```

CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present squence represents a L. pneumophila DNA.
XX
SQ Sequence 295644 BP; 87609 A; 52716 C; 63040 G; 92276 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	5.22e+04	Length:	295644
Score:	39.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	86.7%	Indels:	0
DB:	14	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AEB35721 (1-295644)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||   :::|||||||||||||||||
Db     112705 CCGGACTTTGTTTTTTTACAAGAAGTG 112731
```

RESULT 44
AEB39173
ID AEB39173 standard; DNA; 298667 BP.
XX
AC AEB39173;
XX
DT 08-SEP-2005 (first entry)
XX
DE L. pneumophila DNA SEQ ID NO 3505.
XX
KW detection; infection; Antibacterial; Vaccine; ds; gene.
XX
OS Legionella pneumophila.
XX
PN WO2005049642-A2.
XX
PD 02-JUN-2005.
XX
PF 23-SEP-2004; 2004WO-IB003578.
XX
PR 21-NOV-2003; 2003FR-00013687.
XX
PA (INSP) INST PASTEUR.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (UYLY-) UNIV LYON 1 BERNARD CLAUDE.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;
PI Rusniok C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
PI Jarraud S;
XX
DR WPI; 2005-388305/40.
XX
PT New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX
PS Example 9; SEQ ID NO 3505; 660pp; English.
XX
CC The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain. (I), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (II). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present squence represents a L. pneumophila DNA.
XX
SQ Sequence 298667 BP; 89051 A; 52428 C; 62548 G; 94640 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.28e+04	Length:	298667
Score:	39.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	86.7%	Indels:	0
DB:	14	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AEB39173 (1-298667)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||   :::|||||||||||||||||
Db     117386 CCGGACTTTGTTTTTTTACAAGAAGTG 117412
```

RESULT 45
ABN16687/c
ID ABN16687 standard; cDNA; 244 BP.
XX
AC ABN16687;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX polynucleotide sequence SEQ ID NO:1851.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US010836.
XX
PR 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach MD;
XX
DR WPI; 2002-106308/14.
DR P-PSDB; ABP00935.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
PS Disclosure; SEQ ID NO 1851; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester

```
CC   storage disease, various immune deficiencies and disorders, infectious
CC   diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC   arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC   disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC   useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC   bone degenerative disorders, or periodontal disease, and for gut
CC   protection or regeneration and treatment of lung or liver fibrosis,
CC   reperfusion injury in various tissues and conditions resulting from
CC   systemic cytokine damage. N.B. The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 244 BP; 33 A; 88 C; 84 G; 36 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:                22.2          Length:          244
Score:                    38.00         Matches:         7
Percent Similarity:       88.9%         Conservative:    1
Best Local Similarity:    77.8%         Mismatches:      1
Query Match:              84.4%         Indels:          0
DB:                       6             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ABN16687 (1-244)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   |||||||||||
Db        129 CCCGACGTCGTTCTCCTCCAGGAGGTG 103

RESULT 46
AAH65711
ID   AAH65711 standard; DNA; 912 BP.
XX
AC   AAH65711;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum coding sequence fragment SEQ ID NO: 746.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-00127688.
XX
PR   16-DEC-1999;   99JP-00377484.
PR   07-APR-2000; 2000JP-00159162.
PR   03-AUG-2000; 2000JP-00280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S,  Mizoguchi H,  Ando S,  Hayashi M,  Ochiai K,  Yokoi H;
PI   Tateishi N,  Senoh A,  Ikeda M,  Ozaki A;
XX
DR   WPI; 2001-376931/40.
DR   P-PSDB; AAG90492.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analyzing
PT   expression profile or pattern of a gene and identifying homologous gene.
XX
PS   Claim 8; SEQ ID NO 746; 246pp + Sequence Listing; English.
XX
CC   The present invention provides a number of nucleotide and protein
CC   sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC   are useful for identifying the mutation point of a gene derived from a
CC   mutant of coryneform bacterium, measuring expression amount and analysing
CC   the expression profile or expression pattern of a gene derived from
CC   Coryneform bacterium, and identifying a homologue of a gene derived from
CC   coryneform bacterium. Coryneform bacteria are useful for producing amino
CC   acids, nucleic acids, vitamins, saccharides and organic acids,
CC   particularly L-lysine. The present sequence is a nucleic acid described
CC   in the exemplification of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the European Patent Office
XX
SQ   Sequence 912 BP; 228 A; 276 C; 235 G; 173 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                103           Length:          912
Score:                    38.00         Matches:         7
Percent Similarity:       88.9%         Conservative:    1
Best Local Similarity:    77.8%         Mismatches:      1
Query Match:              84.4%         Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AAH65711 (1-912)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   |||||||||||
Db        106 CCGGACGTTGTCCTCCTCCAAGAAGTC 132

RESULT 47
ACA00116
ID   ACA00116 standard; DNA; 921 BP.
XX
AC   ACA00116;
XX
DT   02-JUN-2003  (first entry)
XX
DE   C. glutamicum derived ORF SEQ ID 107.
XX
KW   Coryneform; nucleic acid array; fermentation; culture; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   DE10128510-A1.
XX
PD   19-DEC-2002.
XX
PF   13-JUN-2001; 2001DE-01028510.
XX
PR   13-JUN-2001; 2001DE-01028510.
XX
PA   (DEGS ) DEGUSSA AG.
XX
PI   Farwick M,  Moeckel B,  Pfefferle W,  Bathe B,  Huthmacher K;
XX
DR   WPI; 2003-279970/28.
XX
PT   New nucleic acid array useful for monitoring mRNA expression of
PT   Corynebacterium glutamicum during fermentation, comprising nucleic acid
PT   from Corynebacterium glutamicum.
XX
PS   Claim 1; Page 110; 709pp; German.
XX
CC   This invention describes a novel nucleic acid array involving
CC   Corynebacterium glutamicum polynucleotides. The arrays are used to
CC   analyse C. glutamicum, particularly for monitoring a fermentation process
CC   to determine expression levels of C. glutamicum cellular mRNA. Such
CC   monitoring particularly differentiates between expression levels of
CC   different strains of C. glutamicum and allows the adjustment of different
CC   culture and fermentation conditions. ACA00010-ACA02188 represent C.
CC   glutamicum derived polynucleotides described in the disclosure of the
CC   invention
XX
SQ   Sequence 921 BP; 231 A; 278 C; 238 G; 174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                105           Length:          921
Score:                    38.00         Matches:         7
Percent Similarity:       88.9%         Conservative:    1
```

```
Best Local Similarity:  77.8%          Mismatches:    1
Query Match:            84.4%          Indels:        0
DB:                     8              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ACA00116 (1-921)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   ||||||||||||
Db        115 CCGGACGTTGTCCTCCTCCAAGAAGTC 141

RESULT 48
ADX50393
ID   ADX50393 standard; cDNA; 1838 BP.
XX
AC   ADX50393;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 25133.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 25133; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This sequence represents a plant full length insert
CC   polynucleotide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 1838 BP; 570 A; 318 C; 493 G; 457 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              234            Length:        1838
Score:                  38.00          Matches:       7
Percent Similarity:     100.0%         Conservative:  1
Best Local Similarity:  87.5%          Mismatches:    0
Query Match:            84.4%          Indels:        0
DB:                     13             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ADX50393 (1-1838)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||||||||||||||:::
Db       1034 CCAGATGTTATTTTTCTACAGCAA 1057

RESULT 49
AAH68526
ID   AAH68526 standard; DNA; 349980 BP.
XX
AC   AAH68526;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum coding sequence fragment SEQ ID NO: 7061.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis; ds.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-00127688.
XX
PR   16-DEC-1999;   99JP-00377484.
PR   07-APR-2000; 2000JP-00159162.
PR   03-AUG-2000; 2000JP-00280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S,  Mizoguchi H,  Ando S,  Hayashi M,  Ochiai K,  Yokoi H;
PI   Tateishi N,  Senoh A,  Ikeda M,  Ozaki A;
XX
DR   WPI; 2001-376931/40.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analyzing
PT   expression profile or pattern of a gene and identifying homologous gene.
XX
PS   Disclosure; SEQ ID NO 7061; 246pp + Sequence Listing; English.
XX
CC   The present invention provides a number of nucleotide and protein
CC   sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC   are useful for identifying the mutation point of a gene derived from a
CC   mutant of coryneform bacterium, measuring expression amount and analysing
CC   the expression profile or expression pattern of a gene derived from
CC   Coryneform bacterium, and identifying a homologue of a gene derived from
CC   coryneform bacterium. Coryneform bacteria are useful for producing amino
CC   acids, nucleic acids, vitamins, saccharides and organic acids,
CC   particularly L-lysine. The present sequence is a nucleic acid described
CC   in the exemplification of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the European Patent Office
XX
SQ   Sequence 349980 BP; 79274 A; 90638 C; 98727 G; 81341 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:                1.06e+05        Length:         349980
Score:                    38.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              84.4%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AAH68526 (1-349980)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   ||||||||||||
Db      86540 CCGGACGTTGTCCTCCTCCAAGAAGTC 86566

RESULT 50
ADE81339/c
ID   ADE81339 standard; cDNA; 516 BP.
XX
AC   ADE81339;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Arabidopsis thaliana expressed polynucleotide seq id 110.
XX
KW   genetically modified organism; transgenic organism; plant;
KW   inhibitor testing; activator testing; modifier testing; fungicide;
KW   insecticide; genetic function; genetic regulation; cellular metabolism;
KW   gene; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   US2003115639-A1.
XX
PD   19-JUN-2003.
XX
PF   26-JAN-2001; 2001US-00770961.
XX
PR   27-JAN-2000; 2000US-0178466P.
XX
PA   (GORL/) GORLACH J.
PA   (ANYY/) AN Y.
PA   (HAMI/) HAMILTON C M.
PA   (PRIC/) PRICE J L.
PA   (RAIN/) RAINES T M.
PA   (YUYY/) YU Y.
PA   (RAME/) RAMEAKA J G.
PA   (PAGE/) PAGE A.
PA   (MATH/) MATHEW A V.
PA   (LEDF/) LEDFORD B L.
PA   (WOES/) WOESSNER J P.
PA   (HAAS/) HAAS W D.
PA   (GARC/) GARCIA C A.
PA   (KRIC/) KRICKER M.
PA   (SLAT/) SLATER T.
PA   (DAVI/) DAVIS K R.
PA   (ALLE/) ALLEN K.
PA   (HOFF/) HOFFMAN N.
PA   (HURB/) HURBAN P.
XX
PI   Gorlach J,  An Y,  Hamilton CM,  Price JL,  Raines TM,  Yu Y;
PI   Rameaka JG,  Page A,  Mathew AV,  Ledford BL,  Woessner JP,  Haas WD;
PI   Garcia CA,  Kricker M,  Slater T,  Davis KR,  Allen K,  Hoffman N;
PI   Hurban P;
XX
DR   WPI; 2003-810930/76.
XX
PT   Novel Arabidopsis thaliana nucleic acids useful for generating
PT   genetically modified transgenic organisms, for screening biologically
PT   active agents such as fungicides, insecticides.
XX
PS   Claim 1; SEQ ID NO 110; 44pp; English.
XX
CC   The invention describes a nucleic acid (I) comprising a sequence capable
CC   of hybridising under stringent conditions to any one of 999 fully defined
CC   Arabidopsis thaliana sequences (1) as given in specification e.g., 360,
CC   1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a
CC   hybridisation probe to complementary molecules in a cDNA library. (I) is
CC   also useful for generating genetically modified and transgenic organisms,
CC   usually plant cells and plants. A proten encoded by (I) is useful in
CC   screening assays to determine the effect of candidate inhibitors,
CC   activators or modifiers of the gene product. The protein is also useful
CC   for screening biologically active agents e.g., fungicides and
CC   insecticides. A genetically modified cell, comprising an exogenous
CC   nucleic acid, where the nucleic acid comprises transcription regulatory
CC   sequences operably linked to a sequence capable of hybridising under
CC   stringent conditions to (1) is useful in the study of genetic function
CC   and regulation, for alteration of the cellular metabolism and for
CC   screening compounds that may affect the biological function of the gene
CC   or gene product. This sequence represents an Arabidopsis thaliana
CC   polynucleotide of the invention.
XX
SQ   Sequence 516 BP; 149 A; 110 C; 139 G; 118 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                89.5            Length:         516
Score:                    37.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              82.2%           Indels:         0
DB:                       10              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x ADE81339 (1-516)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||   ||||||
Db        478 CCTGACATCATCTTCCTTCTTGAGGTT 452

RESULT 51
ADN72674/c
ID   ADN72674 standard; cDNA; 753 BP.
XX
AC   ADN72674;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 569.
XX
KW   gene; ss; plant; transgenic; E2Fa/DPa transcription factor;
KW   growth regulator; animal feed product; thale cress;
KW   cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX
OS   Arabidopsis thaliana.
XX
PN   WO2004035798-A2.
XX
PD   29-APR-2004.
XX
PF   20-OCT-2003; 2003WO-EP011658.
XX
PR   18-OCT-2002; 2002EP-00079408.
XX
PA   (CROP-) CROPDESIGN NV.
XX
PI   Inze D,  De Veylder L,  Vlieghe K;
XX
DR   WPI; 2004-348466/32.
DR   P-PSDB; ADN72675.
XX
PT   Altering plant characteristics, useful for producing plants for enzyme or
PT   pharmaceutical production comprises modifying in a plant, expression of
PT   one or more nucleic acids and/or modifying level or activity of one or
PT   more proteins.
XX
PS   Claim 1; SEQ ID NO 569; 134pp; English.
XX
CC   This invention relates to a novel method for altering one or more plant
```

CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/DPa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/ or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is thale cress cDNA
CC upregulated 1.3 fold or more in plants overexpressing the E2Fa/DPa
CC transcription factor, given in an exemplification of the invention.
XX
SQ Sequence 753 BP; 223 A; 156 C; 205 G; 169 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	139	Length:	753
Score:	37.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	82.2%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADN72674 (1-753)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||:::|||||||||   ||||||
Db        468 CCTGACATCATCTTCCTTCTTGAGGTT 442
```

RESULT 52
AAC44301/c
ID AAC44301 standard; DNA; 766 BP.
XX
AC AAC44301;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 42361.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	142	Length:	766
Score:	37.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	82.2%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AAC44301 (1-766)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||   ||||||
Db        481 CCTGACATCATCTTCCTTCTTGAGGTT 455
```

RESULT 53
AAZ98318/c
ID AAZ98318 standard; DNA; 768 BP.
XX
AC AAZ98318;
XX
DT 14-JUN-2000 (first entry)
XX
DE A. thaliana gene involved in environmental stress tolerance.
XX
KW Environmental stress; plant; transgenic plant; anaerobic; flooding; cold;
KW dehydration; drought; heat stress; salinity; osmotolerance; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200008187-A2.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-EP005652.
XX
PR 04-AUG-1998; 98EP-00202634.
XX
PA (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Lee JH, Verbruggen N;
XX
DR WPI; 2000-205726/18.
DR P-PSDB; AAY77938.
XX
PT Isolation of polynucleic acids useful for producing transgenic plant by
PT isolating genes involved in tolerance to environmental stress.
XX
PS Claim 4; Page 133-134; 312pp; English.
XX
CC The invention relates to isolation of coding sequences and/or genes
CC involved in tolerance to environmental stress in plants. The sequences
CC (AAZ98305-Z98365) are useful for producing a transgenic plant having
CC enhanced tolerance or resistance to environmental stress conditions such
CC as anaerobic, flooding, cold, dehydration, drought, heat stress or

```
CC   salinity. This is useful for producing improved yield, growth,
CC   development and productivity under environmental stress conditions, and
CC   also provides growth of crops in areas where they cannot grow without the
CC   induced osmotolerance. Sequences AAZ98305-365 represent polynucleotide
CC   sequences from A. thaliana that are involved in environmental stress
CC   tolerance
XX
SQ   Sequence 768 BP; 229 A; 157 C; 209 G; 173 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                142             Length:         768
Score:                    37.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              82.2%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x AAZ98318 (1-768)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||   ||||||
Db        467 CCTGACATCATCTTCCTTCTTGAGGTT 441

RESULT 54
ADU92189/c
ID   ADU92189 standard; cDNA; 768 BP.
XX
AC   ADU92189;
XX
DT   27-JAN-2005  (first entry)
XX
DE   Thale cress environmental stress protein encoding gene, SEQ ID 27.
XX
KW   environmental; stress tolerance; transgenic plant; drought resistance;
KW   salt tolerance; cold tolerance; heat tolerance; freezing tolerance;
KW   crop improvement; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             3. .752
FT                   /*tag=  a
FT                   /product= "Arabidopsis thaliana environmental stress
FT                   protein"
XX
PN   AU2003236475-A1.
XX
PD   18-SEP-2003.
XX
PF   26-AUG-2003; 2003AU-00236475.
XX
PR   26-AUG-2003; 2003AU-00236475.
XX
PA   (CROP-) CROPDESIGN NV.
XX
PI   Verbruggen N,  Lee JH;
XX
DR   WPI; 2004-738142/73.
DR   P-PSDB; ADU92190.
XX
PT   Obtaining polynucleic acids with coding sequences involved in
PT   environmental stress in plants, by introducing coding sequences of cDNA
PT   library from siliques in yeast cells, screening stress tolerance
PT   polynucleic acids in transformed cells.
XX
PS   Claim 4; SEQ ID NO 27; 313pp; English.
XX
CC   The invention relates to a novel method for obtaining polynucleic acids,
CC   comprising coding sequences and/or genes involved in environmental stress
CC   in plants. The method involves preparing a cDNA library comprising coding
CC   sequences from siliques, introducing the coding sequences in yeast cells
CC   in a functional format and screening for polynucleic acids leading to an
CC   enhanced tolerance or resistance to environmental stress conditions in
CC   the transformed yeast cells. The method is useful for obtaining
CC   polynucleic acids, comprising coding sequences and/or genes involved in
CC   environmental stress in plants. An isolated polynucleic acid encoding an
CC   HSP 17.6A protein, obtained using the novel method, is useful for the
CC   production of transgenic plants having an enhanced tolerance or
CC   resistance to environmental stress conditions. Isolated polynucleic acids
CC   encoding plant DBF2 kinase or c74 protein can also be used to produce
CC   transgenic plants having an enhanced tolerance or resistance to
CC   environmental stress conditions. Recombinant nucleic acids produced using
CC   the novel method are useful in producing transgenic plants. The method is
CC   useful in obtaining nucleic acids involved in environmental stress, such
CC   as water stress (flooding, drought, dehydration), anaerobic (low level of
CC   oxygen, carbon dioxide), aerobic stress, osmotic stress, salt stress
CC   (hot/heat, cold, freezing, frost) or nutrient/pollutant stress). This
CC   polynucleotide sequence represents the gene encoding an Arabidopsis
CC   thaliana (thale cress) environmental stress protein of the invention.
XX
SQ   Sequence 768 BP; 229 A; 157 C; 209 G; 173 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                142             Length:         768
Score:                    37.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              82.2%           Indels:         0
DB:                       13              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x ADU92189 (1-768)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||   ||||||
Db        467 CCTGACATCATCTTCCTTCTTGAGGTT 441

RESULT 55
ADT47013
ID   ADT47013 standard; cDNA; 774 BP.
XX
AC   ADT47013;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Bacterial polynucleotide #21764.
XX
KW   Recombinant DNA construct; transformed plant; improved plant property;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW   pathogen tolerance; pest tolerance; plant disease resistance;
KW   cell cycle pathway modification; plant growth regulator;
KW   homologous recombination; seed oil yield; protein yield; carbohydrate;
KW   nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW   bacterial polynucleotide; gene; ss.
XX
OS   Bacteria.
XX
PN   US2003233675-A1.
XX
PD   18-DEC-2003.
XX
PF   20-FEB-2003; 2003US-00369493.
XX
PR   21-FEB-2002; 2002US-0360039P.
XX
PA   (CAOY/) CAO Y.
PA   (HINK/) HINKLE G J.
PA   (SLAT/) SLATER S C.
PA   (CHEN/) CHEN X.
PA   (GOLD/) GOLDMAN B S.
XX
PI   Cao Y,  Hinkle GJ,  Slater SC,  Chen X,  Goldman BS;
XX
DR   WPI; 2004-061375/06.
XX
PT   New recombinant DNA construct comprising a promoter positioned to provide
PT   for expression of a polynucleotide encoding a polypeptide from a
```

PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 45451; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 774 BP; 136 A; 251 C; 273 G; 114 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	144	Length:	774
Score:	37.00	Matches:	6
Percent Similarity:	88.9%	Conservative:	2
Best Local Similarity:	66.7%	Mismatches:	1
Query Match:	82.2%	Indels:	0
DB:	13	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADT47013 (1-774)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   |||||||||:::
Db         79 CCCGATGTCGTGCTGCTGCAGGAGATC 105
```

RESULT 56
ADS56055
ID ADS56055 standard; cDNA; 777 BP.
XX
AC ADS56055;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #8042.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 31729; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 777 BP; 122 A; 252 C; 281 G; 122 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	144	Length:	777
Score:	37.00	Matches:	6
Percent Similarity:	88.9%	Conservative:	2
Best Local Similarity:	66.7%	Mismatches:	1
Query Match:	82.2%	Indels:	0
DB:	13	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADS56055 (1-777)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   |||||||||:::
Db         85 CCTGACGTGGTGCTGCTGCAGGAGATC 111
```

RESULT 57
AAC40379/c
ID AAC40379 standard; DNA; 833 BP.
XX
AC AAC40379;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 28098.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.

```
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

Alignment Scores:
Pred. No.:                157             Length:          833
Score:                    37.00           Matches:         7
Percent Similarity:       88.9%           Conservative:    1
Best Local Similarity:    77.8%           Mismatches:      1
Query Match:              82.2%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AAC40379 (1-833)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||   ||||||
Db        368 CCTGACATCATCTTCCTTCTTGAGGTT 342

RESULT 58
ADT48666
ID   ADT48666 standard; cDNA; 852 BP.
XX
AC   ADT48666;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Bacterial polynucleotide #23417.
XX
KW   Recombinant DNA construct; transformed plant; improved plant property;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW   pathogen tolerance; pest tolerance; plant disease resistance;
KW   cell cycle pathway modification; plant growth regulator;
KW   homologous recombination; seed oil yield; protein yield; carbohydrate;
KW   nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW   bacterial polynucleotide; gene; ss.
XX
OS   Bacteria.
XX
PN   US2003233675-A1.
XX
PD   18-DEC-2003.
XX
PF   20-FEB-2003; 2003US-00369493.
XX
PR   21-FEB-2002; 2002US-0360039P.
XX
PA   (CAOY/) CAO Y.
PA   (HINK/) HINKLE G J.
PA   (SLAT/) SLATER S C.
PA   (CHEN/) CHEN X.
PA   (GOLD/) GOLDMAN B S.
XX
PI   Cao Y,  Hinkle GJ,  Slater SC,  Chen X,  Goldman BS;
XX
DR   WPI; 2004-061375/06.
XX
PT   New recombinant DNA construct comprising a promoter positioned to provide
PT   for expression of a polynucleotide encoding a polypeptide from a
PT   microbial source, useful for producing plants with improved properties.
XX
PS   Claim 1; SEQ ID NO 47104; 122pp; English.
XX
CC   The invention relates to a recombinant DNA construct comprising a
CC   promoter functional in a plant cell, where the promoter is positioned to
CC   provide for expression of a polynucleotide encoding a polypeptide from a
CC   microbial source. The invention also relates to a transformed plant
CC   comprising the recombinant DNA construct and a method of producing a
CC   transformed plant having an improved property. The plant is a crop plant
CC   such as maize or soybean. The method of producing a transformed plant
CC   having an improved property comprises transforming a plant with the
CC   recombinant DNA construct and growing the transformed plant, where the
CC   polynucleotide or polypeptide is useful for improving plant properties.
CC   The recombinant DNA construct is useful for producing plants with
CC   improved plant properties, e.g. improved cold, heat or drought tolerance,
CC   tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC   increased resistance to plant disease, better growth rate by modification
CC   of the cell cycle pathway with plant growth regulators, increased rate of
CC   homologous recombination, modified seed oil or protein yield and/or
CC   content, improved yield by modification of carbohydrate, nitrogen or
CC   phosphorus use and/or uptake, by modification of photosynthesis or by
CC   providing improved plant growth and development under at least one stress
CC   condition, improved lignin production or improved galactomannan
CC   production. This sequence represents a bacterial polynucleotide used in
CC   the scope of the invention. Note: The sequence data for this patent did
CC   not form part of the printed specification but was obtained in electronic
CC   format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 852 BP; 143 A; 293 C; 285 G; 131 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                161             Length:          852
Score:                    37.00           Matches:         7
Percent Similarity:       88.9%           Conservative:    1
Best Local Similarity:    77.8%           Mismatches:      1
Query Match:              82.2%           Indels:          0
DB:                       13              Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ADT48666 (1-852)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   ||||||||||||
Db        184 CCCGACGTGCTGCTTCTGCAAGAAGTC 210

RESULT 59
AAC40378/c
ID   AAC40378 standard; DNA; 936 BP.
XX
AC   AAC40378;
XX
```

DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 28095.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.

PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	179	Length:	936
Score:	37.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	82.2%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AAC40378 (1-936)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||:::|||||||||   ||||||
Db        481 CCTGACATCATCTTCCTTCTTGAGGTT 455
```

RESULT 60
ABL08315
ID ABL08315 standard; cDNA; 1350 BP.
XX
AC ABL08315;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 19427.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB64212.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 19427; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1350 BP; 328 A; 376 C; 380 G; 266 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	275	Length:	1350
Score:	37.00	Matches:	8
Percent Similarity:	88.9%	Conservative:	0
Best Local Similarity:	88.9%	Mismatches:	1
Query Match:	82.2%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ABL08315 (1-1350)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||   ||||||||||||
Db        487 CCCGATGTGATTTGTCTGCAAGAGGTA 513
```

RESULT 61
ADF31997
ID ADF31997 standard; DNA; 39949 BP.
XX
AC ADF31997;
XX
DT 26-FEB-2004 (first entry)
XX
DE Full length cosmid 2A7.
XX
KW Streptomyces strain; antibiotics;
KW antibiotic secondary metabolite production; cosmid 2A7; ds.
XX
OS Unidentified.
XX
PN WO2003099993-A2.
XX

PD 04-DEC-2003.
XX
PF 13-MAY-2003; 2003WO-US014972.
XX
PR 17-MAY-2002; 2002US-0381511P.
PR 09-AUG-2002; 2002GB-00018519.
PR 29-AUG-2002; 2002US-0407190P.
PR 24-JAN-2003; 2003GB-00001627.
XX
PA (AVET) AVENTIS PHARM INC.
XX
PI Martinez A;
XX
DR WPI; 2004-035113/03.
XX
PT New isolated DNA sequences, useful for increasing secondary metabolite
PT production in transformed Streptomyces strains, e.g. antibiotics, or for
PT accelerating sporulation in Streptomyces strains.
XX
PS Claim 1; SEQ ID NO 1; 79pp; English.
XX
CC The present sequence is an isolated DNA sequence conferring early and
CC increased secondary metabolite production in a Streptomyces strain. The
CC DNA sequences and transformed Streptomyces strains are useful for
CC increasing (heterologous) secondary metabolite production in Streptomyces
CC strains, e.g. antibiotics, or for accelerating sporulation in
CC Streptomyces strains. The transformed Streptomyces strains may be used as
CC an expression vehicle to assay libraries of soil DNA samples for members
CC having the ability to confer the phenotype of antibiotic secondary
CC metabolite production. The present sequence represents full length cosmid
CC 2A7 encoding sequence.
XX
SQ Sequence 39949 BP; 5170 A; 14694 C; 15000 G; 5082 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	1.43e+04	Length:	39949
Score:	37.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	82.2%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADF31997 (1-39949)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   ||||||||||||
Db      21667 CCCGACGTGCTGCTGCTGCAGGAGGTG 21693
```

RESULT 62
ADF31998
ID ADF31998 standard; DNA; 48200 BP.
XX
AC ADF31998;
XX
DT 26-FEB-2004 (first entry)
XX
DE Cosmid 2A7.
XX
KW Streptomyces strain; antibiotics;
KW antibiotic secondary metabolite production; cosmid 2A7; ds.
XX
OS Unidentified.
XX
PN WO2003099993-A2.
XX
PD 04-DEC-2003.
XX
PF 13-MAY-2003; 2003WO-US014972.
XX
PR 17-MAY-2002; 2002US-0381511P.
PR 09-AUG-2002; 2002GB-00018519.
PR 29-AUG-2002; 2002US-0407190P.
PR 24-JAN-2003; 2003GB-00001627.
XX
PA (AVET) AVENTIS PHARM INC.
XX
PI Martinez A;
XX
DR WPI; 2004-035113/03.
XX
PT New isolated DNA sequences, useful for increasing secondary metabolite
PT production in transformed Streptomyces strains, e.g. antibiotics, or for
PT accelerating sporulation in Streptomyces strains.
XX
PS Claim 2; SEQ ID NO 2; 79pp; English.
XX
CC The present sequence is an isolated DNA sequence conferring early and
CC increased secondary metabolite production in a Streptomyces strain. The
CC DNA sequences and transformed Streptomyces strains are useful for
CC increasing (heterologous) secondary metabolite production in Streptomyces
CC strains, e.g. antibiotics, or for accelerating sporulation in
CC Streptomyces strains. The transformed Streptomyces strains may be used as
CC an expression vehicle to assay libraries of soil DNA samples for members
CC having the ability to confer the phenotype of antibiotic secondary
CC metabolite production. The present sequence represents cosmid 2A7
CC encoding sequence.
XX
SQ Sequence 48200 BP; 6970 A; 17024 C; 17224 G; 6979 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	1.78e+04	Length:	48200
Score:	37.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	82.2%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADF31998 (1-48200)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   ||||||||||||
Db      21957 CCCGACGTGCTGCTGCTGCAGGAGGTG 21983
```

RESULT 63
ADQ97430
ID ADQ97430 standard; DNA; 208765 BP.
XX
AC ADQ97430;
XX
DT 07-OCT-2004 (first entry)
XX
DE Mouse cancer associated sequence MD08-047, SEQ ID 407.
XX
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX
OS Mus musculus.
XX
PN WO2004060304-A2.
XX
PD 22-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US041389.
XX
PR 27-DEC-2002; 2002US-00330773.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2004-543781/52.
XX
PT New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX
PS Claim 1; SEQ ID NO 407; 199pp; English.

```
XX
CC   The present invention relates to cancer associated sequences (ADQ97025-
CC   ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC   treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic formate directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 208765 BP; 53491 A; 42141 C; 46142 G; 66771 T; 0 U; 220 Other;

Alignment Scores:
Pred. No.:                9.85e+04        Length:         208765
Score:                    37.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    75.0%           Mismatches:     0
Query Match:              82.2%           Indels:         0
DB:                       12              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x ADQ97430 (1-208765)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||:::
Db     113412 CCAGACATAATCTTTCTGCAGCAG 113435

RESULT 64
ADQ45678
ID   ADQ45678 standard; DNA; 201 BP.
XX
AC   ADQ45678;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Myocardial infarction-associated SNP flanking transcript, SEQ ID 7341.
XX
KW   Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW   cardiant; gene therapy; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004058052-A2.
XX
PD   15-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US040978.
XX
PR   20-DEC-2002; 2002US-0434778P.
PR   10-MAR-2003; 2003US-0453135P.
PR   30-APR-2003; 2003US-0466412P.
PR   23-SEP-2003; 2003US-0504955P.
XX
PA   (APPL-) APPLERA CORP.
XX
PI   Cargill M,  Devlin JJ,  Iakoubova O;
XX
DR   WPI; 2004-533949/51.
XX
PT   Identifying an individual who has an altered risk for developing
PT   myocardial infarction by detecting a single nucleotide polymorphism in
PT   the individual's nucleic acids.
XX
PS   Claim 7; SEQ ID NO 7341; 145pp; English.
XX
CC   The invention relates to a novel method for identifying an individual who
CC   has an altered risk for developing myocardial infarction. The method
CC   comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC   the nucleotide sequences given in the specification in the individual's
CC   nucleic acids, where the presence of the SNP is correlated with an
CC   altered risk for myocardial infarction in the individual. The invention
CC   further comprises: an isolated nucleic acid molecule comprising at least
CC   8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC   the specification or its complement and encoding any one of the amino
CC   acid sequences given in the specification; an isolated polypeptide
CC   comprising an amino acid sequence given in the specification; an antibody
CC   that specifically binds to the polypeptide or its antigen-binding
CC   fragment; an amplified polynucleotide containing an SNP given in the
CC   specification and which is between about 16 and 1000 nucleotides in
CC   length; a kit for detecting an SNP in a nucleic acid, comprising the
CC   polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC   nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC   method for identifying an agent useful in treating or preventing
CC   myocardial infarction. The novel detection method has cardiant activity.
CC   The nucleic acids of the invention may be used in gene therapy. The
CC   method is useful in identifying an individual who has an increased or
CC   decreased risk for developing myocardial infarction and for preparing a
CC   composition for treating or preventing myocardial infarction. This
CC   polynucleotide sequence represents a transcript-based context sequence
CC   which flanks an SNP found in a human myocardial infarction-associated
CC   gene of the invention. Note: This sequence was not shown in the
CC   specification. The sequence has come from an electronic sequence listing
CC   downloaded from the WIPO website.
XX
SQ   Sequence 201 BP; 46 A; 58 C; 37 G; 59 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                50.1            Length:         201
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       13              Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x ADQ45678 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129

RESULT 65
ADQ45683
ID   ADQ45683 standard; DNA; 201 BP.
XX
AC   ADQ45683;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Myocardial infarction-associated SNP flanking transcript, SEQ ID 7346.
XX
KW   Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW   cardiant; gene therapy; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004058052-A2.
XX
PD   15-JUL-2004.
XX
PF   22-DEC-2003; 2003WO-US040978.
XX
PR   20-DEC-2002; 2002US-0434778P.
PR   10-MAR-2003; 2003US-0453135P.
PR   30-APR-2003; 2003US-0466412P.
PR   23-SEP-2003; 2003US-0504955P.
XX
PA   (APPL-) APPLERA CORP.
XX
PI   Cargill M,  Devlin JJ,  Iakoubova O;
XX
DR   WPI; 2004-533949/51.
XX
PT   Identifying an individual who has an altered risk for developing
PT   myocardial infarction by detecting a single nucleotide polymorphism in
PT   the individual's nucleic acids.
XX
PS   Claim 7; SEQ ID NO 7346; 145pp; English.
XX
CC   The invention relates to a novel method for identifying an individual who
```

CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
CC downloaded from the WIPO website.
XX
SQ Sequence 201 BP; 46 A; 51 C; 45 G; 58 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	50.1	Length:	201
Score:	36.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	13	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADQ45683 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::||||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24
```

RESULT 66
ADQ45674
ID ADQ45674 standard; DNA; 201 BP.
XX
AC ADQ45674;
XX
DT 18-NOV-2004 (first entry)
XX
DE Myocardial infarction-associated SNP flanking transcript, SEQ ID 7337.
XX
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
XX
OS Homo sapiens.
XX
PN WO2004058052-A2.
XX
PD 15-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US040978.
XX
PR 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Cargill M, Devlin JJ, Iakoubova O;
XX
DR WPI; 2004-533949/51.
XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 7; SEQ ID NO 7337; 145pp; English.
XX
CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
CC downloaded from the WIPO website.
XX
SQ Sequence 201 BP; 46 A; 51 C; 45 G; 58 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	50.1	Length:	201
Score:	36.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	13	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADQ45674 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::||||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24
```

RESULT 67
ADQ45687
ID ADQ45687 standard; DNA; 201 BP.
XX
AC ADQ45687;
XX
DT 18-NOV-2004 (first entry)
XX
DE Myocardial infarction-associated SNP flanking transcript, SEQ ID 7350.
XX
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
XX
OS Homo sapiens.
XX
PN WO2004058052-A2.
XX
PD 15-JUL-2004.
XX

PF 22-DEC-2003; 2003WO-US040978.
XX
PR 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Cargill M, Devlin JJ, Iakoubova O;
XX
DR WPI; 2004-533949/51.
XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 7; SEQ ID NO 7350; 145pp; English.
XX
CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
CC downloaded from the WIPO website.
XX
SQ Sequence 201 BP; 46 A; 58 C; 37 G; 59 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	50.1	Length:	201
Score:	36.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	13	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADQ45687 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||  :::|||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129
```

RESULT 68
ADQ06782
ID ADQ06782 standard; DNA; 245 BP.
XX
AC ADQ06782;
XX
DT 09-SEP-2004 (first entry)
XX
DE Soybean transcription factor seqid 3628.
XX
KW maize; soybean; Arabidopsis thaliana; transcription factor; homeobox;
KW HLH; leucine zipper; zinc finger; transformed plant; plant;
KW metabolic pathway; mutation detection; polymorphism; plant trait;
KW genome mapping; gene identification; gene analysis; plant breeding;
KW transgenic; ds.
XX
OS Glycine max.
XX
PN US2004123339-A1.
XX
PD 24-JUN-2004.
XX
PF 06-AUG-2001; 2001US-00922293.
XX
PR 24-NOV-1997; 97US-0067000P.
PR 09-DEC-1997; 97US-0069472P.
PR 13-JAN-1998; 98US-0071479P.
PR 10-FEB-1998; 98US-0074201P.
PR 10-FEB-1998; 98US-0074280P.
PR 10-FEB-1998; 98US-0074281P.
PR 10-FEB-1998; 98US-0074282P.
PR 12-FEB-1998; 98US-0074565P.
PR 12-FEB-1998; 98US-0074566P.
PR 12-FEB-1998; 98US-0074567P.
PR 19-FEB-1998; 98US-0074789P.
PR 19-FEB-1998; 98US-0075459P.
PR 19-FEB-1998; 98US-0075460P.
PR 19-FEB-1998; 98US-0075461P.
PR 19-FEB-1998; 98US-0075462P.
PR 19-FEB-1998; 98US-0075463P.
PR 19-FEB-1998; 98US-0075464P.
PR 09-MAR-1998; 98US-0077229P.
PR 09-MAR-1998; 98US-0077230P.
PR 09-MAR-1998; 98US-0077231P.
PR 18-MAR-1998; 98US-0078368P.
PR 07-APR-1998; 98US-0080844P.
PR 27-APR-1998; 98US-0083067P.
PR 29-APR-1998; 98US-0083386P.
PR 29-APR-1998; 98US-0083387P.
PR 29-APR-1998; 98US-0083388P.
PR 29-APR-1998; 98US-0083389P.
PR 13-MAY-1998; 98US-0085222P.
PR 13-MAY-1998; 98US-0085223P.
PR 13-MAY-1998; 98US-0085224P.
PR 15-MAY-1998; 98US-0085533P.
PR 21-MAY-1998; 98US-0086183P.
PR 21-MAY-1998; 98US-0086184P.
PR 21-MAY-1998; 98US-0086185P.
PR 21-MAY-1998; 98US-0086186P.
PR 21-MAY-1998; 98US-0086187P.
PR 21-MAY-1998; 98US-0086188P.
PR 16-JUN-1998; 98US-0089524P.
PR 18-JUN-1998; 98US-0089793P.
PR 18-JUN-1998; 98US-0089806P.
PR 18-JUN-1998; 98US-0089807P.
PR 18-JUN-1998; 98US-0089808P.
PR 18-JUN-1998; 98US-0089810P.
PR 18-JUN-1998; 98US-0089811P.
PR 18-JUN-1998; 98US-0089812P.
PR 18-JUN-1998; 98US-0089813P.
PR 18-JUN-1998; 98US-0089814P.
PR 30-JUN-1998; 98US-0091247P.
PR 30-JUN-1998; 98US-0091405P.
PR 09-SEP-1998; 98US-0099667P.
PR 09-SEP-1998; 98US-0099668P.
PR 09-SEP-1998; 98US-0099670P.
PR 09-SEP-1998; 98US-0099697P.
PR 16-SEP-1998; 98US-0100672P.
PR 16-SEP-1998; 98US-0100673P.
PR 16-SEP-1998; 98US-0100674P.
PR 17-SEP-1998; 98US-0100963P.
PR 21-SEP-1998; 98US-0101130P.

PR 21-SEP-1998; 98US-0101131P.
PR 21-SEP-1998; 98US-0101132P.
PR 22-SEP-1998; 98US-0101343P.
PR 22-SEP-1998; 98US-0101344P.
PR 22-SEP-1998; 98US-0101347P.
PR 22-SEP-1998; 98US-0101508P.
PR 25-SEP-1998; 98US-0101707P.
PR 13-OCT-1998; 98US-0104124P.
PR 13-OCT-1998; 98US-0104126P.
PR 13-OCT-1998; 98US-0104127P.
PR 13-OCT-1998; 98US-0104128P.
PR 18-NOV-1998; 98US-0108996P.
PR 19-NOV-1998; 98US-0109018P.
PR 24-NOV-1998; 98US-00199129.
PR 08-DEC-1998; 98US-00210297.
PR 11-DEC-1998; 98US-0111981P.
PR 22-DEC-1998; 98US-0113224P.
PR 12-JAN-1999; 99US-00229413.
XX
PA (CONN/) CONNER T W.
PA (HECK/) HECK G R.
PA (LIUJ/) LIU J.
XX
PI Conner TW, Heck GR, Liu J;
XX
DR WPI; 2004-468202/44.
XX
PT New substantially purified nucleic acid molecule that encodes a maize,
PT soybean or Arabidopsis thaliana transcription factor or its fragment,
PT useful for genome mapping, gene identification and analysis or plant
PT breeding.
XX
PS Claim 2; SEQ ID NO 3628; 140pp; English.
XX
CC The invention describes a substantially purified nucleic acid molecule
CC that encodes a maize, soybean or Arabidopsis thaliana transcription
CC factor or its fragment, where the maize or soybean transcription factor
CC is homeobox, HLH, leucine zipper, zinc finger, or other transcription
CC factor. Also described are: a substantially purified maize or soybean
CC transcription factor or its fragment defined above; a substantially
CC purified antibody or its fragment which is capable of specifically
CC binding to the transcription factor or its fragment above; a transformed
CC plant; a method for determining a level or pattern in a plant cell of a
CC transcription factor in a plant metabolic pathway; a method of
CC determining a mutation in a plant; a method of producing a plant
CC containing an overexpressed protein ore reduced levels of plant
CC transcription factor; a method of determining an association between a
CC polymorphism and a plant trait; and a method of isolating a nucleic acid
CC that encodes a plant transcription factor or its fragment. The nucleic
CC acid molecules, proteins and their fragments are useful for genome
CC mapping, gene identification and analysis, plant breeding, preparation of
CC constructs for use in plant gene expression and transgenic plants. The
CC nucleic acid molecules are useful as markers or probes. This sequence
CC represents a soybean transcription factor polynucleotide.
XX
SQ Sequence 245 BP; 82 A; 39 C; 50 G; 74 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	63.2	Length:	245
Score:	36.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x ADQ06782 (1-245)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||   ||||||||||||
Db         42 CCAGACATTATTTGTCTTCAGGAGGTT 68
```

RESULT 69
ADQ06781

ID ADQ06781 standard; DNA; 254 BP.
XX
AC ADQ06781;
XX
DT 09-SEP-2004 (first entry)
XX
DE Soybean transcription factor seqid 3627.
XX
KW maize; soybean; Arabidopsis thaliana; transcription factor; homeobox;
KW HLH; leucine zipper; zinc finger; transformed plant; plant;
KW metabolic pathway; mutation detection; polymorphism; plant trait;
KW genome mapping; gene identification; gene analysis; plant breeding;
KW transgenic; ds.
XX
OS Glycine max.
XX
PN US2004123339-A1.
XX
PD 24-JUN-2004.
XX
PF 06-AUG-2001; 2001US-00922293.
XX
PR 24-NOV-1997; 97US-0067000P.
PR 09-DEC-1997; 97US-0069472P.
PR 13-JAN-1998; 98US-0071479P.
PR 10-FEB-1998; 98US-0074201P.
PR 10-FEB-1998; 98US-0074280P.
PR 10-FEB-1998; 98US-0074281P.
PR 10-FEB-1998; 98US-0074282P.
PR 12-FEB-1998; 98US-0074565P.
PR 12-FEB-1998; 98US-0074566P.
PR 12-FEB-1998; 98US-0074567P.
PR 19-FEB-1998; 98US-0074789P.
PR 19-FEB-1998; 98US-0075459P.
PR 19-FEB-1998; 98US-0075460P.
PR 19-FEB-1998; 98US-0075461P.
PR 19-FEB-1998; 98US-0075462P.
PR 19-FEB-1998; 98US-0075463P.
PR 19-FEB-1998; 98US-0075464P.
PR 09-MAR-1998; 98US-0077229P.
PR 09-MAR-1998; 98US-0077230P.
PR 09-MAR-1998; 98US-0077231P.
PR 18-MAR-1998; 98US-0078368P.
PR 07-APR-1998; 98US-0080844P.
PR 27-APR-1998; 98US-0083067P.
PR 29-APR-1998; 98US-0083386P.
PR 29-APR-1998; 98US-0083387P.
PR 29-APR-1998; 98US-0083388P.
PR 29-APR-1998; 98US-0083389P.
PR 13-MAY-1998; 98US-0085222P.
PR 13-MAY-1998; 98US-0085223P.
PR 13-MAY-1998; 98US-0085224P.
PR 15-MAY-1998; 98US-0085533P.
PR 21-MAY-1998; 98US-0086183P.
PR 21-MAY-1998; 98US-0086184P.
PR 21-MAY-1998; 98US-0086185P.
PR 21-MAY-1998; 98US-0086186P.
PR 21-MAY-1998; 98US-0086187P.
PR 21-MAY-1998; 98US-0086188P.
PR 16-JUN-1998; 98US-0089524P.
PR 18-JUN-1998; 98US-0089793P.
PR 18-JUN-1998; 98US-0089806P.
PR 18-JUN-1998; 98US-0089807P.
PR 18-JUN-1998; 98US-0089808P.
PR 18-JUN-1998; 98US-0089810P.
PR 18-JUN-1998; 98US-0089811P.
PR 18-JUN-1998; 98US-0089812P.
PR 18-JUN-1998; 98US-0089813P.
PR 18-JUN-1998; 98US-0089814P.
PR 30-JUN-1998; 98US-0091247P.
PR 30-JUN-1998; 98US-0091405P.
PR 09-SEP-1998; 98US-0099667P.
PR 09-SEP-1998; 98US-0099668P.

```
PR   09-SEP-1998;   98US-0099670P.
PR   09-SEP-1998;   98US-0099697P.
PR   16-SEP-1998;   98US-0100672P.
PR   16-SEP-1998;   98US-0100673P.
PR   16-SEP-1998;   98US-0100674P.
PR   17-SEP-1998;   98US-0100963P.
PR   21-SEP-1998;   98US-0101130P.
PR   21-SEP-1998;   98US-0101131P.
PR   21-SEP-1998;   98US-0101132P.
PR   22-SEP-1998;   98US-0101343P.
PR   22-SEP-1998;   98US-0101344P.
PR   22-SEP-1998;   98US-0101347P.
PR   22-SEP-1998;   98US-0101508P.
PR   25-SEP-1998;   98US-0101707P.
PR   13-OCT-1998;   98US-0104124P.
PR   13-OCT-1998;   98US-0104126P.
PR   13-OCT-1998;   98US-0104127P.
PR   13-OCT-1998;   98US-0104128P.
PR   18-NOV-1998;   98US-0108996P.
PR   19-NOV-1998;   98US-0109018P.
PR   24-NOV-1998;   98US-00199129.
PR   08-DEC-1998;   98US-00210297.
PR   11-DEC-1998;   98US-0111981P.
PR   22-DEC-1998;   98US-0113224P.
PR   12-JAN-1999;   99US-00229413.
XX
PA   (CONN/) CONNER T W.
PA   (HECK/) HECK G R.
PA   (LIUJ/) LIU J.
XX
PI   Conner TW,  Heck GR,  Liu J;
XX
DR   WPI; 2004-468202/44.
XX
PT   New substantially purified nucleic acid molecule that encodes a maize,
PT   soybean or Arabidopsis thaliana transcription factor or its fragment,
PT   useful for genome mapping, gene identification and analysis or plant
PT   breeding.
XX
PS   Claim 2; SEQ ID NO 3627; 140pp; English.
XX
CC   The invention describes a substantially purified nucleic acid molecule
CC   that encodes a maize, soybean or Arabidopsis thaliana transcription
CC   factor or its fragment, where the maize or soybean transcription factor
CC   is homeobox, HLH, leucine zipper, zinc finger, or other transcription
CC   factor. Also described are: a substantially purified maize or soybean
CC   transcription factor or its fragment defined above; a substantially
CC   purified antibody or its fragment which is capable of specifically
CC   binding to the transcription factor or its fragment above; a transformed
CC   plant; a method for determining a level or pattern in a plant cell of a
CC   transcription factor in a plant metabolic pathway; a method of
CC   determining a mutation in a plant; a method of producing a plant
CC   containing an overexpressed protein ore reduced levels of plant
CC   transcription factor; a method of determining an association between a
CC   polymorphism and a plant trait; and a method of isolating a nucleic acid
CC   that encodes a plant transcription factor or its fragment. The nucleic
CC   acid molecules, proteins and their fragments are useful for genome
CC   mapping, gene identification and analysis, plant breeding, preparation of
CC   constructs for use in plant gene expression and transgenic plants. The
CC   nucleic acid molecules are useful as markers or probes. This sequence
CC   represents a soybean transcription factor polynucleotide.
XX
SQ   Sequence 254 BP; 86 A; 41 C; 52 G; 75 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                65.9          Length:          254
Score:                    36.00         Matches:         7
Percent Similarity:       88.9%         Conservative:    1
Best Local Similarity:    77.8%         Mismatches:      1
Query Match:              80.0%         Indels:          0
DB:                       12            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x ADQ06781 (1-254)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||:::|||   ||||||||||||
Db         33 CCAGACATTATTTGTCTTCAGGAGGTT 59

RESULT 70
AAI87230
ID   AAI87230 standard; cDNA; 346 BP.
XX
AC   AAI87230;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 7290.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US004927.
XX
PR   28-FEB-2000; 2000US-00515126.
PR   18-MAY-2000; 2000US-00577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   P-PSDB; AAO07299.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT   and treating e.g. leukemia, inflammation and immune disorders.
XX
PS   Claim 1; SEQ ID NO 7290; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 346 BP; 89 A; 79 C; 98 G; 80 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                94.5          Length:          346
Score:                    36.00         Matches:         7
Percent Similarity:       87.5%         Conservative:    0
Best Local Similarity:    87.5%         Mismatches:      1
Query Match:              80.0%         Indels:          0
DB:                       4             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x AAI87230 (1-346)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||   ||||||||||||
Db        208 CCAGATGTAACTTTTTTACAGGAA 231
```

RESULT 71
AAI87801/c
ID AAI87801 standard; cDNA; 396 BP.
XX
AC AAI87801;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 7861.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR P-PSDB; AAO07870.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 1; SEQ ID NO 7861; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 396 BP; 124 A; 61 C; 91 G; 120 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	111	Length:	396
Score:	36.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	80.0%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x AAI87801 (1-396)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||::::::
Db         58 CCAGATCTTCTGTTTTTGCAACAAATA 32
```

RESULT 72
ACA48082
ID ACA48082 standard; DNA; 717 BP.
XX
AC ACA48082;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #29739.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Streptococcus mutans.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU44212.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 35952; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 717 BP; 257 A; 76 C; 125 G; 259 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	221	Length:	717
Score:	36.00	Matches:	7

```
Percent Similarity:     88.9%          Conservative:  1
Best Local Similarity:  77.8%          Mismatches:    1
Query Match:            80.0%          Indels:        0
DB:                     8              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ACA48082 (1-717)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||   |||||||||:::||||||
Db        385 CCAGATAAAATATTTCTAGAAGAAGTC 411

RESULT 73
ACH96047/c
ID   ACH96047 standard; DNA; 768 BP.
XX
AC   ACH96047;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Klebsiella pneumoniae polynucleotide seqid 1842.
XX
KW   Recombinant expression vector; transcription regulatory element;
KW   Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX
OS   Klebsiella pneumoniae.
XX
PN   US6610836-B1.
XX
PD   26-AUG-2003.
XX
PF   27-JAN-2000; 2000US-00489039.
XX
PR   29-JAN-1999;   99US-0117747P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Breton GL,  Osborne M;
XX
DR   WPI; 2003-895346/82.
DR   P-PSDB; ABO62496.
XX
PT   New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT   preparing a vaccine composition against Klebsiella pneumoniae.
XX
PS   Disclosure; SEQ ID NO 1842; 932pp; English.
XX
CC   The invention describes a new isolated nucleic acid encoding a Klebsiella
CC   pneumoniae polypeptide. Also described are: a recombinant expression
CC   vector comprising the nucleic acid, operably linked to a transcription
CC   regulatory element; and a cell comprising the recombinant expression
CC   vector. The nucleic acid is useful for preparing a vaccine composition
CC   against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC   pneumoniae polypeptide of the invention
XX
SQ   Sequence 768 BP; 165 A; 212 C; 247 G; 144 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              239            Length:        768
Score:                  36.00          Matches:       6
Percent Similarity:     100.0%         Conservative:  2
Best Local Similarity:  75.0%          Mismatches:    0
Query Match:            80.0%          Indels:        0
DB:                     11             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ACH96047 (1-768)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||:::|||||||||:::
Db        546 CCAGACGTTCTCTTCCTGCAGCAG 523

RESULT 74
ACL68818
ID   ACL68818 standard; DNA; 864 BP.
XX
AC   ACL68818;
XX
DT   02-JUN-2005  (first entry)
XX
DE   M. xanthus gene sequence, seq id 5281.
XX
KW   Transgenic plant; DNA replication; gene regulation; gene expression;
KW   gene; ds.
XX
OS   Myxococcus xanthus.
XX
PN   US6833447-B1.
XX
PD   21-DEC-2004.
XX
PF   10-JUL-2001; 2001US-00902540.
XX
PR   10-JUL-2000; 2000US-0217883P.
XX
PA   (MONS ) MONSANTO TECHNOLOGY LLC.
XX
PI   Goldman BS,  Hinkle GJ,  Slater SC,  Wiegand RC;
XX
DR   WPI; 2005-028716/03.
XX
PT   New substantially purified Myxococcus xanthus nucleic acid molecule
PT   encoding a nitrite reductase, useful for determining gene expression,
PT   identifying mutations in a gene of interest, and for constructing
PT   mutations in a gene of interest.
XX
PS   Example 2; SEQ ID NO 5281; 25pp; English.
XX
CC   The invention relates to a substantially purified nucleic acid molecule
CC   encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC   recombinant DNA construct for expression of a nitrite reductase gene in a
CC   plant cell, and a plant cell comprising the recombinant DNA construct.
CC   The nucleic acid is useful for determining gene expression, identifying
CC   mutations in a gene of interest, and for constructing mutations in a gene
CC   of interest. Sequences given in records for SEQ IDs 1850-9691 represent a
CC   set of about 7842 genes or partial genes from the genome of the bacterium
CC   Mxyococcus xanthus. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from USPTO
XX
SQ   Sequence 864 BP; 106 A; 308 C; 310 G; 140 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              275            Length:        864
Score:                  36.00          Matches:       7
Percent Similarity:     88.9%          Conservative:  1
Best Local Similarity:  77.8%          Mismatches:    1
Query Match:            80.0%          Indels:        0
DB:                     14             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ACL68818 (1-864)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||:::   ||||||||||||
Db        199 CCGGACGTCGTCGCCTTGCAGGAGGTG 225

RESULT 75
ADV92147
ID   ADV92147 standard; DNA; 1040 BP.
XX
AC   ADV92147;
XX
DT   10-MAR-2005  (first entry)
XX
DE   Mouse CX3CR1 chemokine coding sequence.
XX
KW   CX3CR1 chemokine; G-protein coupled receptor; analgesic; pain;
KW   neurological disease; trauma; injury; neuropathy; drug screening; gene;
```

```
KW   ds.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             2. .1036
PT                   /*tag=  a
FT                   /product= "Mouse CR3CR1"
FT                   /organism
FT                   /note= "No start codon"
XX
PN   WO2004110416-A2.
XX
PD   23-DEC-2004.
XX
PF   04-JUN-2004; 2004WO-GB002391.
XX
PR   12-JUN-2003; 2003GB-00013632.
XX
PA   (NEUR-) NEUROTARGETS LTD.
XX
PI   Wynick D;
XX
DR   WPI; 2005-039974/04.
DR   P-PSDB; ADV92148.
XX
PT   Use of agonist of CX3C receptor in preparation of medicament for
PT   prevention or treatment of pain such as centrally mediated neuropathic
PT   pain, chronic, allodynia, hyperalgesic pain, or for promotion of nerve
PT   regeneration or repair.
XX
PS   Disclosure; Fig 1A-B; 36pp; English.
XX
CC   A claimed method of selecting candidate drugs for use in the prevention
CC   or treatment of pain comprises determining whether a compound is an
CC   agonist of a CXC3 receptor, and selecting such compounds. The receptor is
CC   preferably human CX3CR1. The agonist is useful in the treatment of pain,
CC   especially neuropathic pain, centrally mediated neuropathic pain, or
CC   chronic, allodynia, hyperalgesic pain, phantom pain or inflammatory pain,
CC   diabetic peripheral neuropathy, post-herpetic neuralgia, trigeminal
CC   neuralgia, post-stroke pain, multiple sclerosis-associated pain,
CC   neuropathy-associated pain such as in idiopathic or post-traumatic
CC   neuropathy and mononeuritis, HIV-associated neuropathic pain, cancer-
CC   associated neuropathic pain, carpal tunnel-associated neuropathic pain,
CC   spinal cord injury-associated pain, complex regional pain syndrome,
CC   fibromyalgia-associated neuropathic pain, lumbar and cervical pain,
CC   reflex sympathic dystrophy, phantom limb syndrome or peripheral nerve or
CC   spinal cord trauma, entrapment neuropathy, nerve transection including
CC   surgery, Lissauer tract section, limb amputation and stump pain,
CC   neuroma/tumor compression, arteriovenous malformation, vitamin B12
CC   deficiency, diabetic neuropathy, alcoholic neuropathy, pain caused by the
CC   side effects of anti-cancer and anti-AIDS therapies, HIV-associated
CC   neuropathic pain, pain associated with inflammation or infection of a
CC   tooth (toothache), visceral pain, pain caused by chemical burns, pain
CC   caused by local or systemic infection, pain caused by connective tissue
CC   disease such as rheumatoid arthritis, Wallenberg's syndrome, systemic
CC   lupus erythematosus, multiple sclerosis or polyarteritis nodosa, or pain
CC   is that associated with cancer, surgery, visceral damage, headache or
CC   trauma (all claimed). The agonist can also be used to promote nerve
CC   regeneration or repair. The agonist is fractalkine or its derivative. The
CC   present sequence is that of a murine CX3CR1 partial coding sequence.
XX
SQ   Sequence 1040 BP; 207 A; 313 C; 256 G; 264 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              341          Length:        1040
Score:                  36.00        Matches:       6
Percent Similarity:     87.5%        Conservative:  1
Best Local Similarity:  75.0%        Mismatches:    1
Query Match:            80.0%        Indels:        0
DB:                     14           Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x ADV92147 (1-1040)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        619 CCAGACGCTGTTTTCCTGCAAGAA 642


Search completed: March 10, 2006, 15:27:27
Job time : 97.8261 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2006, 05:53:56 ; Search time 877 Seconds
(without alignments)
9970.441 Million cell updates/sec

Title: US-10-757-745-3
Perfect score: 1312
Sequence: 1 agctattaatgattcgaatt.....tctcaacctttcaggacatc 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1312	100.0	1312	3	AAZ47119	Aaz47119 Mouse CD4
c	2	682.8	52.0	774	10	ADD34375	Add34375 Mouse mit
	3	668.2	50.9	1920	3	AAZ47118	Aaz47118 Human CD4
	4	668.2	50.9	1948	4	AAI58997	Aai58997 Human pol
	5	668.2	50.9	1948	5	ADQ99219	Adq99219 DNA encod
	6	668.2	50.9	1948	9	ADB48979	Adb48979 Novel hum
	7	668.2	50.9	1958	4	AAI60783	Aai60783 Human pol
	8	668.2	50.9	2499	2	AAX28153	Aax28153 Topoisome
	9	668.2	50.9	3152	9	ADA10970	Ada10970 Human cDN
	10	667	50.8	1296	3	AAC98160	Aac98160 Human col
	11	666.6	50.8	1936	10	ADD19013	Add19013 Human dis
	12	666.6	50.8	1936	13	ADP25361	Adp25361 PRO polyp
	13	666.6	50.8	1940	14	ADX06356	Adx06356 Cyclin-de
	14	665	50.7	1898	4	AAH15146	Aah15146 Human cDN
	15	600.4	45.8	602	10	ADD34376	Add34376 Mouse mit
	16	493.2	37.6	1088	5	AAS86254	Aas86254 DNA encod
c	17	490	37.3	1079	2	AAX84209	Aax84209 DNA encod
c	18	490	37.3	1079	3	AAC79438	Aac79438 cDNA sequ
c	19	490	37.3	1079	6	ABK28982	Abk28982 Human bre
	20	461.6	35.2	674	11	ADT95548	Adt95548 Colon can
	21	461.6	35.2	674	11	ADX42030	Adx42030 Human cDN
	22	436.4	33.3	644	11	ADT95551	Adt95551 Colon can
	23	436.4	33.3	644	11	ADX42033	Adx42033 Human cDN
	24	428.4	32.7	625	11	ADT95512	Adt95512 Colon can
	25	428.4	32.7	625	11	ADX41994	Adx41994 Human cDN
	26	428.4	32.7	633	11	ADT95565	Adt95565 Colon can
	27	428.4	32.7	633	11	ADX42047	Adx42047 Human cDN
	28	418	31.9	625	11	ADT95029	Adt95029 Colon can
	29	418	31.9	625	11	ADX41511	Adx41511 Human cDN
	30	385.8	29.4	553	11	ADT95464	Adt95464 Colon can
	31	385.8	29.4	553	11	ADX41946	Adx41946 Human cDN
	32	377.8	28.8	752	4	AAH08073	Aah08073 Human cDN
	33	290.2	22.1	401	11	ADT95844	Adt95844 Colon can
	34	290.2	22.1	401	11	ADX42326	Adx42326 Human cDN
	35	231.8	17.7	483	2	AAX40590	Aax40590 Human sec
	36	198.2	15.1	1227	5	AAS86255	Aas86255 DNA encod
c	37	140.8	10.7	391	11	ADT95307	Adt95307 Colon can
c	38	140.8	10.7	391	11	ADX41789	Adx41789 Human cDN
c	39	135.6	10.3	179	6	ABL36461	Abl36461 Human col
	40	134.4	10.2	176	6	ABK27684	Abk27684 Human col
c	41	134.4	10.2	176	6	ABV95959	Abv95959 Human pan
c	42	101.4	7.7	268	6	ABT07150	Abt07150 Human ova
c	43	101.4	7.7	268	8	ABX73028	Abx73028 Human ova
	44	75.2	5.7	258	2	AAX41008	Aax41008 Human sec
	45	74.6	5.7	444	4	AAI15311	Aai15311 Probe #52
	46	74.6	5.7	691	4	AAI24484	Aai24484 Probe #14
	47	67.4	5.1	705	6	ABT09027	Abt09027 Phase-1 R
	48	67.4	5.1	705	12	ADG45549	Adg45549 Liver inf
	49	67.4	5.1	705	12	ADH22854	Adh22854 Partial D
	50	64.8	4.9	631	6	ABT09296	Abt09296 Phase-1 R
	51	64.8	4.9	631	12	ADG45660	Adg45660 Liver inf
	52	64.8	4.9	631	12	ADH22946	Adh22946 Partial D
	53	64.8	4.9	631	13	ADR91260	Adr91260 Spleen ne
	54	63.8	4.9	690	6	ABT09191	Abt09191 Phase-1 R
	55	63.8	4.9	690	10	ADG30969	Adg30969 Liver tox
	56	63.8	4.9	690	12	ADG45586	Adg45586 Liver inf
	57	63.8	4.9	690	12	ADH22884	Adh22884 Partial D
	58	63.6	4.8	969	4	AAK52001	Aak52001 Human pol
	59	63.2	4.8	696	6	ABT09200	Abt09200 Phase-1 R
	60	63.2	4.8	696	10	ADG30974	Adg30974 Liver tox
	61	63.2	4.8	696	12	ADG45593	Adg45593 Liver inf
	62	63.2	4.8	696	13	ADR91217	Adr91217 Spleen ne
	63	63.2	4.8	704	6	ABT09004	Abt09004 Phase-1 R
	64	63.2	4.8	704	10	ADG30925	Adg30925 Liver tox
	65	63.2	4.8	704	13	ADR91162	Adr91162 Spleen ne
	66	62.8	4.8	722	6	ABT09153	Abt09153 Phase-1 R
	67	62.8	4.8	722	12	ADG45729	Adg45729 Liver inf
	68	62.2	4.7	630	6	ABT09311	Abt09311 Phase-1 R
	69	62.2	4.7	630	12	ADG45672	Adg45672 Liver inf
	70	62.2	4.7	630	13	ADR91268	Adr91268 Spleen ne
	71	62.2	4.7	1549	3	AAA38745	Aaa38745 Murine Hs
	72	61.6	4.7	701	6	ABT08958	Abt08958 Phase-1 R
	73	61.6	4.7	701	10	ADG30880	Adg30880 Liver tox
	74	61.6	4.7	701	12	ADG45448	Adg45448 Liver inf
	75	61.6	4.7	701	12	ADH22750	Adh22750 Partial D

ALIGNMENTS

RESULT 1
AAZ47119
ID AAZ47119 standard; cDNA; 1312 BP.
XX
AC AAZ47119;
XX
DT 15-MAR-2000 (first entry)
XX
DE Mouse CD40 receptor associated protein gene.
XX
KW Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;

```
KW   mouse; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW   tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW   TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW   diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW   arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW   graft versus host disease; autoimmune disease; ds.
XX
OS   Mus musculus.
XX
PN   WO9955859-A2.
XX
PD   04-NOV-1999.
XX
PF   28-APR-1999;   99WO-EP003025.
XX
PR   29-APR-1998;   98EP-00201392.
XX
PA   (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI   Pype SMC,  Remacle JEFJG,  Huylebroeck DFE;
XX
DR   WPI; 2000-062029/05.
DR   P-PSDB; AAY56020.
XX
PT   Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT   diseases and for improvement of anti-tumor treatments.
XX
PS   Claim 10; Page 41-43; 48pp; English.
XX
CC   This sequence represents the gene encoding mouse CD40 receptor associated
CC   protein (CRAP). CRAP is a functional protein capable of interacting with
CC   the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC   necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC   I, where the protein has no homology to TNF receptor associated factor
CC   (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC   the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC   related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC   diseases, and for the improvement of anti-tumour diseases. Diseases which
CC   may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC   systemic lupus erythematosus, graft rejection, graft versus host disease,
CC   allergy, and autoimmune disease. The proteins can be used to sensitize
CC   tumour cells to anti-tumour treatments and to screen for compounds which
CC   interfere with the interaction of the proteins with other protein
CC   components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ   Sequence 1312 BP; 359 A; 279 C; 352 G; 322 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 1312;  DB 3;  Length 1312;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1312;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60

Qy         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120

Qy        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180

Qy        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240

Qy        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300

Qy        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360

Qy        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420

Qy        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480

Qy        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540

Qy        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600

Qy        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660

Qy        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720

Qy        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780

Qy        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840

Qy        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900

Qy        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960

Qy        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020

Qy       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080

Qy       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140

Qy       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200

Qy       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260

Qy       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312

RESULT 2
ADD34375/c
ID   ADD34375 standard; DNA; 774 BP.
XX
AC   ADD34375;
XX
```

DT 15-JAN-2004 (first entry)
XX
DE Mouse mitochondrial DNA sequence SEQ ID NO:2153.
XX
KW ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW mitochondrial disease; oxidative phosphorylation dysfunction;
KW oxidative stress; apoptosis; aging.
XX
OS Mus musculus.
XX
PN WO2003020220-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027886.
XX
PR 30-AUG-2001; 2001US-0316323P.
PR 31-AUG-2001; 2001CA-02356540.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Wallace DC, Levy S, Kerstann K, Procaccio V;
XX
DR WPI; 2003-300821/29.
XX
PT Array containing probes for genes involved in mitochondrial biology,
PT useful for determining mitochondrial biology gene expression profiles for
PT use in diagnosing pathologies and identifying biochemical pathways.
XX
PS Claim 2; SEQ ID NO 2153; 201pp; English.
XX
CC The invention relates to a novel array comprising at least two isolated
CC nucleotide molecules, each molecule having a sequence capable of uniquely
CC hybridising to a nucleic acid molecule which is an expression product of
CC a gene involved in mitochondrial biology. The array comprises two or more
CC isolated nucleic acid molecules or spots, each molecule having a sequence
CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC of the invention is useful for determining an expression profile of a
CC mouse or human sample containing nucleic acid, by contacting the array
CC with the sample under conditions allowing selective hybridisation, and
CC measuring hybridisation of nucleic acid in the sample to the array to
CC produce an expression profile. The array is also useful for determining
CC an expression profile of a first labelled sample containing nucleic acid
CC relative to a second, differently labelled sample containing nucleic
CC acid. The second sample is a reference or a standard. An array is useful
CC for determining an expression profile diagnostic of an energy-metabolism-
CC related physiological condition. An array of the invention is useful for
CC determining mitochondrial biology gene expression profiles of organisms,
CC such as human, mice and closely related species, tissue and organs of
CC such organisms, which are useful for determining expression profiles
CC diagnostic of energy metabolism-related physiological conditions,
CC diagnosing such physiological conditions, identifying biochemical
CC pathways, genes, and mutations involved in such physiological conditions,
CC identifying therapeutic agents useful for preventing and/or treating such
CC physiological conditions, evaluating and/or monitoring the efficacy of
CC such therapies, and creating and identifying animal models of human
CC energy metabolism-related physiological conditions. An array is also
CC useful for defining expression signatures or profiles for mitochondrial
CC diseases, as well as distinguishing clinical disorders that result from
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC apoptosis and aging. An array of the invention contains probes of genes
CC not previously recognised to participate in mitochondrial biology. The
CC sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC clones used to make the probes of the invention. Some sequences are not
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC 1906, 2408 and 2643.
XX
SQ Sequence 774 BP; 218 A; 169 C; 135 G; 252 T; 0 U; 0 Other;

Query Match 52.0%; Score 682.8; DB 10; Length 774;
Best Local Similarity 99.4%; Pred. No. 9.4e-179;
Matches 706; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```
Qy        593 GTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGT 652
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        768 GTATTTCTACAGGAAG-TATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAG- 711

Qy        653 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        710 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 651

Qy        713 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAAC 772
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        650 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAGC 591

Qy        773 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        590 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 531

Qy        833 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        530 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 471

Qy        893 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 411

Qy        953 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 351

Qy       1013 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 291

Qy       1073 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTT 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTT 231

Qy       1133 ATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGT 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 ATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGT 171

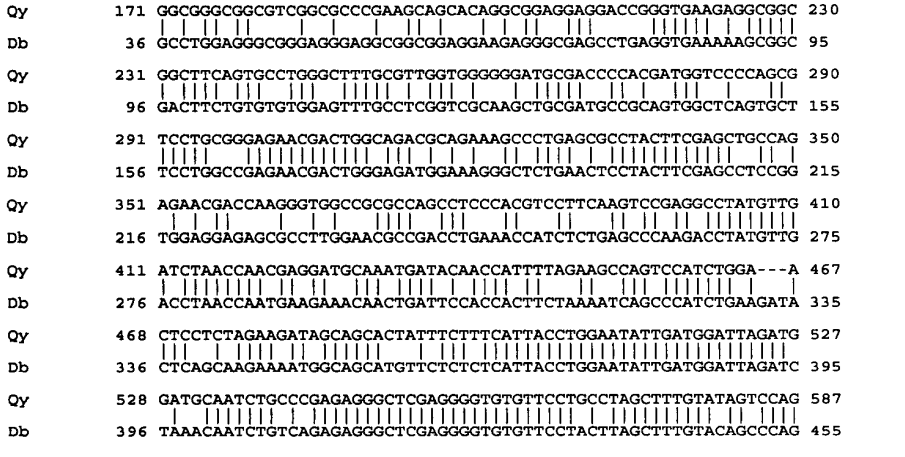
Qy       1193 GATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAG 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 GATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAG 111

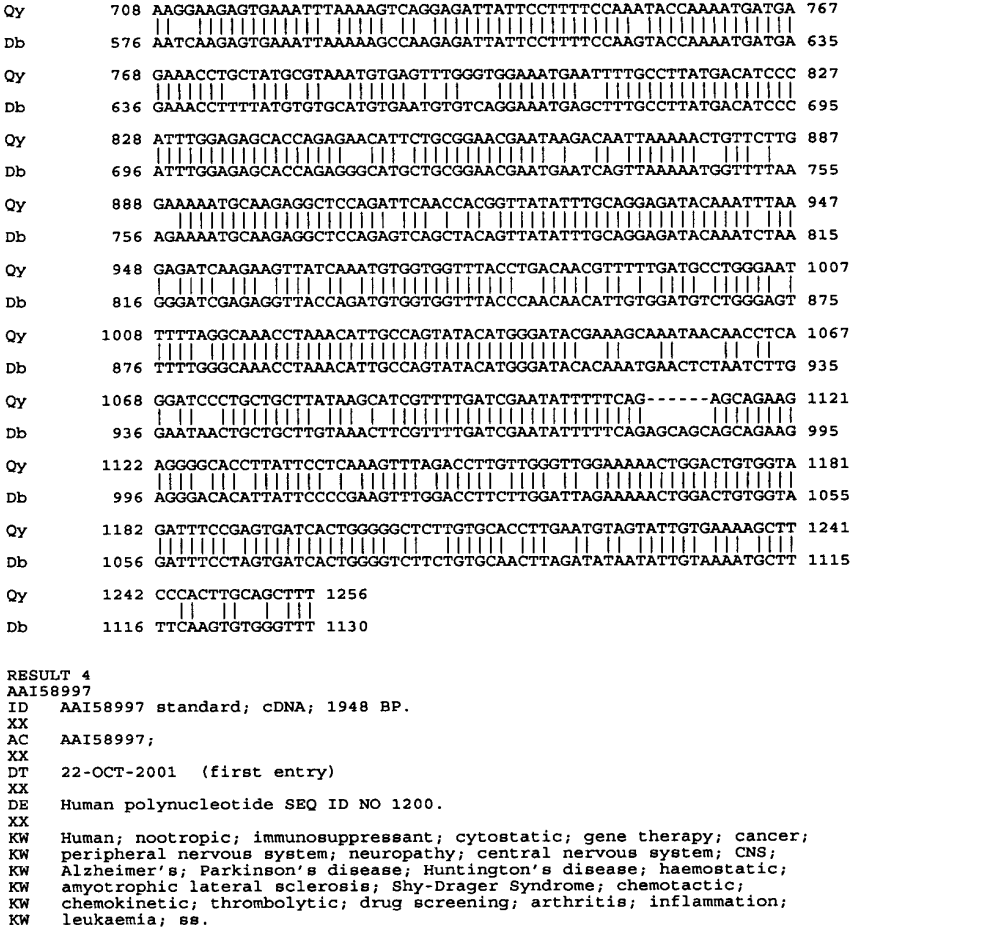
Qy       1253 CTTTACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTT 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        110 CTTTACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACTTT 61
```

RESULT 3
AAZ47118
ID AAZ47118 standard; cDNA; 1920 BP.
XX
AC AAZ47118;
XX
DT 15-MAR-2000 (first entry)
XX
DE Human CD40 receptor associated protein gene.
XX
KW Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW human; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW graft versus host disease; autoimmune disease; ds.
XX
OS Homo sapiens.
XX
PN WO9955859-A2.
XX
PD 04-NOV-1999.

XX
PF 28-APR-1999; 99WO-EP003025.
XX
PR 29-APR-1998; 98EP-00201392.
XX
PA (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Pype SMC, Remacle JEFJG, Huylebroeck DFE;
XX
DR WPI; 2000-062029/05.
DR P-PSDB; AAY56019.
XX
PT Novel proteins used to treat inflammatory diseases, NF-kappaB related
PT diseases and for improvement of anti-tumor treatments.
XX
PS Claim 9; Page 37-39; 48pp; English.
XX
CC This sequence represents the gene encoding human CD40 receptor associated
CC protein (CRAP). CRAP is a functional protein capable of interacting with
CC the cytoplasmic domain of CD40 and/or other receptors of the tumour
CC necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor
CC I, where the protein has no homology to TNF receptor associated factor
CC (TRAF)-proteins. The CD40 binding proteins can be used as modulators of
CC the CD40 signalling pathway, especially to diagnose and treat TRAF-
CC related, CD40-related, NF-kappaB related and/or Jun (kinase)-related
CC diseases, and for the improvement of anti-tumour diseases. Diseases which
CC may be treated include atherosclerosis, arthritis, multiple sclerosis,
CC systemic lupus erythematosus, graft rejection, graft versus host disease,
CC allergy, and autoimmune disease. The proteins can be used to sensitize
CC tumour cells to anti-tumour treatments and to screen for compounds which
CC interfere with the interaction of the proteins with other protein
CC components of the TRAF, CD40 or NF-kappaB related pathway
XX
SQ Sequence 1920 BP; 599 A; 327 C; 435 G; 557 T; 0 U; 2 Other;

Query Match 50.9%; Score 668.2; DB 3; Length 1920;
Best Local Similarity 77.0%; Pred. No. 1.7e-174;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;





RESULT 4
AAI58997
ID AAI58997 standard; cDNA; 1948 BP.
XX
AC AAI58997;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1200.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM39841.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 1200; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Query Match 50.9%; Score 668.2; DB 4; Length 1948;
Best Local Similarity 75.6%; Pred. No. 1.7e-174;
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;

```
Qy       131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
Db        15 GCCAAGAATTCGGCACGAGGGAAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAG 74

Qy       191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
Db        75 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 134

Qy       251 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 310
Db       135 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 194

Qy       311 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 370
Db       195 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 254

Qy       371 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 430
Db       255 CGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 314

Qy       431 AATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGC 487
Db       315 ACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGC 374

Qy       488 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 547
Db       375 ATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGG 434

Qy       548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
Db       435 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 494

Qy       608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
Db       495 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 554

Qy       668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
Db       555 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 614

Qy       728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
Db       615 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 674

Qy       788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
Db       675 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 734

Qy       848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
Db       735 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 794

Qy       908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
Db       795 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 854

Qy       968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
Db       855 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 914

Qy      1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
Db       915 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 974

Qy      1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
Db       975 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 1034

Qy      1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
Db      1035 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 1094

Qy      1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
Db      1095 GGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 1149
```

RESULT 5
ADQ99219
ID ADQ99219 standard; cDNA; 1948 BP.
XX
AC ADQ99219;
XX
DT 23-SEP-2004 (first entry)
XX
DE DNA encoding human GPCR-like protein seqid 889.
XX
KW ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
KW antidiabetic; GPCR-like protein; ophthalmic disorder;
KW neurological disorder; immunological disorder; nephritic disorder;
KW hormonal dysfunction; cancer; atherosclerosis; diabetes;
KW molecular weight marker; food supplement; human; ss.
XX
OS Homo sapiens.
XX
PN US6569662-B1.
XX
PD 27-MAY-2003.
XX
PF 19-JUL-2000; 2000US-00620312.
XX
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
XX
PA (HYSE-) HYSEQ INC.

```
XX
PI   Tang YT,  Zhou P,  Drmanac RT;
XX
DR   WPI; 2001-442255/47.
XX
PT   New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT   useful for treating diseases of ophthalmic, neurological, immunological
PT   and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT   and diabetes.
XX
PS   Example 2; SEQ ID NO 889; 92pp; English.
XX
CC   The invention describes an isolated polynucleotide (I) comprising a fully
CC   defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC   1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC   given in the specification, its translated or protein coding portion, its
CC   extracellular portion or its active domain. The GPCR-like polypeptides
CC   and polynucleotides are useful for the treatment of diseases of
CC   ophthalmic, neurological, immunological and nephritic systems. They may
CC   also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC   diabetes. The antibodies are useful for detecting or quantitating the
CC   polypeptide in tissue. The polypeptides can also be used as molecular
CC   weight markers and as a food supplement. This sequence represents a human
CC   polynucleotide of the invention.
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

  Query Match             50.9%;  Score 668.2;  DB 5;  Length 1948;
  Best Local Similarity   75.6%;  Pred. No. 1.7e-174;
  Matches  858;  Conservative    0;  Mismatches  268;  Indels    9;  Gaps    2;

Qy        131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
              | ||  | ||| |      | | ||   |    |  |    | | || ||    |  |
Db         15 GCCAAGAATTCGGCACGAGGGAAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAG 74

Qy        191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
              |  || |   | | ||  |||     ||||||| | |||||| |||| |||  |||  |||
Db         75 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 134

Qy        251 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 310
              || | ||| |    | |||||  || |  |||  |    || |||||   ||||||||||||
Db        135 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 194

Qy        311 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 370
               ||| | | |  || |||| | ||||||||||||  || | | | ||    |  | |
Db        195 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 254

Qy        371 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 430
              ||||  |||    ||  ||  || || || |||||||||| |||||||| || ||  ||
Db        255 CGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 314

Qy        431 AATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGC 487
              | |||| | |||| ||    ||  ||| |||||||  |   ||||  | |||| || |||||
Db        315 ACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGC 374

Qy        488 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 547
              |    | ||| ||||||||||||||||||||||||||||||  |  ||||||| | ||||||
Db        375 ATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGG 434

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||  | |||||||||| || |||||||| ||||||||||||||
Db        435 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 494

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              |||||  |||||||| | |  ||||||||||||||||| ||    | |||     |||||||||
Db        495 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 554

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        555 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 614

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| |||||||||||||||||||  |||| ||  ||
```

```
Db        615 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 674

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||| ||||||||||||||||||||||||||||||||||||||
Db        675 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 734

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||  |  || |||||||  |||  |    ||||||||||||||||||
Db        735 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 794

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || |||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        795 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 854

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||
Db        855 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 914

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||||  ||   ||     || ||  | ||  |||||||| |||
Db        915 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 974

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
              | |||||||||||||||||||||||      |||||||||||| ||| ||||||| | |
Db        975 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 1034

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
              ||||| |||||| |||| || ||||||||||||||||||||||||| ||||||||||||
Db       1035 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 1094

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              || ||  |||||| |||  || || |||||| ||| ||||  ||  ||  | |||
Db       1095 GGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 1149
```

```
RESULT 6
ADB48979
ID   ADB48979 standard; cDNA; 1948 BP.
XX
AC   ADB48979;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Novel human cDNA SEQ ID NO 889.
XX
KW   ss; cancer; neurodegenerative disease; human.
XX
OS   Homo sapiens.
XX
PN   US2003104529-A1.
XX
PD   05-JUN-2003.
XX
PF   04-JAN-2002; 2002US-00037270.
XX
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   19-JUL-2000; 2000US-00620312.
XX
PA   (ZHOU/) ZHOU P.
PA   (TANG/) TANG Y T.
PA   (LIUC/) LIU C.
PA   (ASUN/) ASUNDI V.
PA   (DRMA/) DRMANAC R T.
XX
PI   Zhou P,  Tang YT,  Liu C,  Asundi V,  Drmanac RT;
XX
DR   WPI; 2003-678194/64.
XX
PT   New polynucleotide, useful for treating diseases e.g., cancer or
PT   neurodegenerative diseases.
XX
PS   Claim 1; SEQ ID NO 889; 99pp; English.
```

```
XX
CC   The invention relates to a polynucleotide comprising a sequence given in
CC   the specification, or its mature protein-coding portion, or its
CC   complement. The polynucleotide is useful for treating diseases e.g.,
CC   cancer or neurodegenerative diseases and many others listed in the
CC   specification. The present sequence represents a novel human cDNA. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=20030104529.
XX
SQ   Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

  Query Match           50.9%;  Score 668.2;  DB 9;  Length 1948;
  Best Local Similarity  75.6%;  Pred. No. 1.7e-174;
  Matches  858;  Conservative    0;  Mismatches  268;  Indels    9;  Gaps    2;

Qy       131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
Db        15 GCCAAGAATTCGGCACGAGGGAAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAG 74

Qy       191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
Db        75 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 134

Qy       251 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 310
Db       135 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 194

Qy       311 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 370
Db       195 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 254

Qy       371 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 430
Db       255 CGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 314

Qy       431 AATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGC 487
Db       315 ACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGC 374

Qy       488 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 547
Db       375 ATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGG 434

Qy       548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
Db       435 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 494

Qy       608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
Db       495 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 554

Qy       668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
Db       555 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 614

Qy       728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
Db       615 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 674

Qy       788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
Db       675 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 734

Qy       848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
Db       735 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 794

Qy       908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
Db       795 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 854

Qy       968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
Db       855 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 914

Qy      1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
Db       915 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 974

Qy      1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
Db       975 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 1034

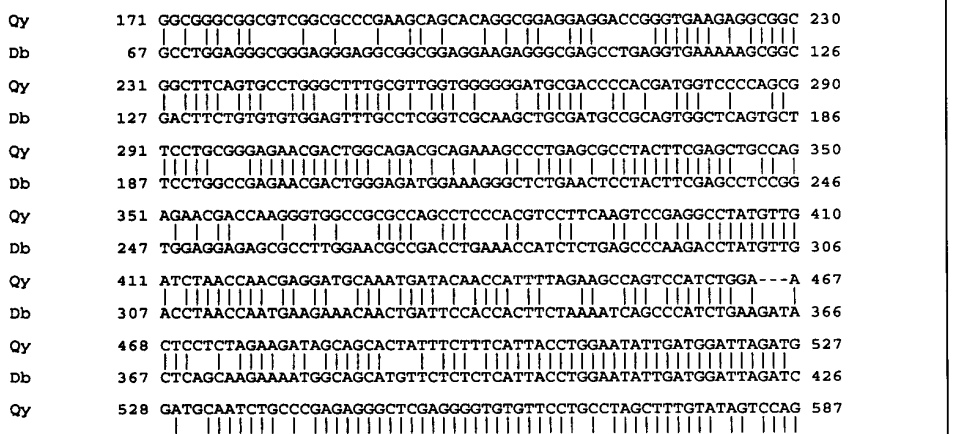
Qy      1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
Db       1035 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 1094

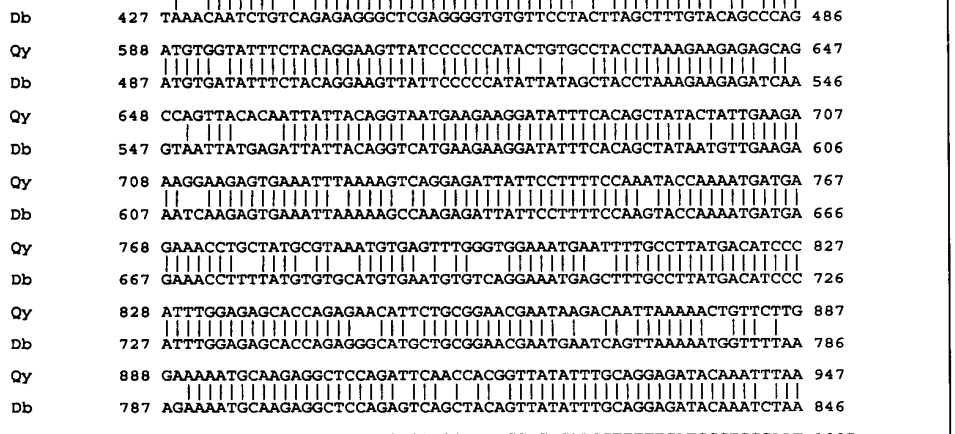
Qy      1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
Db       1095 GGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 1149

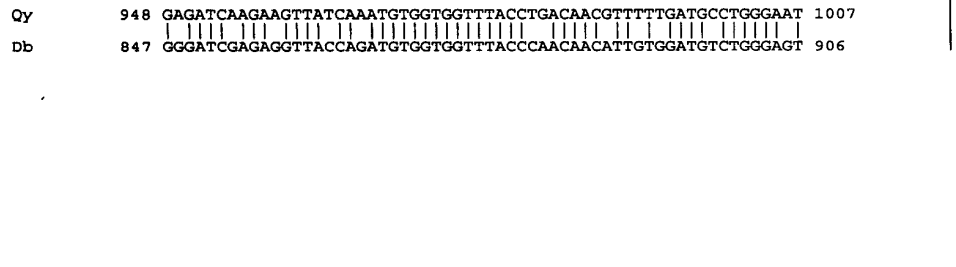

RESULT 7
AAI60783
ID   AAI60783 standard; cDNA; 1958 BP.
XX
AC   AAI60783;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4772.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J,  Zhao QA;
PI   Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41627.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders such
PT   as central nervous system injuries.
XX
PS   Claim 1; SEQ ID NO 4772; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC   encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
```

CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1958 BP; 613 A; 339 C; 442 G; 564 T; 0 U; 0 Other;

Query Match 50.9%; Score 668.2; DB 4; Length 1958;
Best Local Similarity 77.0%; Pred. No. 1.7e-174;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;



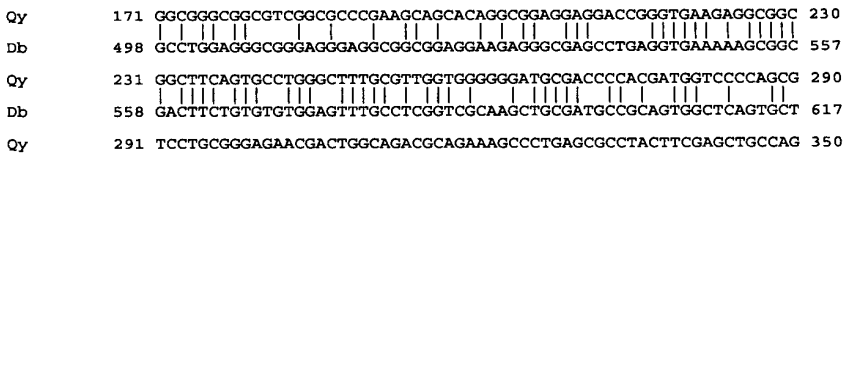


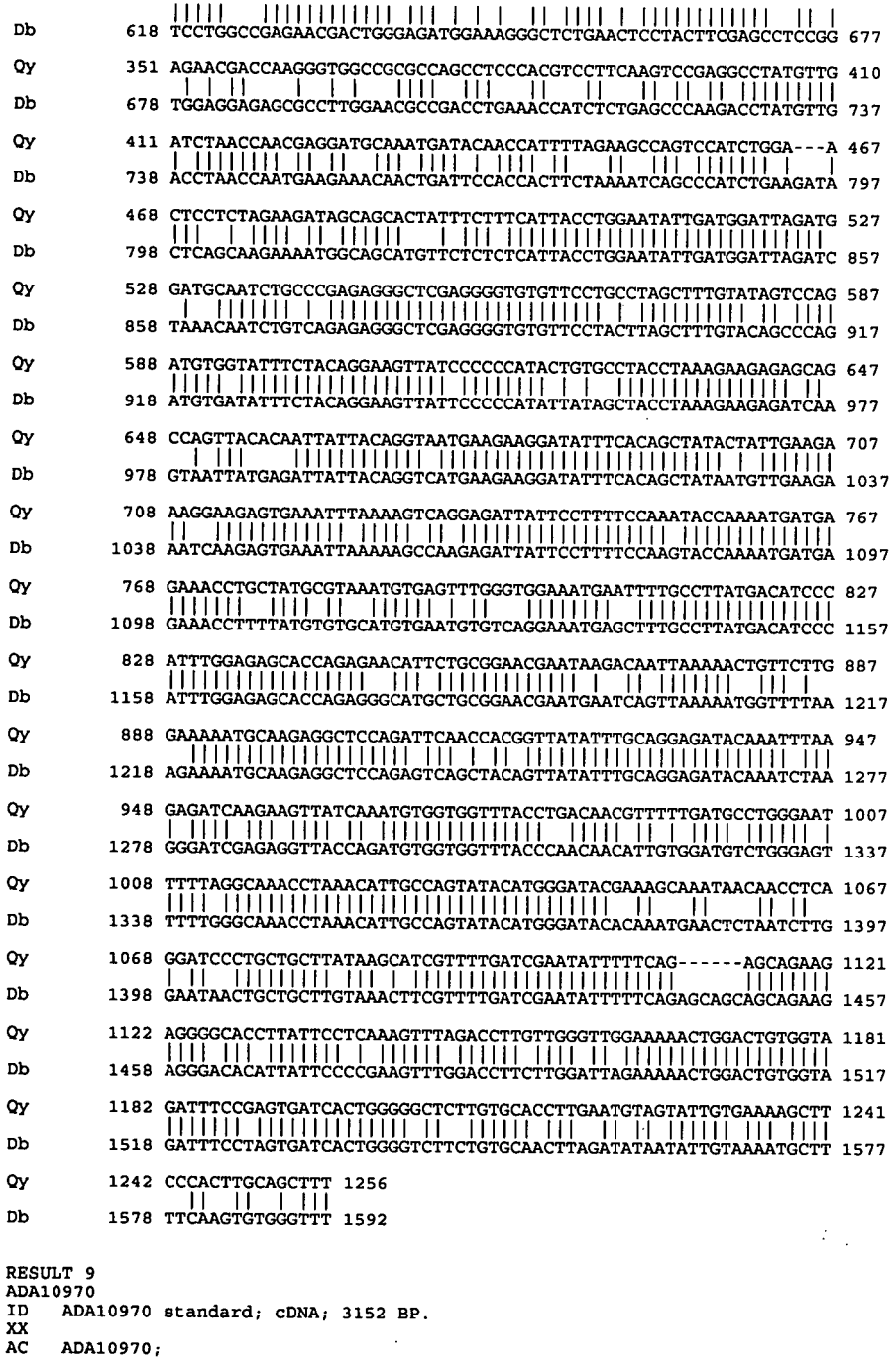




RESULT 8
AAX28153
ID AAX28153 standard; DNA; 2499 BP.
XX
AC AAX28153;
XX
DT 16-JUN-1999 (first entry)
XX
DE Topoisomerase II binding protein 1 coding sequence.
XX
KW Topoisomerase II binding protein; TopBP; anticancer agent; ds.
XX
OS Homo sapiens.
XX
PN JP11075856-A.
XX
PD 23-MAR-1999.
XX
PF 17-SEP-1997; 97JP-00251544.
XX
PR 17-SEP-1997; 97JP-00251544.
XX
PA (TSUR/) TSURUO T.
PA (CHUS) CHUGAI PHARM CO LTD.
XX
DR WPI; 1999-257704/22.
DR P-PSDB; AAY03182.
XX
PT New Topoisomerase II- binding protein - useful as an anticancer agent.
XX
PS Disclosure; Page 18-19; 28pp; Japanese.
XX
CC This sequence encodes the topoisomerase II binding protein (TopBP) of the
CC invention. The TopBP protein is useful as an anticancer agent. TopBP can
CC be used as the target molecule for anticancer agent
XX
SQ Sequence 2499 BP; 720 A; 505 C; 587 G; 687 T; 0 U; 0 Other;

Query Match 50.9%; Score 668.2; DB 2; Length 2499;
Best Local Similarity 77.0%; Pred. No. 2e-174;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;





RESULT 9
ADA10970
ID ADA10970 standard; cDNA; 3152 BP.
XX
AC ADA10970;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human cDNA differentially expressed in colon cancer #60.
XX
KW ss; differential expression; colon cancer; cancer; human.
XX
OS Homo sapiens.
XX
PN US2002160382-A1.
XX
PD 31-OCT-2002.
XX
PF 11-OCT-2001; 2001US-00981353.
XX
PR 11-OCT-2000; 2000US-0239841P.
XX
PA (LASE/) LASEK A W.
PA (JONE/) JONES D A.
XX
PI Lasek AW, Jones DA;
XX
DR WPI; 2003-265756/26.
XX
PT New combination comprising cDNAs that are differentially expressed in
PT colon disorder, useful for diagnosing, treating, staging or monitoring
PT treatment for colon cancers.
XX
PS Claim 1; SEQ ID NO 88; 231pp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in colon disorder. The methods and compositions
CC of the present invention are useful for diagnosing, treating, staging or
CC monitoring treatment for colon cancer. They are also useful in high
CC throughput methods for using cDNAs to detect differential expression of
CC nucleic acids in a sample, screening molecules or compounds to identify a
CC ligand which specifically binds a cDNA and using a protein to screen
CC molecules or compounds to identify at least one ligand which specifically
CC binds the protein. The present sequence represents a human cDNA
CC differentially expressed in colon cancer.
XX
SQ Sequence 3152 BP; 875 A; 643 C; 707 G; 926 T; 0 U; 1 Other;

Query Match 50.9%; Score 668.2; DB 9; Length 3152;
Best Local Similarity 77.0%; Pred. No. 2.2e-174;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;



```
Db       1628 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 1687

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db       1688 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 1747

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db       1748 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 1807

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db       1808 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 1867

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db       1868 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 1927

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db       1928 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 1987

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db       1988 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 2047

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db       2048 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 2107

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db       2108 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 2167

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db       2168 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 2227

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       2228 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 2287

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       2288 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 2347

Qy       1242 CCCACTTGCAGCTTT 1256

Db       2348 TTCAAGTGTGGGTTT 2362
```

RESULT 10
AAC98160
ID AAC98160 standard; cDNA; 1296 BP.
XX
AC AAC98160;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human colon cancer antigen nucleotide sequence SEQ ID NO:170.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200055351-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005883.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587534/55.
DR P-PSDB; AAB53403.
XX
PT Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer.
XX
PS Claim 1; Page 597; 2104pp; English.
XX
CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins may
CC also be used to prevent diseases such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, wounds, renal disorders, infectious diseases, and
CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 1296 BP; 376 A; 250 C; 326 G; 333 T; 0 U; 11 Other;

Query Match 50.8%; Score 667; DB 3; Length 1296;
Best Local Similarity 76.7%; Pred. No. 3e-174;
Matches 840; Conservative 3; Mismatches 243; Indels 9; Gaps 2;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         40 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 99

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        100 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 159

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        160 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 219

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        220 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 279

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        280 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 339

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        340 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 399

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        400 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 459

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        460 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 519
```

```
Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db        520 GTAATTATGAGATTATTACAGGTCATGAAGAAGGAKATTTCACAGCTATAATGTTGAAGA 579

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db        580 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 639

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db        640 GAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 699

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db        700 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 759

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db        760 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 819

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db        820 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 879

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db        880 TTTTGGGYAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 939

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db        940 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 999

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db       1000 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1059

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db       1060 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1119

Qy       1242 CCCACTTGCAGCTTT 1256
Db       1120 TTCAAGTGTGGGTTT 1134
```

RESULT 11
ADD19013
ID ADD19013 standard; DNA; 1936 BP.
XX
AC ADD19013;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human disease related protein DNA sequence SeqID502.
XX
KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW antiarteriosclerotic; vulnerary; gene therapy;
KW hypoxia-regulated condition; tumourigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003018621-A2.
XX
PD 06-MAR-2003.
XX
PF 23-AUG-2002; 2002WO-GB003892.
XX
PR 23-AUG-2001; 2001GB-00020558.
PR 05-OCT-2001; 2001GB-00024037.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
DR WPI; 2003-290046/28.
DR P-PSDB; ADD19012.
XX
PT New substantially purified polypeptide, useful for diagnosing or treating
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT wound healing.
XX
PS Claim 27; SEQ ID NO 502; 424pp; English.
XX
CC This invention relates to novel human genes and gene product which are
CC implicated in certain disease states. Compounds which modulate the
CC proteins of the invention may have cytostatic, antiinflammatory,
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The
CC sequences of the invention may be useful for gene therapy. The invention
CC may be useful for diagnosing or treating a hypoxia-regulated condition,
CC such as tumourigenesis, angiogenesis, apoptosis, inflammation,
CC erythropoiesis, or the biological response to hypoxia conditions
CC including processes such as glycolysis, gluconeogenesis, glucose
CC transportation, catecholamine synthesis, iron transport or nitric oxide
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC inflammatory conditions or wound healing. The present sequence is that of
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Query Match 50.8%; Score 666.6; DB 10; Length 1936;
Best Local Similarity 76.9%; Pred. No. 4.8e-174;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db         33 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 92

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db         93 GACTTCTGTGTGTGAGGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 152

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db        153 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 212

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db        213 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 272

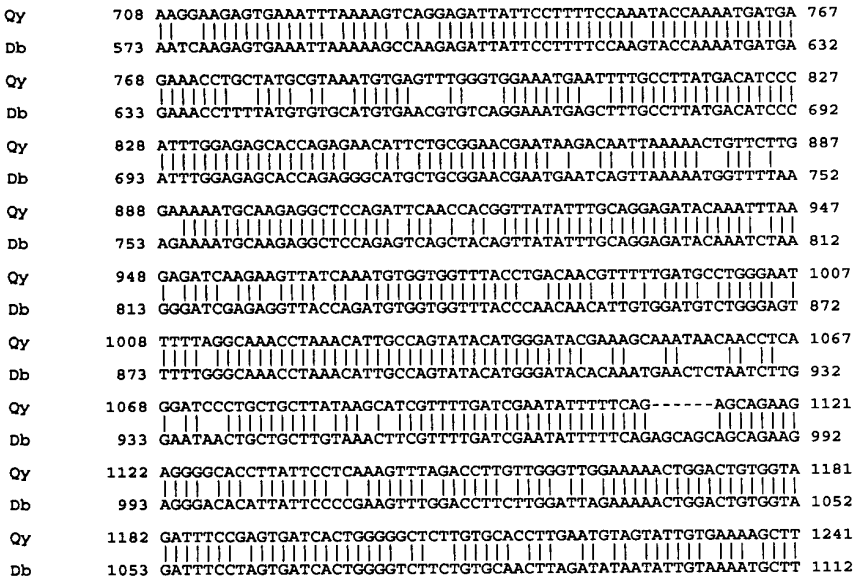
Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db        273 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 332

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db        333 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 392

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db        393 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 452

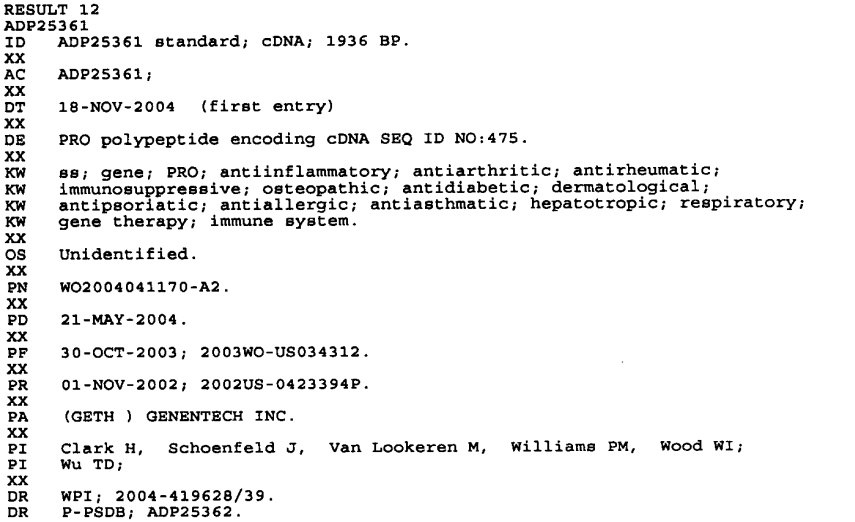
Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db        453 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 512

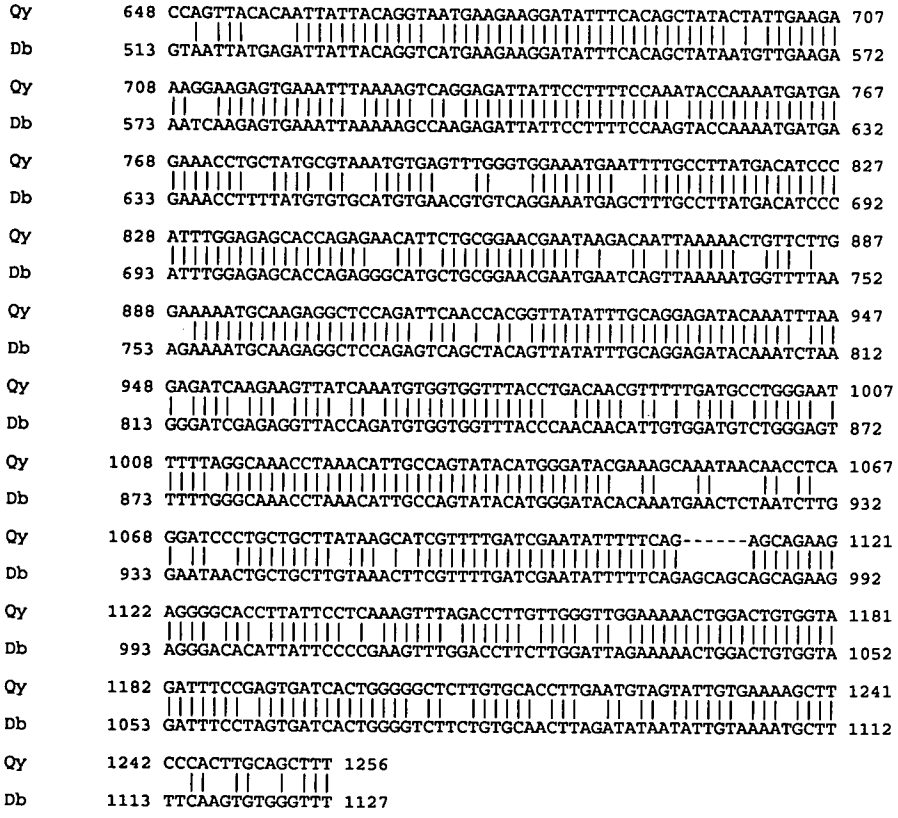
Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db        513 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 572
```



RESULT 12
ADP25361
ID ADP25361 standard; cDNA; 1936 BP.
XX
AC ADP25361;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide encoding cDNA SEQ ID NO:475.
XX
KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-419628/39.
DR P-PSDB; ADP25362.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 1; SEQ ID NO 475; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence encodes a PRO protein of
CC the invention.
XX
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

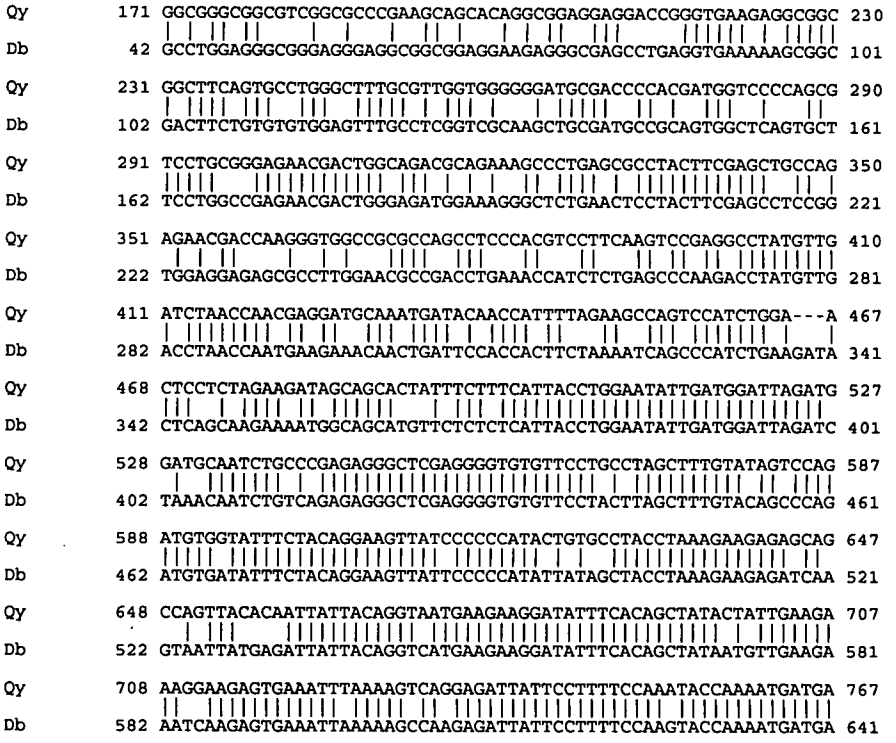
Query Match 50.8%; Score 666.6; DB 13; Length 1936;
Best Local Similarity 76.9%; Pred. No. 4.8e-174;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;





RESULT 13
ADX06356
ID ADX06356 standard; DNA; 1940 BP.
XX
AC ADX06356;
XX
DT 21-APR-2005 (first entry)
XX
DE Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 921.
XX
KW cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2005012875-A2.
XX
PD 10-FEB-2005.
XX
PF 29-JUL-2004; 2004WO-US024424.
XX
PR 29-JUL-2003; 2003US-0490890P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
XX
DR WPI; 2005-163068/17.
DR P-PSDB; ADX06357.
XX
PT Biomarkers useful for predicting or determining the response of a mammal
PT to a cancer treatment comprising administration of a modulator of cyclin-
PT dependent kinase activity.
XX
PS Claim 5; SEQ ID NO 921; 141pp; English.
XX
CC This invention describes a novel method of predicting or determining
CC whether a mammal will respond or is responding to an anti-cancer agent
CC that modulates cyclin-dependent kinase (cdk) activity. The method
CC comprises measuring the level of one or more biomarkers selected from
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC invention is utilized in a kit for determining or predicting whether
CC patient would be susceptible or resistant to treatment by an agent
CC modulating cdk activity. The invention also describes a method for
CC utilizing individualized genetic profiles for treating diseases and
CC disorders based on patient's response and molecular level, specialized
CC microarrays comprising the biomarkers described, antibodies directed
CC against the biomarkers and a cell culture model to identify biomarkers.
CC The cdk modulator is preferably N-5-[[5-(1,1-Dimethylethyl)-2-
CC oxazolyl]methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC tartaric acid salt. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC sequence encodes a biomarker used in the method of the invention.
XX
SQ Sequence 1940 BP; 613 A; 330 C; 439 G; 558 T; 0 U; 0 Other;

Query Match 50.8%; Score 666.6; DB 14; Length 1940;
Best Local Similarity 76.9%; Pred. No. 4.8e-174;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;



```
Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  ||||  ||  ||||||   ||   ||||||||  ||||||||||||||||
Db        642 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 701

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||||  |||  ||||||||||||  |  ||  ||||||  ||| |
Db        702 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 761

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||||||||||||||||  |||  |  ||  ||||||||||||||||||||||||  |||
Db        762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | ||||  |||  ||||  ||  ||||||||||||||  |||||  ||  |  ||||  ||||||  |
Db        822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              ||||  ||||||||||||||||||||||||||||||||  ||   ||    || ||
Db        882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||  |||  |  ||||||||||||||||||||||||      ||||||||
Db        942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              ||||  |||  |||||||  |  ||||||  ||||  ||  ||||||||||||||||||
Db       1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              |||||||  ||||||||||||||  ||  ||||||  |||  ||  ||  ||||||  |||  ||||
Db       1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

Qy       1242 CCCACTTGCAGCTTT 1256
                ||  ||  | |||
Db       1122 TTCAAGTGTGGGTTT 1136
```

RESULT 14
AAH15146
ID AAH15146 standard; cDNA; 1898 BP.
XX
AC AAH15146;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:13209.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 13209; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 1898 BP; 590 A; 327 C; 426 G; 555 T; 0 U; 0 Other;

Query Match 50.7%; Score 665; DB 4; Length 1898;
Best Local Similarity 76.8%; Pred. No. 1.3e-173;
Matches 841; Conservative 0; Mismatches 245; Indels 9; Gaps 2;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||      |||||| | |||||
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | ||||  |||  |||  |||||  | |||  |   |  |||||  ||  ||  |||  |   ||
Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   ||||||||||||  |||  |  |  ||  ||||  |  |||  ||||||||   ||  |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              |  |  ||    |  |  ||||  |||   ||    ||  ||  ||  ||||||||||
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | ||||||||  ||  ||  |||  ||||  |  ||||  ||    ||  |||  |||||||  |    |
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  |  ||||  ||  ||||||    |  |||  ||||||||||||||||||||||||||||||
Db        317 CTCAGCAAGAAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||||  |  ||||||||||||||||||||  ||||  |  ||||||||||  ||  ||||
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              |||||  |||||||||||||||||||  ||||||||  |  |  ||||||||||||||  ||
Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              |  |||      ||||||||||||  ||||||||||||||||||||||||||  |  |||||||
Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  ||||||||||||  |||||  ||  ||||||||||||||||||  ||||||||||||||
Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616
```

```
Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              ||||||| |||| ||  ||||||   ||   |||||||| ||||||||||||||||||
Db        617 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 676

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| |||||||||||| |  || |||||||  ||| |
Db        677 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 736

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        737 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 796

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  ||||| || | |||| |||||| |
Db        797 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 856

Qy        1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
               |||| ||||||||||||||||||||||||||||||||||  ||   ||    || ||
Db         857 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 916

Qy        1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
               | ||  ||||||||| ||| | |||||||||||||||||||||      ||||||||
Db         917 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 976

Qy        1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
               |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db         977 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1036

Qy        1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
               ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db        1037 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1096

Qy        1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db        1097 TTCAAGTGTGGGTTT 1111
```

RESULT 15
ADD34376
ID ADD34376 standard; DNA; 602 BP.
XX
AC ADD34376;
XX
DT 15-JAN-2004 (first entry)
XX
DE Mouse mitochondrial DNA sequence SEQ ID NO:2154.
XX
KW ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW mitochondrial disease; oxidative phosphorylation dysfunction;
KW oxidative stress; apoptosis; aging.
XX
OS Mus musculus.
XX
PN WO2003020220-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027886.
XX
PR 30-AUG-2001; 2001US-0316323P.
PR 31-AUG-2001; 2001CA-02356540.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Wallace DC, Levy S, Kerstann K, Procaccio V;
XX
DR WPI; 2003-300821/29.
XX
PT Array containing probes for genes involved in mitochondrial biology,
PT useful for determining mitochondrial biology gene expression profiles for
PT use in diagnosing pathologies and identifying biochemical pathways.
XX
PS Claim 2; SEQ ID NO 2154; 201pp; English.
XX
CC The invention relates to a novel array comprising at least two isolated
CC nucleotide molecules, each molecule having a sequence capable of uniquely
CC hybridising to a nucleic acid molecule which is an expression product of
CC a gene involved in mitochondrial biology. The array comprises two or more
CC isolated nucleic acid molecules or spots, each molecule having a sequence
CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC of the invention is useful for determining an expression profile of a
CC mouse or human sample containing nucleic acid, by contacting the array
CC with the sample under conditions allowing selective hybridisation, and
CC measuring hybridisation of nucleic acid in the sample to the array to
CC produce an expression profile. The array is also useful for determining
CC an expression profile of a first labelled sample containing nucleic acid
CC relative to a second, differently labelled sample containing nucleic
CC acid. The second sample is a reference or a standard. An array is useful
CC for determining an expression profile diagnostic of an energy-metabolism-
CC related physiological condition. An array of the invention is useful for
CC determining mitochondrial biology gene expression profiles of organisms,
CC such as human, mice and closely related species, tissue and organs of
CC such organisms, which are useful for determining expression profiles
CC diagnostic of energy metabolism-related physiological conditions,
CC diagnosing such physiological conditions, identifying biochemical
CC pathways, genes, and mutations involved in such physiological conditions,
CC identifying therapeutic agents useful for preventing and/or treating such
CC physiological conditions, evaluating and/or monitoring the efficacy of
CC such therapies, and creating and identifying animal models of human
CC energy metabolism-related physiological conditions. An array is also
CC useful for defining expression signatures or profiles for mitochondrial
CC diseases, as well as distinguishing clinical disorders that result from
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC apoptosis and aging. An array of the invention contains probes of genes
CC not previously recognised to participate in mitochondrial biology. The
CC sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC clones used to make the probes of the invention. Some sequences are not
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC 1906, 2408 and 2643.
XX
SQ Sequence 602 BP; 194 A; 119 C; 134 G; 155 T; 0 U; 0 Other;

  Query Match   45.8%; Score 600.4; DB 10; Length 602;
  Best Local Similarity 99.8%; Pred. No. 6.3e-156;
  Matches 601; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

```
Qy        357 ACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAA 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAA 60

Qy        417 CCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAG 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAG 120

Qy        477 AAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATC 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATC 180

Qy        537 TGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTAT 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTAT 240

Qy        597 TTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACA 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACA 300

Qy        657 CAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAG 360

Qy        717 TGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGC 776
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        361 TGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAGCCTGC 420

Qy        777 TATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGA 836
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGA 480

Qy        837 GCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGC 896
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGC 540

Qy        897 AAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAG 956
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAG 600

Qy        957 AA 958
                ||
Db        601 AA 602
```

RESULT 16
AAS86254
ID AAS86254 standard; cDNA; 1088 BP.
XX
AC AAS86254;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #22058.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG22067.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 22058; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1088 BP; 345 A; 182 C; 236 G; 325 T; 0 U; 0 Other;

```
  Query Match             37.6%;  Score 493.2;  DB 5;  Length 1088;
  Best Local Similarity   82.9%;  Pred. No. 5.1e-126;
  Matches  577;  Conservative    0;  Mismatches  113;  Indels    6;  Gaps    1;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || |||||||| ||||||||||||||||
Db          1 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 60

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  |||||||||||||||| ||   | |||    |||||||||
Db         61 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 120

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| | |
Db        121 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAACA 180

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| |||||||||||||||||||||  |||| ||  ||
Db        181 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   ||||||||  |||||||||||||||||||||||||||||||||| ||
Db        241 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAAAGGG 300

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| ||||||||||||| |  || |||||||  | | |    |||||||||||||||||
Db        301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAATGCAAGAGGCTCCA 360

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || |||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        361 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              ||||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||
Db        421 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 480

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||  ||   ||    || ||  | ||  |||||||| |||
Db        481 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 540

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
              | |||||||||||||||||||||||      |||||||||||| ||| ||||||| | |
Db        541 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 600

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
              ||||| |||||| |||| || ||||||||||||||||||||||||||| ||||||||||||
Db        601 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 660

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAA 1237
              || ||  |||||| |||  || || ||||   ||||
Db        661 GGTCTTCTGTGCAACTTAGATATAATATTTGTAAAA 696
```

RESULT 17
AAX84209/c
ID AAX84209 standard; cDNA; 1079 BP.
XX
AC AAX84209;
XX
DT 08-SEP-1999 (first entry)
XX
DE DNA encoding human breast tumour protein immunogenic fragment.
XX
KW Breast tumour protein; immunogenic fragment; vaccine; detection;
KW breast cancer development; therapy; ss.
XX
OS Homo sapiens.
XX

PN WO9933869-A2.
XX
PD 08-JUL-1999.
XX
PF 22-DEC-1998; 98WO-US027416.
XX
PR 24-DEC-1997; 97US-00998253.
PR 24-DEC-1997; 97US-00998255.
PR 17-JUL-1998; 98US-00118554.
PR 17-JUL-1998; 98US-00118627.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J;
XX
DR WPI; 1999-405486/34.
XX
PT New breast tumor protein genes used, in vaccines for immunotherapy, or
PT for diagnosis of breast cancer.
XX
PS Claim 3; Page 60; 70pp; English.
XX
CC This sequence encodes a human breast tumour protein immunogenic fragment
CC of the invention. The polypeptides or nucleic acids encoding them are
CC useful in vaccines and pharmaceutical compositions for manufacture of
CC medicaments for inhibiting the development of breast cancer in a patient.
CC They can also be used to treat breast cancer. Antibodies against these
CC polypeptides can be used to detect and monitor progression of breast
CC cancer in patients. Primers and probes derived from the polynucleotides
CC encoding the breast proteins are useful for detection of breast cancer.
CC Peripheral blood cells from a patient incubated in the presence of at
CC least one polypeptide, such that T cells proliferate, are useful in
CC manufacture of a medicament for treating breast cancer in a patient.
CC Antigen presenting cells incubated in the presence of at least one
CC polypeptide are also useful for treating breast cancer
XX
SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Query Match 37.3%; Score 490; DB 2; Length 1079;
Best Local Similarity 82.8%; Pred. No. 3.9e-125;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy       555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
            ||||||||||||| | |||||||||| || ||||||||| ||||||||||||||||||| |
Db      1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy       615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
            ||||||| | |  ||||||||||||||||| ||   | |||    |||||||||||| |||
Db      1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy       675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
            |||||||||||||||||||||| | |||||||||  |||||||||||| ||||| || |
Db       959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy       735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
            |||||||||||||||||| ||||||||||||||||||||  |||| ||  |||||| |
Db       899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy       795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
            ||   ||||||||  ||||||||||||||||||||||||||||||||||||  ||| |||
Db       839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy       855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
            |||||||||| |  || |||||||  ||| |    |||||||||||||||||||| |||
Db       779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy       915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
            | || |||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db       719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy       975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
            |||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||||||||
Db       659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600
```

Qy 1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
Db 599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy 1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
Db 539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy 1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
Db 480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy 1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
Db 420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373

RESULT 18
AAC79438/c
ID AAC79438 standard; cDNA; 1079 BP.
XX
AC AAC79438;
XX
DT 07-FEB-2001 (first entry)
XX
DE cDNA sequence of human breast tumour clone 1015D11.
XX
KW Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;
KW vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200061756-A2.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US009688.
XX
PR 09-APR-1999; 99US-00288950.
PR 02-JUL-1999; 99US-00346327.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J, Dillon DC;
XX
DR WPI; 2000-638568/61.
XX
PT A novel isolated polypeptide comprising an immunogenic portion of a
PT breast cancer protein useful in the detection and treatment of breast
PT cancer.
XX
PS Claim 4; Page 77; 95pp; English.
XX
CC The present sequence was isolated from a breast tumour cDNA library. It
CC is provided in a specification relating to compounds for immunotherapy
CC and diagnosis of breast cancer. Breast tumour antigens and the
CC polynucleotides that encode them may be used in the production of a
CC pharmaceutical composition to be used in the treatment of breast cancer.
CC Proliferated T cells and incubated antigen presenting cells are also
CC required. The polypeptides and polynucleotides may also be used to
CC produce a vaccine
XX
SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Query Match 37.3%; Score 490; DB 3; Length 1079;
Best Local Similarity 82.8%; Pred. No. 3.9e-125;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy       555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
            ||||||||||||| | |||||||||| || ||||||||| ||||||||||||||||||| |
Db      1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020
```

```
Qy       615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
Db      1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy       675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
Db       959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy       735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
Db       899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy       795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
Db       839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy       855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
Db       779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy       915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
Db       719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy       975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
Db       659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy      1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
Db       599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy      1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
Db       539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy      1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
Db       480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy      1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
Db       420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

RESULT 19
ABK28982/c
ID ABK28982 standard; cDNA; 1079 BP.
XX
AC ABK28982;
XX
DT 23-APR-2002 (first entry)
XX
DE Human breast tumour polypeptide cDNA clone #11.
XX
KW Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
KW immunostimulant.
XX
OS Homo sapiens.
XX
PN WO200198339-A2.
XX
PD 27-DEC-2001.
XX
PF 12-JUN-2001; 2001WO-US019032.
XX
PR 22-JUN-2000; 2000US-00602877.
PR 12-OCT-2000; 2000US-00687507.
PR 06-FEB-2001; 2001US-00778381.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;

XX
DR WPI; 2002-147792/19.
XX
PT Polynucleotides encoding breast tumor polypeptides, useful for treating
PT breast cancer or stimulating an immune response.
XX
PS Claim 1; Page 131; 150pp; English.
XX
CC The invention relates to polynucleotides encoding breast tumour
CC polypeptides. The sequences are useful for treating cancer, preferably
CC breast cancer, in a patient or for stimulating an immune response. The
CC polynucleotides and polypeptides are also useful in the diagnosis and
CC monitoring of breast cancer. A method for detecting the presence of a
CC cancer in a patient, comprises obtaining a biological sample from the
CC patient, contacting the biological sample with a binding agent that binds
CC to a breast tumour polypeptide, detecting in the sample an amount of
CC polypeptide that binds to the binding agent, and comparing the amount of
CC polypeptide to a predetermined cut-off value, therefore determining the
CC presence of a cancer in the patient. Sequences ABK28920-ABK29025
CC represent cDNA clones encoding human breast tumour polypeptides of the
CC invention
XX
SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Query Match 37.3%; Score 490; DB 6; Length 1079;
Best Local Similarity 82.8%; Pred. No. 3.9e-125;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy       555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
Db      1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy       615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
Db      1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy       675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
Db       959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy       735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
Db       899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy       795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
Db       839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy       855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
Db       779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy       915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
Db       719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy       975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
Db       659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy      1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
Db       599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy      1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
Db       539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy      1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
Db       480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy      1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
```

```
Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

RESULT 20
ADT95548
ID ADT95548 standard; cDNA; 674 BP.
XX
AC ADT95548;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #1067.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1067; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Query Match 35.2%; Score 461.6; DB 11; Length 674;
Best Local Similarity 83.4%; Pred. No. 2.3e-117;
Matches 548; Conservative 0; Mismatches 102; Indels 7; Gaps 2;

```
Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
Db         18 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA  77

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
Db         78 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 137

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
Db        138 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 197

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
Db        198 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
Db        258 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
Db        318 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCN 377

Qy        908 -GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAA 966
Db        378 TGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437

Qy        967 ATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACA 1026
Db        438 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 497

Qy       1027 TTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAA 1086
Db        498 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 557

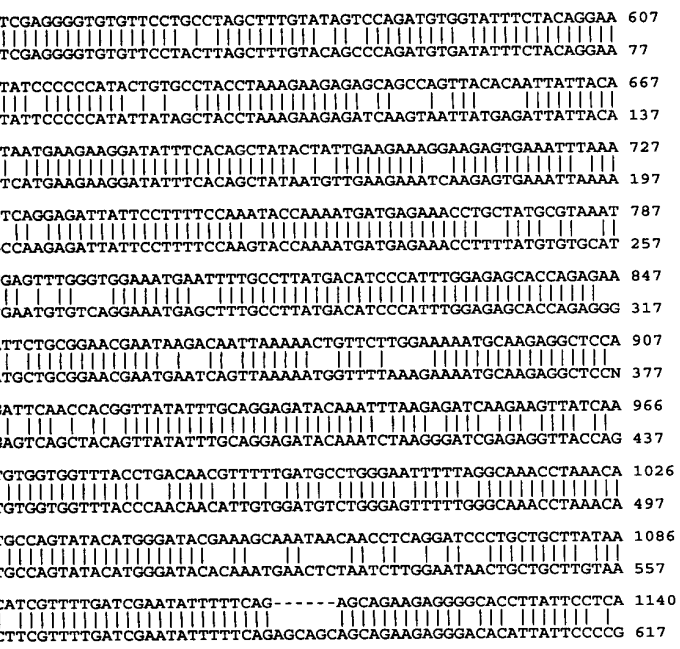
Qy       1087 GCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCA 1140
Db        558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCG 617

Qy       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCA 1197
Db        618 AAGTTTGGACCTTCTTGGATTANAAAAACTGGACTGNGGTAGATTTCCTAGTGATCA 674
```

RESULT 21
ADX42030
ID ADX42030 standard; cDNA; 674 BP.
XX
AC ADX42030;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 1067.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.

PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1067; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Query Match 35.2%; Score 461.6; DB 11; Length 674;
Best Local Similarity 83.4%; Pred. No. 2.3e-117;
Matches 548; Conservative 0; Mismatches 102; Indels 7; Gaps 2;



RESULT 22
ADT95551
ID ADT95551 standard; cDNA; 644 BP.
XX
AC ADT95551;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #1070.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1070; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;

Query Match 33.3%; Score 436.4; DB 11; Length 644;

Best Local Similarity 83.3%; Pred. No. 2.3e-110;
Matches 510; Conservative 0; Mismatches 96; Indels 6; Gaps 1;

```
Qy       548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
Db        32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy       608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
Db        92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy       668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
Db       152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy       728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
Db       212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy       788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
Db       272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy       848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
Db       332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy       908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
Db       392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy       968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
Db       452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy      1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
Db       512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy      1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
Db       572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy      1142 AGTTTAGACCTT 1153
Db       632 AGTTTGGACCTT 643
```

RESULT 23
ADX42033
ID ADX42033 standard; cDNA; 644 BP.
XX
AC ADX42033;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 1070.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1070; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;

Query Match 33.3%; Score 436.4; DB 11; Length 644;
Best Local Similarity 83.3%; Pred. No. 2.3e-110;
Matches 510; Conservative 0; Mismatches 96; Indels 6; Gaps 1;

```
Qy       548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
Db        32 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 91

Qy       608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
Db        92 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 151

Qy       668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
Db       152 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 211

Qy       728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
Db       212 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271

Qy       788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
Db       272 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 331

Qy       848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
Db       332 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCA 391

Qy       908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
Db       392 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451

Qy       968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
Db       452 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 511

Qy      1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
Db       512 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 571

Qy      1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
Db       572 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 631

Qy      1142 AGTTTAGACCTT 1153
```

```
                |||||| ||||||
Db        632 AGTTTGGACCTT 643


RESULT 24
ADT95512
ID   ADT95512 standard; cDNA; 625 BP.
XX
AC   ADT95512;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1031.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1031; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

  Query Match             32.7%;  Score 428.4;  DB 11;  Length 625;
  Best Local Similarity   84.1%;  Pred. No. 3.7e-108;
  Matches  483;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||    | |||    |||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||||| | |||||||||  ||||||||||| |||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||| ||||||||||||||||||||||||||||||||||||||
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||  || |||||||  ||| |   ||||||||||||||||||||
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || ||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||
Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||  ||   ||    || ||  | ||  |||||||| |||
Db        522 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 581

Qy       1088 CATCGTTTTGATCGAATATTTTTCAGAGCAGAAG 1121
              | |||||||||||||||||||||||||||| ||
Db        582 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAG 615


RESULT 25
ADX41994
ID   ADX41994 standard; cDNA; 625 BP.
XX
AC   ADX41994;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1031.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
```

```
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1031; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

  Query Match             32.7%;  Score 428.4;  DB 11;  Length 625;
  Best Local Similarity   84.1%;  Pred. No. 3.7e-108;
  Matches  483;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||| ||    | |||    |||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||  |||||||||||||||||||||||||||||||||||||
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||| |  || |||||||  ||| |    ||||||||||||||||||
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || |||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              ||||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||
Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||| ||   ||    || ||  | ||  ||||||||| |||
Db        522 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 581

Qy       1088 CATCGTTTTGATCGAATATTTTTCAGAGCAGAAG 1121
              | |||||||||||||||||||||||||||||| ||
Db        582 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAG 615


RESULT 26
ADT95565
ID   ADT95565 standard; cDNA; 633 BP.
XX
AC   ADT95565;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #1084.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1084; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

  Query Match             32.7%;  Score 428.4;  DB 11;  Length 633;
  Best Local Similarity   84.1%;  Pred. No. 3.8e-108;
  Matches  483;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
```

```
Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727

Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787

Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847

Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907

Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967

Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027

Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087

Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy       1088 CATCGTTTTGATCGAATATTTTTCAGAGCAGAAG 1121

Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAG 607


RESULT 27
ADX42047
ID   ADX42047 standard; cDNA; 633 BP.
XX
AC   ADX42047;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1084.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1084; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

  Query Match             32.7%;  Score 428.4;  DB 11;  Length 633;
  Best Local Similarity   84.1%;  Pred. No. 3.8e-108;
  Matches  483;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607

Db         34 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 93

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667

Db         94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727

Db        154 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 213

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787

Db        214 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847

Db        274 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907

Db        334 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 393

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967

Db        394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027

Db        454 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 513

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087

Db        514 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 573

Qy       1088 CATCGTTTTGATCGAATATTTTTCAGAGCAGAAG 1121

Db        574 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAG 607


RESULT 28
ADT95029
ID   ADT95029 standard; cDNA; 625 BP.
XX
AC   ADT95029;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Colon cancer associated human cDNA sequence #548.
XX
KW   Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW   humoral immune response; cellular immune response; cytostatic;
```

```
KW   immunostimulant; human; ss.
XX
OS   Homo sapiens.
XX
PN   US2003087818-A1.
XX
PD   08-MAY-2003.
XX
PF   01-FEB-2002; 2002US-00066543.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI   Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 548; 87pp; English.
XX
CC   The invention relates to polynucleotide and polypeptide sequences
CC   associated with cancer, particularly colon cancer. Also disclosed are (i)
CC   an expression vector comprising the polynucleotide, (ii) a host cell
CC   transformed or transfected with the expression vector, (iii) an isolated
CC   antibody, or its antigen-binding fragment, which specifically binds to
CC   the polypeptide, (iv) a method of detecting or determining the presence
CC   of cancer in a patient, (v) a fusion protein comprising at least one of
CC   the polypeptides, (vi) an oligonucleotide that hybridises to the
CC   polynucleotide sequence under highly stringent conditions, and (vii) a
CC   method of stimulating and/or expanding T cells specific for a tumour
CC   protein. The polypeptide specifically comprises the amino acid sequence
CC   of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC   frames (ORFs) of C636S. These polypeptides are encoded by the
CC   polynucleotide sequences, where both are capable of eliciting a humoral
CC   and/or cellular immune response. The polynucleotides, polypeptides, and
CC   antibodies are useful for diagnosing, preventing or treating cancer,
CC   particularly colon cancer. The polynucleotide and polypeptide sequences
CC   are also useful in DNA strand invasion, antisense inhibition, mutational
CC   analysis, nucleic acid purification, isolation of transcriptionally
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;

  Query Match           31.9%;  Score 418;  DB 11;  Length 625;
  Best Local Similarity  84.0%;  Pred. No. 2.9e-105;
  Matches  483;  Conservative    0;  Mismatches   91;  Indels    1;  Gaps    1;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||    | |||     |||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy        668 -GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAA 726
               ||| ||||||||||||||||||||||||| | |||||||||  |||||||||||| ||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy        727 AAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAA 786
              ||| || |||||||||||||||||| |||||||||||||||||||||  |||| ||  |
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy        787 TGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGA 846
              ||||| | ||   ||||||||  ||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        847 ACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCC 906
              ||| ||||||||||||| |  || |||||||  ||| |     ||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy        907 AGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAA 966
              ||| ||| | || ||||||||||||||||||||||||| |||| |||| ||| |||| ||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        967 ATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACA 1026
              |||||||||||||||  ||||| || | |||| |||||| |||||| ||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy       1027 TTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAA 1086
              |||||||||||||||||||  ||   ||    || ||  | ||  ||||||||| |||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy       1087 GCATCGTTTTGATCGAATATTTTTCAGAGCAGAAG 1121
              | |||||||||||||||||||||||||||||| ||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAG 600


RESULT 29
ADX41511
ID   ADX41511 standard; cDNA; 625 BP.
XX
AC   ADX41511;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 548.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI   Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 548; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
```

CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;

Query Match 31.9%; Score 418; DB 11; Length 625;
Best Local Similarity 84.0%; Pred. No. 2.9e-105;
Matches 483; Conservative 0; Mismatches 91; Indels 1; Gaps 1;

```
Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         26 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGNTATTTCTACAGGAA 85

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||| ||    | |||    |||||||||
Db         86 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 145

Qy        668 -GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAA 726
               ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| ||
Db        146 GGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAA 205

Qy        727 AAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAA 786
              ||| || |||||||||||||||||| ||||||||||||||||||||  |||| ||  |
Db        206 AAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 265

Qy        787 TGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGA 846
              ||||| | ||   |||||||  ||||||||||||||||||||||||||||||||||||
Db        266 TGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

Qy        847 ACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCC 906
              ||| ||||||||||||| |  || |||||||  ||| |    ||||||||||||||||||
Db        326 GCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCC 385

Qy        907 AGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAA 966
              ||| ||| | || ||||||||||||||||||||||| |||| |||| ||| |||| ||
Db        386 AGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445

Qy        967 ATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACA 1026
              |||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||
Db        446 ATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACA 505

Qy       1027 TTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAA 1086
              |||||||||||||||||||  ||    ||    || ||  | ||  ||||||||| |||
Db        506 TTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAA 565

Qy       1087 GCATCGTTTTGATCGAATATTTTTCAGAGCAGAAG 1121
              | |||||||||||||||||||||||||||||| ||
Db        566 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAG 600
```

RESULT 30
ADT95464
ID ADT95464 standard; cDNA; 553 BP.
XX
AC ADT95464;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #983.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 983; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Query Match 29.4%; Score 385.8; DB 11; Length 553;
Best Local Similarity 84.9%; Pred. No. 2.4e-96;
Matches 432; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```
Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||| ||    | |||    |||||||||
Db        102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||  ||||||||||||||||||||||||||||||||||||||
Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
```

```
Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967

Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027

Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAA 1056

Db        522 TGCCAGTATACATGGGATACACAAATGAA 550
```

RESULT 31
ADX41946
ID ADX41946 standard; cDNA; 553 BP.
XX
AC ADX41946;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 983.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 983; 244pp; English.
XX
CC The invention relates to a new isolated nucleic acid. The nucleic acids,
CC polypeptides, antibodies are useful for diagnosing, preventing or
CC treating cancer, particularly colon cancer. The nucleic acid and
CC polypeptides are also useful in DNA strand invasion, antisense
CC inhibition, mutational analysis, nucleic acid purification, isolation of
CC transcriptionally active genes, blocking or transcription factor binding,
CC genome cleavage or in situ hybridization, and as enhancers of
CC transcription or biomarkers. The kits are useful for detecting antibody
CC binding. The present sequence represents a human cDNA encoding a colon
CC cancer protein.
XX
SQ Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Query Match 29.4%; Score 385.8; DB 11; Length 553;
Best Local Similarity 84.9%; Pred. No. 2.4e-96;
Matches 432; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```
Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607

Db         42 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 101

Qy        608 GTTATCCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667

Db        102 GTTATTCCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727

Db        162 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 221

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787

Db        222 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847

Db        282 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907

Db        342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 401

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967

Db        402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027

Db        462 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 521

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAA 1056

Db        522 TGCCAGTATACATGGGATACACAAATGAA 550
```

RESULT 32
AAH08073
ID AAH08073 standard; cDNA; 752 BP.
XX
AC AAH08073;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:4908.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length

PT cDNAs.
XX
PS Claim 1; SEQ ID NO 4908; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 752 BP; 225 A; 148 C; 200 G; 176 T; 0 U; 3 Other;

Query Match 28.8%; Score 377.8; DB 4; Length 752;
Best Local Similarity 72.5%; Pred. No. 4.7e-94;
Matches 528; Conservative 0; Mismatches 195; Indels 5; Gaps 3;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     |||||| | |||||
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| ||| ||| ||||| | ||| |   | |||||  || || |||  |   ||
Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||  |||||||||||| ||| | | |  || |||| | ||| ||||||||  || |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   |  | |  |||| |||   ||   ||   || || || |||||||||
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | ||||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||
Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | ||||||||||||||||||||| |||| | |||||||||| || ||||
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||   |||||||||||| ||||||||||||||||||||||||| | |||||||
Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616
```

```
Qy        768 G-AAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 826
              | ||||||  |||| ||  ||||||   ||   ||||||||  |||||||| |||||||
Db        617 GAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTAAGACATCC 676

Qy        827 C-ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCT 885
              | ||| || || ||||| ||  |||     |||||||| |  || || ||||  | | |
Db        677 CAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATCAGTTTAAAATGGGTTT 736

Qy        886 TGGAAAAA 893
                 |||
Db        737 AAANGAAA 744
```

RESULT 33
ADT95844
ID ADT95844 standard; cDNA; 401 BP.
XX
AC ADT95844;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #1363.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1363; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally

```
CC   active genes, blocking or transcription factor binding, genome cleavage
CC   or in situ hybridisation, and as enhancers of transcription or
CC   biomarkers. This sequence represents a human colon cancer.associated
CC   cDNA. Note: The sequence data for this patent was obtained in electronic
CC   format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ   Sequence 401 BP; 132 A; 67 C; 89 G; 113 T; 0 U; 0 Other;

  Query Match            22.1%;  Score 290.2;  DB 11;  Length 401;
  Best Local Similarity  84.9%;  Pred. No. 7.1e-70;
  Matches  325;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         19 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 78

Qy        608 GTTATCCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||   | |||    |||||||||
Db         79 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 138

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        139 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 198

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        199 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||| ||||||||||||||||||||||||||||||||||||||
Db        259 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 318

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||| |  || |||||||  ||| |    ||||||||||||||||||
Db        319 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 378

Qy        908 GATTCAACCACGGTTATATTTGC 930
              || ||| | || ||||||||||
Db        379 GAGTCAGCTACAGTTATATTTGC 401


RESULT 34
ADX42326
ID   ADX42326 standard; cDNA; 401 BP.
XX
AC   ADX42326;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human cDNA encoding colon cancer protein SEQ ID NO 1363.
XX
KW   Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW   ss; gene.
XX
OS   Homo sapiens.
XX
PN   WO200274156-A2.
XX
PD   26-SEP-2002.
XX
PF   01-FEB-2002; 2002WO-US002870.
XX
PR   02-FEB-2001; 2001US-0267400P.
PR   07-FEB-2001; 2001US-0267382P.
PR   11-MAY-2001; 2001US-0290322P.
PR   12-JUL-2001; 2001US-0305265P.
PR   16-AUG-2001; 2001US-0313077P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 1363; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 401 BP; 132 A; 67 C; 89 G; 113 T; 0 U; 0 Other;

  Query Match            22.1%;  Score 290.2;  DB 11;  Length 401;
  Best Local Similarity  84.9%;  Pred. No. 7.1e-70;
  Matches  325;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||||| | |||||||||| || ||||||||| ||||||||||||||
Db         19 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 78

Qy        608 GTTATCCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  ||||||||||||||||| ||   | |||    |||||||||
Db         79 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 138

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  |||||||||||| |||
Db        139 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 198

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        199 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||| ||||||||||||||||||||||||||||||||||||||
Db        259 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 318

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||| |  || |||||||  ||| |    ||||||||||||||||||
Db        319 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 378

Qy        908 GATTCAACCACGGTTATATTTGC 930
              || ||| | || ||||||||||
Db        379 GAGTCAGCTACAGTTATATTTGC 401


RESULT 35
AAX40590
ID   AAX40590 standard; cDNA; 483 BP.
XX
AC   AAX40590;
XX
DT   18-JUN-1999  (first entry)
XX
DE   Human secreted protein 5' EST SEQ ID No: 190.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS   Homo sapiens.
XX
```

PN WO9906550-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB001232.
XX
PR 01-AUG-1997; 97US-00905144.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
DR WPI; 1999-153780/13.
DR P-PSDB; AAY11868.
XX
PT New isolated prostate-derived nucleic acids - used to develop products
PT which may have cytokine, immune regulatory, haematopoiesis regulating,
PT anti-inflammatory or tumour inhibition activity.
XX
PS Claim 1; Page 298; 675pp; English.
XX
CC AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins expressed in prostate, and encode the proteins
CC given in AAY11716 to AAY11993 respectively. The proteins given represent
CC the signal peptide and an N-terminal fragment of a secreted protein. The
CC nucleic acid sequences can be used for producing secreted human gene
CC products. They can also be used to develop products for diagnosis and
CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation and differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptides can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 483 BP; 123 A; 111 C; 139 G; 110 T; 0 U; 0 Other;

Query Match 17.7%; Score 231.8; DB 2; Length 483;
Best Local Similarity 70.0%; Pred. No. 1.3e-53;
Matches 327; Conservative 0; Mismatches 137; Indels 3; Gaps 1;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    | | |  || | | | ||  |||     |||||| | |||||
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||  |  | | |  |||| |||   ||   ||  || || || |||||||||
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||
Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||||| | ||||||||||| || ||||
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 436
```

Qy 588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTA 634
||||| |||||||||||||||||| |||||||| | | |||||||
Db 437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTA 483

RESULT 36
AAS86255
ID AAS86255 standard; cDNA; 1227 BP.
XX
AC AAS86255;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #22059.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG22068.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 22059; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1227 BP; 311 A; 239 C; 360 G; 317 T; 0 U; 0 Other;

Query Match 15.1%; Score 198.2; DB 5; Length 1227;
Best Local Similarity 70.3%; Pred. No. 4.5e-44;
Matches 282; Conservative 0; Mismatches 113; Indels 6; Gaps 1;

Qy 851 TCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGAT 910
| | |||||||||| | || ||||||| ||| | ||||||||||||||||||||

```
Db        554 TATACGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAG 613

Qy        911 TCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGT 970
              ||| | || ||||||||||||||||||||||||| |||| |||| ||| || |   |
Db        614 TCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTGAGTGACAAA 673

Qy        971 GGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGC 1030
              |   ||||      | |                ||    |    |     ||||
Db        674 GTCAGTTTGATGTTTGAGGACAGATTATTGCTATGGAGACTAATTTAGGTTCCTTTTTGC 733

Qy       1031 CAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCAT 1090
              |||||||||||||||||  ||    ||    || ||  | ||  ||||||||| ||| | |
Db        734 CAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTT 793

Qy       1091 CGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGT 1144
              ||||||||||||||||||||||      |||||||||||| ||| ||||||| | ||||
Db        794 CGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGT 853

Qy       1145 TTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGG 1204
              || |||||| |||| || |||||||||||||||||||||||||| ||||||||||||||
Db        854 TTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGT 913

Qy       1205 CTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCA 1245
              ||  |||||| |||  || || |||||| ||| ||||  ||
Db        914 CTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCA 954
```

RESULT 37
ADT95307/c
ID ADT95307 standard; cDNA; 391 BP.
XX
AC ADT95307;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #826.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
DR WPI; 2003-040540/03.
XX
PT New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 826; 87pp; English.
XX
CC The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 391 BP; 83 A; 100 C; 103 G; 104 T; 0 U; 1 Other;

Query Match 10.7%; Score 140.8; DB 11; Length 391;
Best Local Similarity 69.6%; Pred. No. 2.2e-28;
Matches 206; Conservative 0; Mismatches 87; Indels 3; Gaps 1;

```
Qy        262 GGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAA 321
              | | |||||  || |  |||  |   || |||||   |||||||||||| ||| | | |
Db        338 GAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAG 279

Qy        322 AGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCC 381
              || |||| | ||||||||||||  || | | | ||    |  | |   ||||  |||
Db        278 GGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGA 219

Qy        382 CACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAAC 441
              ||   ||   || || || |||||||||| |||||||| || ||  ||| |||| | ||
Db        218 AACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCAC 159

Qy        442 CATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTT 498
              || ||   ||  ||| ||||||| |   ||||  | |||| || ||||||   | ||| |
Db        158 CACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCT 99

Qy        499 CATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAG 554
              ||||||||||||||||||||||||||  |  |||||||  | ||||||||||||
Db         98 CATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAG 43
```

RESULT 38
ADX41789/c
ID ADX41789 standard; cDNA; 391 BP.
XX
AC ADX41789;
XX
DT 21-APR-2005 (first entry)
XX
DE Human cDNA encoding colon cancer protein SEQ ID NO 826.
XX
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
KW ss; gene.
XX
OS Homo sapiens.
XX
PN WO200274156-A2.
XX
PD 26-SEP-2002.
XX
PF 01-FEB-2002; 2002WO-US002870.
XX
PR 02-FEB-2001; 2001US-0267400P.
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305265P.
PR 16-AUG-2001; 2001US-0313077P.

```
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Chenault RA,  Xu J,  Indirias CY,  Lodes MJ,  Secrist H;
PI   Carter D,  Fanger GR,  Smith CL,  Durham M,  Stolk JA;
XX
DR   WPI; 2003-040540/03.
XX
PT   New isolated nucleic acids and polypeptides capable of eliciting a
PT   humoral and/or cellular immune response, useful for diagnosing,
PT   preventing or treating cancer, particularly colon cancer.
XX
PS   Claim 1; SEQ ID NO 826; 244pp; English.
XX
CC   The invention relates to a new isolated nucleic acid. The nucleic acids,
CC   polypeptides, antibodies are useful for diagnosing, preventing or
CC   treating cancer, particularly colon cancer. The nucleic acid and
CC   polypeptides are also useful in DNA strand invasion, antisense
CC   inhibition, mutational analysis, nucleic acid purification, isolation of
CC   transcriptionally active genes, blocking or transcription factor binding,
CC   genome cleavage or in situ hybridization, and as enhancers of
CC   transcription or biomarkers. The kits are useful for detecting antibody
CC   binding. The present sequence represents a human cDNA encoding a colon
CC   cancer protein.
XX
SQ   Sequence 391 BP; 83 A; 100 C; 103 G; 104 T; 0 U; 1 Other;

  Query Match           10.7%;  Score 140.8;  DB 11;  Length 391;
  Best Local Similarity   69.6%;  Pred. No. 2.2e-28;
  Matches  206;  Conservative    0;  Mismatches   87;  Indels    3;  Gaps    1;

Qy        262 GGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAA 321
              | | |||||  || |  |||  |   || |||||   ||||||||||||| ||| | | |
Db        338 GAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAG 279

Qy        322 AGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCC 381
               || |||| | ||||||||||||  || | | | ||   |  | |   |||| |||
Db        278 GGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGA 219

Qy        382 CACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAAC 441
               ||   ||   || || || |||||||||| |||||||| || ||  ||| |||| | ||
Db        218 AACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCAC 159

Qy        442 CATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTT 498
              || ||   ||  ||| ||||||| |   ||||  | |||| || ||||||   | ||| |
Db        158 CACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCT 99

Qy        499 CATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAG 554
              |||||||||||||||||||||||||||||   |  ||||||| | ||||||||||||
Db         98 CATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAG 43


RESULT 39
ABL36461/c
ID   ABL36461 standard; cDNA; 179 BP.
XX
AC   ABL36461;
XX
DT   08-APR-2002  (first entry)
XX
DE   Human colon tumour antigen polynucleotide SEQ ID NO:50.
XX
KW   Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW   colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200196388-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US018557.
XX
PR   09-JUN-2000; 2000US-0210899P.
PR   20-FEB-2001; 2001US-0270216P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Harlocker SL,  Secrist H;
XX
DR   WPI; 2002-114514/15.
XX
PT   Novel isolated colon tumor polynucleotide differentially expressed in
PT   colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT   useful for inhibiting development of cancer in patient.
XX
PS   Claim 1; SEQ ID NO 50; 105pp; English.
XX
CC   ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC   which were isolated from human colon tumour and colon metastatic tumour
CC   cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC   production. (I) can be used for stimulating and/or expanding T cells
CC   specific for a tumour protein on contact with the T cells. They are also
CC   useful for inhibiting the development of cancer in a patient. (I) can be
CC   used as probes or primers for nucleic acid hybridisation, for preparing
CC   mutant species primers, or primers for use in genetic constructions. (I)
CC   can be used in the diagnosis of a colon tumour
XX
SQ   Sequence 179 BP; 51 A; 31 C; 27 G; 68 T; 0 U; 2 Other;

  Query Match           10.3%;  Score 135.6;  DB 6;  Length 179;
  Best Local Similarity   84.4%;  Pred. No. 4e-27;
  Matches  151;  Conservative    1;  Mismatches   27;  Indels    0;  Gaps    0;

Qy        576 TGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAA 635
              | :| || ||||||||| |||||||||||||||||||| |||||||| | |  ||||||||
Db        179 TNHACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAA 120

Qy        636 AGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTA 695
              |||||||| ||   | |||     |||||||||||| |||||||||||||||||||||||
Db        119 AGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTA 60

Qy        696 TACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              || | |||||||||  |||||||||||| ||||| || |||||||||||||||||||| |
Db         59 TAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1


RESULT 40
ABK27684
ID   ABK27684 standard; cDNA; 176 BP.
XX
AC   ABK27684;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human colon cancer expressed sequence tag, Seq ID no 121.
XX
KW   Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
KW   expressed sequence tag.
XX
OS   Homo sapiens.
XX
PN   WO200196390-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US018577.
XX
PR   09-JUN-2000; 2000US-0210821P.
PR   18-DEC-2000; 2000US-0256571P.
PR   10-MAY-2001; 2001US-0290240P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Hepler WT,  Clapper JD,  Wang A,  Secrist H;
XX
```

DR WPI; 2002-139708/18.
XX
PT Novel isolated polynucleotide encoding a polypeptide comprising a portion
PT of colon tumor protein, useful for detection, diagnosis and therapy of
PT human colon cancer.
XX
PS Claim 1; Page 182; 220pp; English.
XX
CC The invention relates to an isolated polynucleotide (I) encoding a
CC polypeptide (II) comprising at least a portion of a colon tumour protein.
CC (I), (II) and antibody (III) to (II) are useful for determining the
CC presence of a cancer in a patient. (I), (II) or antigen presenting cells
CC expressing (I) is useful for stimulating and/or expanding T cells
CC specific for a tumour protein, by contacting T cells with (I), (II) or
CC antigen-presenting cells that express (I), under conditions and for a
CC time sufficient to permit the stimulation and/or expansion of T cells.
CC (I), (II), or antigen presenting cells that express (II) are useful for
CC treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
CC isolated from a patient with (I), (II) or antigen presenting cells that
CC express (II), such that T cells proliferate, and administering to the
CC patient an effective amount of the proliferated T cells, thus inhibiting
CC the development of a cancer in the patient. (I) or (II) is useful in
CC vaccines and pharmaceutical compositions for prevention and treatment of
CC colon malignancies and for the diagnosis and monitoring of such cancers.
CC (I), (II) or (III) is useful for detection, diagnosis and/or therapy of
CC human colon cancer. (I) is useful as a probe or primer for nucleic acid
CC hybridisation, and in the design and preparation of ribozyme molecules
CC for inhibiting expression of (II) in tumour cells. ABK27564-ABK27807
CC represent novel human colon cancer coding sequences and primers of the
CC invention
XX
SQ Sequence 176 BP; 68 A; 27 C; 31 G; 50 T; 0 U; 0 Other;

Query Match 10.2%; Score 134.4; DB 6; Length 176;
Best Local Similarity 85.2%; Pred. No. 8.6e-27;
Matches 150; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy        579 ATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGA 638
              | || ||||||||| |||||||||||||||||||| |||||||| | |  ||||||||||
Db          1 ACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGA 60

Qy        639 AGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATAC 698
              ||||| ||   | |||    |||||||||||| ||||||||||||||||||||||||||
Db         61 AGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAA 120

Qy        699 TATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              | |||||||||  |||||||||||| ||||| || ||||||||||||||||||| |
Db        121 TGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176
```

RESULT 41
ABV95959/c
ID ABV95959 standard; cDNA; 176 BP.
XX
AC ABV95959;
XX
DT 14-JAN-2003 (first entry)
XX
DE Human pancreatic cancer expressed cDNA SEQ ID NO 1367.
XX
KW Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW cytostatic; tumour; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200260317-A2.
XX
PD 08-AUG-2002.
XX
PF 30-JAN-2002; 2002WO-US002781.
XX
PR 30-JAN-2001; 2001US-0265305P.
PR 31-JAN-2001; 2001US-0265682P.
PR 09-FEB-2001; 2001US-0267568P.
PR 21-MAR-2001; 2001US-0278651P.
PR 28-APR-2001; 2001US-0287112P.
PR 16-MAY-2001; 2001US-0291631P.
PR 12-JUL-2001; 2001US-0305484P.
PR 20-AUG-2001; 2001US-0313999P.
PR 27-NOV-2001; 2001US-0333626P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX
DR WPI; 2002-627435/67.
XX
PT New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT diagnosing, preventing and/or treating cancer, particularly pancreatic
PT cancer.
XX
PS Claim 1; SEQ ID NO 1367; 300pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising: (a)
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
CC complements of (a); (c) sequences consisting of at least 20 contiguous
CC residues of (a); (d) sequences that hybridize to (a), under moderately
CC stringent conditions; (e) sequences having at least 75% or 90% identity
CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
CC ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer
CC in a patient and compositions comprising polypeptides, polynucleotides,
CC antibodies, fusion proteins, T cell populations and antigen presenting
CC cells expressing the polypeptide are useful in treating pancreatic cancer
CC and stimulating an immune response. The polynucleotides can be used as
CC probes or primers for nucleic acid hybridisation, in the design and
CC preparation of ribozyme molecules for inhibiting expression of the tumour
CC polypeptides and proteins in the tumour cells, in vaccines and for gene
CC therapy. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 176 BP; 50 A; 31 C; 27 G; 68 T; 0 U; 0 Other;

Query Match 10.2%; Score 134.4; DB 6; Length 176;
Best Local Similarity 85.2%; Pred. No. 8.6e-27;
Matches 150; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy        579 ATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGA 638
              | || ||||||||| |||||||||||||||||||| |||||||| | |  ||||||||||
Db        176 ACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGA 117

Qy        639 AGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATAC 698
              ||||| ||   | |||    |||||||||||| ||||||||||||||||||||||||||
Db        116 AGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAA 57

Qy        699 TATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              | |||||||||  |||||||||||| ||||| || ||||||||||||||||||| |
Db         56 TGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1
```

RESULT 42
ABT07150/c
ID ABT07150 standard; cDNA; 268 BP.
XX
AC ABT07150;
XX
DT 07-NOV-2002 (first entry)
XX
DE Human ovarian cancer associated coding sequence SEQ ID NO: 312.
XX
KW Human; ovarian cancer; cancer; gene; ss.
XX
OS Homo sapiens.
XX
PN US2002076715-A1.
XX

PD 20-JUN-2002.
XX
PF 06-JUN-2001; 2001US-00876889.
XX
PR 23-SEP-1998; 98US-00159320.
PR 08-FEB-1999; 99US-00246429.
PR 16-SEP-1999; 99US-00397787.
XX
PA (BENS/) BENSON D R.
PA (LODE/) LODES M J.
PA (MITC/) MITCHAM J L.
PA (KING/) KING G E.
XX
PI Benson DR, Lodes MJ, Mitcham JL, King GE;
XX
DR WPI; 2002-598720/64.
XX
PT Composition for detecting and treating ovarian cancer, comprises a
PT specific polypeptide, polynucleotide, T cell population, or antigen
PT presenting cell.
XX
PS Example 1; Page 116; 188pp; English.
XX
CC The present invention relates to a method of detecting the presence of
CC ovarian cancer in a patient, involving detecting ovarian cancer
CC associated polynucleotides. The method is not only used to detect the
CC presence of cancer, preferably ovarian cancer in a patient, but also is
CC used to stimulate and/or expand T cells specific for an ovarian tumour
CC protein. The sequences can be used in vaccines used to treat cancer. The
CC present sequence is an ovarian cancer associated coding sequence
XX
SQ Sequence 268 BP; 87 A; 62 C; 43 G; 76 T; 0 U; 0 Other;

Query Match 7.7%; Score 101.4; DB 6; Length 268;
Best Local Similarity 81.8%; Pred. No. 1.6e-17;
Matches 117; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy       1114 AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGA 1173
             |||||||||||| ||| ||||||| | |||||| |||||| |||| || ||||||||||
Db        262 AGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGA 203

Qy       1174 CTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTG 1233
             |||||||||||||||| |||||||||||||| ||  |||||| |||  || || ||||||
Db        202 CTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTA 143

Qy       1234 AAAAGCTTCCCACTTGCAGCTTT 1256
              ||| ||||  ||  ||  | |||
Db        142 AAATGCTTTTCAAGTGTGGGTTT 120
```

RESULT 43
ABX73028/c
ID ABX73028 standard; cDNA; 268 BP.
XX
AC ABX73028;
XX
DT 14-MAR-2003 (first entry)
XX
DE Human ovarian carcinoma antigen partial cDNA sequence #294.
XX
KW Human; cancer detection; ovarian carcinoma antigen; ovarian cancer;
KW tumour antigen; tumour; OV2; OV3; OV6; OV9; OV10; OV12; OV14; OV17; OV18;
KW OV23; OV24; OV27; OV41; OV54; OV57; gene; ss.
XX
OS Homo sapiens.
XX
PN US6468758-B1.
XX
PD 22-OCT-2002.
XX
PF 16-SEP-1999; 99US-00397787.
XX
PR 23-SEP-1998; 98US-00159320.
PR 08-FEB-1999; 99US-00246429.
XX
PA (CORI-) CORIXA CORP.
XX
PI Benson DR, Lodes MJ, Mitcham JL, King GE;
XX
DR WPI; 2003-147101/14.
XX
PT Determining presence or absence of cancer in patient by contacting
PT patient sample with oligonucleotide that hybridizes to polynucleotide
PT encoding ovarian carcinoma antigen, and detecting amount of hybridization
PT in sample.
XX
PS Example 1; Col 223-224; 152pp; English.
XX
CC The present invention relates to compositions and methods for determining
CC presence or absence of cancer in patient. The method comprising
CC contacting a biological sample with an oligonucleotide that hybridises to
CC partial polynucleotide sequence encoding for human ovarian carcinoma
CC antigen, or its complement. The method and compositions are useful for
CC the therapy and diagnosis of ovarian cancer. The method can be used to
CC identify tumour antigens that are secreted from ovarian carcinoma and/or
CC other tumours. Effective cancer detection is achieved using the method of
CC the invention. ABX72720-ABX73050 represent partial cDNA sequences
CC encoding human ovarian carcinoma antigens
XX
SQ Sequence 268 BP; 87 A; 62 C; 43 G; 76 T; 0 U; 0 Other;

Query Match 7.7%; Score 101.4; DB 8; Length 268;
Best Local Similarity 81.8%; Pred. No. 1.6e-17;
Matches 117; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy       1114 AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGA 1173
             |||||||||||| ||| ||||||| | |||||| |||||| |||| || ||||||||||
Db        262 AGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGA 203

Qy       1174 CTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTG 1233
             |||||||||||||||| |||||||||||||| ||  |||||| |||  || || ||||||
Db        202 CTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTA 143

Qy       1234 AAAAGCTTCCCACTTGCAGCTTT 1256
              ||| ||||  ||  ||  | |||
Db        142 AAATGCTTTTCAAGTGTGGGTTT 120
```

RESULT 44
AAX41008
ID AAX41008 standard; cDNA; 258 BP.
XX
AC AAX41008;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human secreted protein 5' EST SEQ ID NO: 220.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS Homo sapiens.
XX
PN WO9906554-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB001238.
XX
PR 01-AUG-1997; 97US-00905134.
XX
PA (GEST) GENSET.

XX
PI Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
DR WPI; 1999-153784/13.
DR P-PSDB; AAY12175.
XX
PT New nucleic acids encoding human secreted proteins - obtained from cDNA
PT libraries prepared from kidney, fetal kidney, dystrophic muscle, muscle
PT and heart tissue.
XX
PS Claim 1; Page 319; 622pp; English.
XX
CC AAX40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY01602 and
CC AAY11994 to AAY12260, respectively. The proteins given represent the
CC signal peptide and an N-terminal fragment of a secreted protein. The
CC nucleic acid sequences can be used for producing secreted human gene
CC products. They can also be used to develop products for diagnosis and
CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 258 BP; 54 A; 54 C; 105 G; 40 T; 0 U; 5 Other;

Query Match 5.7%; Score 75.2; DB 2; Length 258;
Best Local Similarity 58.7%; Pred. No. 2.9e-10;
Matches 122; Conservative 4; Mismatches 82; Indels 0; Gaps 0;

```
Qy      171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
            | | || ||    |  |   |  || |   | | ||  |||     ||||||| | |||||
Db       47 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 106

Qy      231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
            | |||| ||| :|:|  |||:| | ||| |   | |||||  || |  |||  |   ||
Db      107 GACTTCTGTGTSTRGAGTTTRCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 166

Qy      291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
            |||||   ||:|||||||| ||| | | |  || |||| | ||||||||||||  || |
Db      167 TCCTGGCCGAKAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 226

Qy      351 AGAACGACCAAGGGTGGCCGCGCCAGCC 378
            | | ||    |  | |   |||| ||
Db      227 TGGAGGAGAGCGCCTTGGAACGCCGACC 254
```

RESULT 45
AAI15311
ID AAI15311 standard; DNA; 444 BP.
XX
AC AAI15311;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #5244 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 5244; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 444 BP; 138 A; 60 C; 80 G; 166 T; 0 U; 0 Other;

Query Match 5.7%; Score 74.6; DB 4; Length 444;
Best Local Similarity 71.3%; Pred. No. 5.6e-10;
Matches 112; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

```
Qy      577 GTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAA 636
            ||| || |||||||||| ||||||||||||||||||||| |||||||| | |  |||||||||
Db      253 GTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAA 312

Qy      637 GAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTAT 696
            ||||||| ||    | |||    |||||||||||||  |     | ||||  ||   | |
Db      313 GAAGAGATCAAGTAATTATGAGATTATTACAGGTAACAATTTTTGTTATTCAACTTGTCT 372

Qy      697 ACT-ATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 732
               |  ||||| ||  ||  |||||||   | || |||
Db      373 TTTCTTTGAAAAACAGATCAGTGAAAGAGATAAATCA 409
```

RESULT 46
AAI24484
ID AAI24484 standard; DNA; 691 BP.
XX
AC AAI24484;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #14417 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.

```
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 14417; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 691 BP; 230 A; 81 C; 112 G; 268 T; 0 U; 0 Other;

  Query Match             5.7%;  Score 74.6;  DB 4;  Length 691;
  Best Local Similarity   71.3%;  Pred. No. 7.1e-10;
  Matches  112;  Conservative    0;  Mismatches   44;  Indels    1;  Gaps    1;

Qy        577 GTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAA 636
              ||| || |||||||||| ||||||||||||||||||| |||||||| | |  |||||||||
Db        115 GTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAA 174

Qy        637 GAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTAT 696
              ||||||| ||   | |||    ||||||||||||| |    | ||||  ||    || |
Db        175 GAAGAGATCAAGTAATTATGAGATTATTACAGGTAACAATTTTTGTTATTCAACTTGTCT 234

Qy        697 ACT-ATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 732
                |  ||||| ||  ||  |||||||   | ||| |||
Db        235 TTTCTTTGAAAAACAGATCAGTGAAAGAGATAAATCA 271


RESULT 47
ABT09027
ID   ABT09027 standard; DNA; 705 BP.
XX
AC   ABT09027;
XX
DT   05-DEC-2002  (first entry)
XX
DE   Phase-1 Rat CT gene SEQ ID No 115.
XX
KW   Rat; toxicity study; rat toxic response gene; toxicological response;
KW   drug development; phase-1 rat CT gene; ds.
XX
OS   Rattus sp.
XX
PN   WO200266682-A2.
XX
PD   29-AUG-2002.
XX
PF   29-JAN-2002; 2002WO-US002935.
XX
PR   29-JAN-2001; 2001US-0264933P.
PR   26-JUL-2001; 2001US-0308161P.
XX
PA   (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI   Farris G,  Hicken SH,  Farr SB;
XX
DR   WPI; 2002-674961/72.
XX
PT   Evaluating the toxicity of an agent, useful in drug development or in
PT   determining toxicological responses to a new drug, by determining the
PT   expression of rat toxicologically relevant genes in the test animal in
PT   response to the test agent.
XX
PS   Disclosure; Page 134; 388pp; English.
XX
CC   The invention relates to a method used for evaluating the toxicity of an
CC   agent comprising determining the expression of a rat toxic response
CC   gene(s) in the test animal in response to the agent. The method is useful
CC   in drug development, particularly for conducting toxicity studies and
CC   analysis before a new drug or compound is approved for human consumption
CC   or use. The method is also useful in determining toxicological responses
CC   to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC   gene of the invention
XX
SQ   Sequence 705 BP; 220 A; 141 C; 173 G; 170 T; 0 U; 1 Other;

  Query Match             5.1%;  Score 67.4;  DB 6;  Length 705;
  Best Local Similarity   60.8%;  Pred. No. 7.2e-08;
  Matches  110;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
              ||  |||| || ||||| | |   |       |   |||     |   |  ||| | |
Db         76 GGACAGCAACGGGAAGACCAGCCATGAGTCAAGTCAGCTGGATGAACCAAGCGTGGAAAC 135

Qy        137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
              |      ||  || | |  | ||     |||| | |  | |   | || |  ||||
Db        136 ACACAGCCTGGAGCAGTCCAAGGAGTATAAGCAGAGGGCCAGCCACGAGAGCACTGAGCA 195

Qy        197 G 197
              |
Db        196 G 196


RESULT 48
ADG45549
ID   ADG45549 standard; DNA; 705 BP.
XX
AC   ADG45549;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Liver inflammatory predictive gene related DNA sequence.
XX
KW   liver; liver toxicity; liver toxicity predictive gene;
KW   liver inflammation predictive gene; inflammation; gene; ds; rat.
XX
OS   Rattus norvegicus.
XX
PN   WO2003095624-A2.
XX
PD   20-NOV-2003.
XX
PF   09-MAY-2003; 2003WO-US014832.
XX
PR   10-MAY-2002; 2002US-0379831P.
XX
PA   (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI   Kier L,  Nolan TD,  Sankar U,  Derbel M;
XX
DR   WPI; 2004-042456/04.
XX
PT   Predicting the liver toxicity in an individual to an agent by measuring
```

PT the expression of one or more liver toxicity predictive genes in the
PT sample from the individual treated with the agent.
XX
PS Claim 7; Fig 4; 336pp; English.
XX
CC The present invention describes a method for predicting the liver
CC toxicity in an individual to an agent. The method comprises: (a)
CC obtaining a biological sample from the individual treated with the agent;
CC (b) measuring the expression of one or more liver toxicity predictive
CC genes in the sample; and (c) using the test expression profile with a set
CC of reference expression profiles in a Predictive Model to determine
CC whether the agent will induce liver toxicity in the individual. Also
CC described: (1) predicting the liver toxicity of an agent using an in
CC vitro system; (2) predicting the liver toxicity in a biological sample
CC from an individual, in vitro cell cultures or explants to an agent via a
CC programmable machine; (3) a computer program product for enabling a
CC computer to perform Predictive Model analysis for liver toxicity on a
CC biological sample from an individual, in vitro cell cultures or explants
CC to an agent; (4) a computer system adopted to predict liver toxicity in a
CC biological system from an individual, in vitro cell cultures or explants
CC to an agent; (5) a computer program product for predicting liver toxicity
CC from a test sample expression profile; (6) mining genes predictive of
CC liver toxicity; (7) an integrated system for predicting liver toxicity;
CC and (8) identifying one or more liver inflammation predictive genes. The
CC method is useful for predicting the liver toxicity in an individual to an
CC agent. The present sequence represents a liver toxicity predictive gene
CC sequence that represents a 24 hour combo 3 gene, which is specifically
CC claimed in the exemplification of the present invention.
XX
SQ Sequence 705 BP; 220 A; 141 C; 173 G; 170 T; 0 U; 1 Other;

Query Match 5.1%; Score 67.4; DB 12; Length 705;
Best Local Similarity 60.8%; Pred. No. 7.2e-08;
Matches 110; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
              ||  |||| || ||||| | |   |       |    |||     |    |  ||| | |
Db         76 GGACAGCAACGGGAAGACCAGCCATGAGTCAAGTCAGCTGGATGAACCAAGCGTGGAAAC 135

Qy        137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
              |     ||  || | |  | ||    |||| | |   | |    | || |  ||||
Db        136 ACACAGCCTGGAGCAGTCCAAGGAGTATAAGCAGAGGGCCAGCCACGAGAGCACTGAGCA 195

Qy        197 G 197
              |
Db        196 G 196
```

RESULT 49
ADH22854
ID ADH22854 standard; DNA; 705 BP.
XX
AC ADH22854;
XX
DT 11-MAR-2004 (first entry)
XX
DE Partial DNA seqeunce of a rat kidney toxicity predictive gene (182).
XX
KW kidney toxicity; toxicology; predictive model; gene expression profile;
KW toxic damage; kidney tubule necrosis; acute renal failure; rat; ds.
XX
OS Rattus sp.
XX
PN WO2003100030-A2.
XX
PD 04-DEC-2003.
XX
PF 27-FEB-2003; 2003WO-US006196.
XX
PR 27-FEB-2002; 2002US-0361128P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Kier L, Nolan TD, Sankar U, Derbel M;
XX
DR WPI; 2004-035137/03.
XX
PT Predicting kidney toxicity in an individual to an agent, useful for
PT predicting toxic responses to one or more agents comprising measuring the
PT expression of one or more kidney toxicity predictive genes.
XX
PS Claim 1; Page 350; 388pp; English.
XX
CC This invention relates to novel isolated kidney toxicity predictive genes
CC and methods of using such genes in the field of toxicology. Specifically,
CC it refers to novel genes that can be used for generating predictive
CC models, which in turn are useful for predicting the in vivo toxic
CC response to one or more agents. The present invention describes obtaining
CC a gene expression profile from a biological sample and using the
CC predictive model to determine whether an agent will induce kidney
CC toxicity in the individual. As such, it can be used to detect any toxic
CC effects that may be manifested as long lasting or chronic consequences
CC including irreversible toxicity or carcinogenesis. Furthermore, the
CC predictive genes can be considered as therapeutic targets for toxic
CC damage or to ameliorate specific disease conditions such as kidney tubule
CC necrosis or acute renal failure, as well as for additional screening
CC assays. This polynucleotide sequence is the partial gene sequence of a
CC rat kidney toxicity predictive gene of the invention.
XX
SQ Sequence 705 BP; 220 A; 141 C; 173 G; 170 T; 0 U; 1 Other;

Query Match 5.1%; Score 67.4; DB 12; Length 705;
Best Local Similarity 60.8%; Pred. No. 7.2e-08;
Matches 110; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
              ||  |||| || ||||| | |   |       |    |||     |    |  ||| | |
Db         76 GGACAGCAACGGGAAGACCAGCCATGAGTCAAGTCAGCTGGATGAACCAAGCGTGGAAAC 135

Qy        137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
              |     ||  || | |  | ||    |||| | |   | |    | || |  ||||
Db        136 ACACAGCCTGGAGCAGTCCAAGGAGTATAAGCAGAGGGCCAGCCACGAGAGCACTGAGCA 195

Qy        197 G 197
              |
Db        196 G 196
```

RESULT 50
ABT09296
ID ABT09296 standard; DNA; 631 BP.
XX
AC ABT09296;
XX
DT 05-DEC-2002 (first entry)
XX
DE Phase-1 Rat CT gene SEQ ID No 384.
XX
KW Rat; toxicity study; rat toxic response gene; toxicological response;
KW drug development; phase-1 rat CT gene; ds.
XX
OS Rattus sp.
XX
PN WO200266682-A2.
XX
PD 29-AUG-2002.
XX
PF 29-JAN-2002; 2002WO-US002935.

XX
PR 29-JAN-2001; 2001US-0264933P.
PR 26-JUL-2001; 2001US-0308161P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Farris G, Hicken SH, Farr SB;
XX
DR WPI; 2002-674961/72.
XX
PT Evaluating the toxicity of an agent, useful in drug development or in
PT determining toxicological responses to a new drug, by determining the
PT expression of rat toxicologically relevant genes in the test animal in
PT response to the test agent.
XX
PS Disclosure; Page 201; 388pp; English.
XX
CC The invention relates to a method used for evaluating the toxicity of an
CC agent comprising determining the expression of a rat toxic response
CC gene(s) in the test animal in response to the agent. The method is useful
CC in drug development, particularly for conducting toxicity studies and
CC analysis before a new drug or compound is approved for human consumption
CC or use. The method is also useful in determining toxicological responses
CC to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC gene of the invention
XX
SQ Sequence 631 BP; 177 A; 107 C; 154 G; 193 T; 0 U; 0 Other;

Query Match 4.9%; Score 64.8; DB 6; Length 631;
Best Local Similarity 78.0%; Pred. No. 3.6e-07;
Matches 78; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 77

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTG 116
              ||| | || |   | |   | | |  |||  |||| ||
Db         78 TGGATGAAGAGAACATATGAGCTTGTGAGAGAAGACCGTG 117
```

RESULT 51
ADG45660
ID ADG45660 standard; DNA; 631 BP.
XX
AC ADG45660;
XX
DT 26-FEB-2004 (first entry)
XX
DE Liver inflammatory predictive gene related DNA sequence.
XX
KW liver; liver toxicity; liver toxicity predictive gene;
KW liver inflammation predictive gene; inflammation; gene; ds; rat.
XX
OS Rattus norvegicus.
XX
PN WO2003095624-A2.
XX
PD 20-NOV-2003.
XX
PF 09-MAY-2003; 2003WO-US014832.
XX
PR 10-MAY-2002; 2002US-0379831P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Kier L, Nolan TD, Sankar U, Derbel M;
XX
DR WPI; 2004-042456/04.
XX
PT Predicting the liver toxicity in an individual to an agent by measuring
PT the expression of one or more liver toxicity predictive genes in the
PT sample from the individual treated with the agent.
XX
PS Claim 7; Fig 4; 336pp; English.
XX
CC The present invention describes a method for predicting the liver
CC toxicity in an individual to an agent. The method comprises: (a)
CC obtaining a biological sample from the individual treated with the agent;
CC (b) measuring the expression of one or more liver toxicity predictive
CC genes in the sample; and (c) using the test expression profile with a set
CC of reference expression profiles in a Predictive Model to determine
CC whether the agent will induce liver toxicity in the individual. Also
CC described: (1) predicting the liver toxicity of an agent using an in
CC vitro system; (2) predicting the liver toxicity in a biological sample
CC from an individual, in vitro cell cultures or explants to an agent via a
CC programmable machine; (3) a computer program product for enabling a
CC computer to perform Predictive Model analysis for liver toxicity on a
CC biological sample from an individual, in vitro cell cultures or explants
CC to an agent; (4) a computer system adopted to predict liver toxicity in a
CC biological system from an individual, in vitro cell cultures or explants
CC to an agent; (5) a computer program product for predicting liver toxicity
CC from a test sample expression profile; (6) mining genes predictive of
CC liver toxicity; (7) an integrated system for predicting liver toxicity;
CC and (8) identifying one or more liver inflammation predictive genes. The
CC method is useful for predicting the liver toxicity in an individual to an
CC agent. The present sequence represents a liver toxicity predictive gene
CC sequence that represents a 24 hour combo 3 gene, which is specifically
CC claimed in the exemplification of the present invention.
XX
SQ Sequence 631 BP; 177 A; 107 C; 154 G; 193 T; 0 U; 0 Other;

Query Match 4.9%; Score 64.8; DB 12; Length 631;
Best Local Similarity 78.0%; Pred. No. 3.6e-07;
Matches 78; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 77

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTG 116
              ||| | || |   | |   | | |  |||  |||| ||
Db         78 TGGATGAAGAGAACATATGAGCTTGTGAGAGAAGACCGTG 117
```

RESULT 52
ADH22946
ID ADH22946 standard; DNA; 631 BP.
XX
AC ADH22946;
XX
DT 11-MAR-2004 (first entry)
XX
DE Partial DNA seqeunce of a rat kidney toxicity predictive gene (274).
XX
KW kidney toxicity; toxicology; predictive model; gene expression profile;
KW toxic damage; kidney tubule necrosis; acute renal failure; rat; ds.
XX
OS Rattus sp.
XX
PN WO2003100030-A2.
XX
PD 04-DEC-2003.
XX
PF 27-FEB-2003; 2003WO-US006196.
XX
PR 27-FEB-2002; 2002US-0361128P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Kier L, Nolan TD, Sankar U, Derbel M;
XX
DR WPI; 2004-035137/03.
XX
PT Predicting kidney toxicity in an individual to an agent, useful for
PT predicting toxic responses to one or more agents comprising measuring the

PT expression of one or more kidney toxicity predictive genes.
XX
PS Claim 1; Page 368; 388pp; English.
XX
CC This invention relates to novel isolated kidney toxicity predictive genes
CC and methods of using such genes in the field of toxicology. Specifically,
CC it refers to novel genes that can be used for generating predictive
CC models, which in turn are useful for predicting the in vivo toxic
CC response to one or more agents. The present invention describes obtaining
CC a gene expression profile from a biological sample and using the
CC predictive model to determine whether an agent will induce kidney
CC toxicity in the individual. As such, it can be used to detect any toxic
CC effects that may be manifested as long lasting or chronic consequences
CC including irreversible toxicity or carcinogenesis. Furthermore, the
CC predictive genes can be considered as therapeutic targets for toxic
CC damage or to ameliorate specific disease conditions such as kidney tubule
CC necrosis or acute renal failure, as well as for additional screening
CC assays. This polynucleotide sequence is the partial gene sequence of a
CC rat kidney toxicity predictive gene of the invention.
XX
SQ Sequence 631 BP; 177 A; 107 C; 154 G; 193 T; 0 U; 0 Other;

Query Match 4.9%; Score 64.8; DB 12; Length 631;
Best Local Similarity 78.0%; Pred. No. 3.6e-07;
Matches 78; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 77

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTG 116
                ||| | || |   | |  | |  |  |||  |||| ||
Db         78 TGGATGAAGAGAACATATGAGCTTGTGAGAGAAGACCGTG 117
```

RESULT 53
ADR91260
ID ADR91260 standard; cDNA; 631 BP.
XX
AC ADR91260;
XX
DT 16-DEC-2004 (first entry)
XX
DE Spleen necrosis predictive cDNA sequence, SEQ ID No 216.
XX
KW altered expression; toxic response; spleen; toxicity; lymphoid; gene; ss.
XX
OS Unidentified.
XX
PN WO2004083402-A2.
XX
PD 30-SEP-2004.
XX
PF 17-MAR-2004; 2004WO-US008371.
XX
PR 17-MAR-2003; 2003US-0455443P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Sankar U, Kier L, Derbel M, Nolan T;
XX
DR WPI; 2004-691048/67.
XX
PT New composition comprises cDNAs useful for detecting altered expression
PT of genes in a toxic response of the spleen or for predicting toxic
PT responses to one or more agents including lymphoid tissue types or other
PT species.
XX
PS Claim 1; SEQ ID NO 216; 249pp; English.
XX
CC The invention relates to a novel composition comprising cDNAs for use in
CC detecting the altered expression of genes in a toxic response of the
CC spleen, where the cDNAs comprises 50-816 base pairs (ADR91045-ADR91348)
CC or their complete complements. The invention further comprises:
CC monitoring the treatment of compound toxicity in a sample; predicting the
CC spleen toxicity in an individual to an agent; predicting the spleen
CC toxicity of an agent using an in vitro system; a computer program product
CC for predicting spleen toxicity from an expression profile of nucleic
CC acids from a sample under test, comprising a computer readable medium
CC bearing an encrypted training data set, encrypted lists of genes selected
CC from the cDNAs, and a predictive model for causing a general purpose
CC computer to predict the spleen toxicity of the sample based upon the
CC training data set, the list of genes selected from the cDNAs, and the
CC expression profile of nucleic acids from the sample; and an integrated
CC system for predicting spleen toxicity, comprising means for measuring
CC gene expression profiles of spleen predictive genes from samples exposed
CC to the test agent and a computer system operably linked to the means that
CC is capable of implementing a predictive model. The composition comprising
CC cDNAs is useful for detecting altered expression of genes in a toxic
CC response of the spleen or for predicting toxic responses to one or more
CC agents including lymphoid tissue types or other species. The predictive
CC genes and models of the invention are useful for identifying and
CC evaluating various in vitro systems that can be used to accurately
CC predict in vivo toxicity. This polynucleotide sequence represents one of
CC the genes predictive for spleen necrosis of the invention.
XX
SQ Sequence 631 BP; 177 A; 107 C; 154 G; 193 T; 0 U; 0 Other;

Query Match 4.9%; Score 64.8; DB 13; Length 631;
Best Local Similarity 78.0%; Pred. No. 3.6e-07;
Matches 78; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 77

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTG 116
                ||| | || |   | |  | |  |  |||  |||| ||
Db         78 TGGATGAAGAGAACATATGAGCTTGTGAGAGAAGACCGTG 117
```

RESULT 54
ABT09191
ID ABT09191 standard; DNA; 690 BP.
XX
AC ABT09191;
XX
DT 05-DEC-2002 (first entry)
XX
DE Phase-1 Rat CT gene SEQ ID No 279.
XX
KW Rat; toxicity study; rat toxic response gene; toxicological response;
KW drug development; phase-1 rat CT gene; ds.
XX
OS Rattus sp.
XX
PN WO200266682-A2.
XX
PD 29-AUG-2002.
XX
PF 29-JAN-2002; 2002WO-US002935.
XX
PR 29-JAN-2001; 2001US-0264933P.
PR 26-JUL-2001; 2001US-0308161P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Farris G, Hicken SH, Farr SB;
XX
DR WPI; 2002-674961/72.
XX
PT Evaluating the toxicity of an agent, useful in drug development or in
PT determining toxicological responses to a new drug, by determining the
PT expression of rat toxicologically relevant genes in the test animal in
PT response to the test agent.
XX

PS Disclosure; Page 175; 388pp; English.
XX
CC The invention relates to a method used for evaluating the toxicity of an
CC agent comprising determining the expression of a rat toxic response
CC gene(s) in the test animal in response to the agent. The method is useful
CC in drug development, particularly for conducting toxicity studies and
CC analysis before a new drug or compound is approved for human consumption
CC or use. The method is also useful in determining toxicological responses
CC to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC gene of the invention
XX
SQ Sequence 690 BP; 200 A; 169 C; 195 G; 126 T; 0 U; 0 Other;

Query Match 4.9%; Score 63.8; DB 6; Length 690;
Best Local Similarity 61.8%; Pred. No. 7.1e-07;
Matches 118; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGAT-GGCGTCTGGCAG 135
                | |    ||| |  | |  | |       ||  ||||  | ||  |  | || | |
Db         77 CGAACCTCGCGGGGCGGGAAGCCGCGAGATGGACACCCCTCCGCTCTCAGACTCGGACTC 136

Qy        136 CAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGC 195
              | | || |||| |     | || |||  |  |    || ||  ||||   | | || |
Db        137 CGGGTCGGATGAGTGCCTGGCCTCAGATCAAGAGTTGCAGGATGCGTTTTCCCGCGGACT 196

Qy        196 AGCACAGGCGG 206
                | || | |
Db        197 CCTAAAGCCAG 207
```

RESULT 55
ADG30969
ID ADG30969 standard; DNA; 690 BP.
XX
AC ADG30969;
XX
DT 26-FEB-2004 (first entry)
XX
DE Liver toxicity predictive DNA 149.
XX
KW liver toxicity; Norway rat; 24 hour combo All; ds.
XX
OS Unidentified.
XX
PN WO2003085083-A2.
XX
PD 16-OCT-2003.
XX
PF 01-APR-2003; 2003WO-US010141.
XX
PR 01-APR-2002; 2002US-0369287P.
XX
PA (PHAS-) PHASE 1 MOLECULAR TOXICOLOGY.
XX
PI Kier L, Nolan TD, Sankar U, Derbel M;
XX
DR WPI; 2003-804300/75.
XX
PT Predicting the liver toxicity of an agent to an individual by using the
PT test expression profile with a set of reference expression profiles in a
PT Predictive Model to determine whether the agent will reduce liver
PT toxicity in the individual.
XX
PS Claim 2; Page 177; 379pp; English.
XX
CC The invention relates to a novel method for predicting the liver toxicity
CC of an agent to an individual comprising obtaining a biological sample
CC from an individual treated with the agent, measuring the expression of
CC one or more liver toxicity predictive genes in the sample and using the
CC test expression profile with a set of reference expression profiles in a
CC predictive model to determine whether the agent will reduce liver
CC toxicity in the individual. The method of the invention may be useful for
CC predicting the liver toxicity of an agent to an individual. The current
CC sequence is that of the liver toxicity predictive DNA of the invention
CC which represents a 24 hour combo All gene.
XX
SQ Sequence 690 BP; 200 A; 169 C; 195 G; 126 T; 0 U; 0 Other;

Query Match 4.9%; Score 63.8; DB 10; Length 690;
Best Local Similarity 61.8%; Pred. No. 7.1e-07;
Matches 118; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGAT-GGCGTCTGGCAG 135
                | |    ||| |  | |  | |       ||  ||||  | ||  |  | || | |
Db         77 CGAACCTCGCGGGGCGGGAAGCCGCGAGATGGACACCCCTCCGCTCTCAGACTCGGACTC 136

Qy        136 CAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGC 195
              | | || |||| |     | || |||  |  |    || ||  ||||   | | || |
Db        137 CGGGTCGGATGAGTGCCTGGCCTCAGATCAAGAGTTGCAGGATGCGTTTTCCCGCGGACT 196

Qy        196 AGCACAGGCGG 206
                | || | |
Db        197 CCTAAAGCCAG 207
```

RESULT 56
ADG45586
ID ADG45586 standard; DNA; 690 BP.
XX
AC ADG45586;
XX
DT 26-FEB-2004 (first entry)
XX
DE Liver inflammatory predictive gene related DNA sequence.
XX
KW liver; liver toxicity; liver toxicity predictive gene;
KW liver inflammation predictive gene; inflammation; gene; ds; rat.
XX
OS Rattus norvegicus.
XX
PN WO2003095624-A2.
XX
PD 20-NOV-2003.
XX
PF 09-MAY-2003; 2003WO-US014832.
XX
PR 10-MAY-2002; 2002US-0379831P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Kier L, Nolan TD, Sankar U, Derbel M;
XX
DR WPI; 2004-042456/04.
XX
PT Predicting the liver toxicity in an individual to an agent by measuring
PT the expression of one or more liver toxicity predictive genes in the
PT sample from the individual treated with the agent.
XX
PS Claim 7; Fig 4; 336pp; English.
XX
CC The present invention describes a method for predicting the liver
CC toxicity in an individual to an agent. The method comprises: (a)
CC obtaining a biological sample from the individual treated with the agent;
CC (b) measuring the expression of one or more liver toxicity predictive
CC genes in the sample; and (c) using the test expression profile with a set
CC of reference expression profiles in a Predictive Model to determine
CC whether the agent will induce liver toxicity in the individual. Also
CC described: (1) predicting the liver toxicity of an agent using an in

CC vitro system; (2) predicting the liver toxicity in a biological sample
CC from an individual, in vitro cell cultures or explants to an agent via a
CC programmable machine; (3) a computer program product for enabling a
CC computer to perform Predictive Model analysis for liver toxicity on a
CC biological sample from an individual, in vitro cell cultures or explants
CC to an agent; (4) a computer system adopted to predict liver toxicity in a
CC biological system from an individual, in vitro cell cultures or explants
CC to an agent; (5) a computer program product for predicting liver toxicity
CC from a test sample expression profile; (6) mining genes predictive of
CC liver toxicity; (7) an integrated system for predicting liver toxicity;
CC and (8) identifying one or more liver inflammation predictive genes. The
CC method is useful for predicting the liver toxicity in an individual to an
CC agent. The present sequence represents a liver toxicity predictive gene
CC sequence that represents a 24 hour combo 3 gene, which is specifically
CC claimed in the exemplification of the present invention.
XX
SQ Sequence 690 BP; 200 A; 169 C; 195 G; 126 T; 0 U; 0 Other;

Query Match 4.9%; Score 63.8; DB 12; Length 690;
Best Local Similarity 61.8%; Pred. No. 7.1e-07;
Matches 118; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
Db 17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy 77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGAT-GGCGTCTGGCAG 135
Db 77 CGAACCTCGCGGGGCGGGAAGCCGCGAGATGGACACCCCTCCGCTCTCAGACTCGGACTC 136

Qy 136 CAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGC 195
Db 137 CGGGTCGGATGAGTGCCTGGCCTCAGATCAAGAGTTGCAGGATGCGTTTTCCCGCGGACT 196

Qy 196 AGCACAGGCGG 206
Db 197 CCTAAAGCCAG 207

RESULT 57
ADH22884
ID ADH22884 standard; DNA; 690 BP.
XX
AC ADH22884;
XX
DT 11-MAR-2004 (first entry)
XX
DE Partial DNA seqeunce of a rat kidney toxicity predictive gene (212).
XX
KW kidney toxicity; toxicology; predictive model; gene expression profile;
KW toxic damage; kidney tubule necrosis; acute renal failure; rat; ds.
XX
OS Rattus sp.
XX
PN WO2003100030-A2.
XX
PD 04-DEC-2003.
XX
PF 27-FEB-2003; 2003WO-US006196.
XX
PR 27-FEB-2002; 2002US-0361128P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Kier L, Nolan TD, Sankar U, Derbel M;
XX
DR WPI; 2004-035137/03.
XX
PT Predicting kidney toxicity in an individual to an agent, useful for
PT predicting toxic responses to one or more agents comprising measuring the
PT expression of one or more kidney toxicity predictive genes.
XX
PS Claim 1; Page 356; 388pp; English.

XX
CC This invention relates to novel isolated kidney toxicity predictive genes
CC and methods of using such genes in the field of toxicology. Specifically,
CC it refers to novel genes that can be used for generating predictive
CC models, which in turn are useful for predicting the in vivo toxic
CC response to one or more agents. The present invention describes obtaining
CC a gene expression profile from a biological sample and using the
CC predictive model to determine whether an agent will induce kidney
CC toxicity in the individual. As such, it can be used to detect any toxic
CC effects that may be manifested as long lasting or chronic consequences
CC including irreversible toxicity or carcinogenesis. Furthermore, the
CC predictive genes can be considered as therapeutic targets for toxic
CC damage or to ameliorate specific disease conditions such as kidney tubule
CC necrosis or acute renal failure, as well as for additional screening
CC assays. This polynucleotide sequence is the partial gene sequence of a
CC rat kidney toxicity predictive gene of the invention.
XX
SQ Sequence 690 BP; 200 A; 169 C; 195 G; 126 T; 0 U; 0 Other;

Query Match 4.9%; Score 63.8; DB 12; Length 690;
Best Local Similarity 61.8%; Pred. No. 7.1e-07;
Matches 118; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
Db 17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy 77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGAT-GGCGTCTGGCAG 135
Db 77 CGAACCTCGCGGGGCGGGAAGCCGCGAGATGGACACCCCTCCGCTCTCAGACTCGGACTC 136

Qy 136 CAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGC 195
Db 137 CGGGTCGGATGAGTGCCTGGCCTCAGATCAAGAGTTGCAGGATGCGTTTTCCCGCGGACT 196

Qy 196 AGCACAGGCGG 206
Db 197 CCTAAAGCCAG 207

RESULT 58
AAK52001
ID AAK52001 standard; cDNA; 969 BP.
XX
AC AAK52001;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 546.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.

```
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y;
PI   Ma Y,  Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wejhrman T,  Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   P-PSDB; AAM78868.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT   in diagnosis and gene therapy.
XX
PS   Claim 1; Page 1968-1969; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC   (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC   sequence listing were missing at the time of publication
XX
SQ   Sequence 969 BP; 248 A; 269 C; 220 G; 232 T; 0 U; 0 Other;

  Query Match             4.8%;  Score 63.6;  DB 4;  Length 969;
  Best Local Similarity   88.5%;  Pred. No. 9.6e-07;
  Matches   69;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          9 ATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 68
              ||||  ||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db          9 ATGACACGAATTTTTACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 68

Qy         69 ACGAGGGCGGGAAGCAGC 86
              ||||||| ||     |||
Db         69 ACGAGGGTGGCCGATAGC 86


RESULT 59
ABT09200
ID   ABT09200 standard; DNA; 696 BP.
XX
AC   ABT09200;
XX
DT   05-DEC-2002  (first entry)
XX
DE   Phase-1 Rat CT gene SEQ ID No 288.
XX
KW   Rat; toxicity study; rat toxic response gene; toxicological response;
KW   drug development; phase-1 rat CT gene; ds.
XX
OS   Rattus sp.
XX
PN   WO200266682-A2.
XX
PD   29-AUG-2002.
XX
PF   29-JAN-2002; 2002WO-US002935.
XX
PR   29-JAN-2001; 2001US-0264933P.
PR   26-JUL-2001; 2001US-0308161P.
XX
PA   (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI   Farris G,  Hicken SH,  Farr SB;
XX
DR   WPI; 2002-674961/72.
XX
PT   Evaluating the toxicity of an agent, useful in drug development or in
PT   determining toxicological responses to a new drug, by determining the
PT   expression of rat toxicologically relevant genes in the test animal in
PT   response to the test agent.
XX
PS   Disclosure; Page 177; 388pp; English.
XX
CC   The invention relates to a method used for evaluating the toxicity of an
CC   agent comprising determining the expression of a rat toxic response
CC   gene(s) in the test animal in response to the agent. The method is useful
CC   in drug development, particularly for conducting toxicity studies and
CC   analysis before a new drug or compound is approved for human consumption
CC   or use. The method is also useful in determining toxicological responses
CC   to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC   gene of the invention
XX
SQ   Sequence 696 BP; 216 A; 147 C; 184 G; 146 T; 0 U; 3 Other;

  Query Match             4.8%;  Score 63.2;  DB 6;  Length 696;
  Best Local Similarity   84.5%;  Pred. No. 1e-06;
  Matches   71;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              | ||||| || |     |  || ||
Db         76 GAGAAGCTGCCTTCTCATTGGCTG 99


RESULT 60
ADG30974
ID   ADG30974 standard; DNA; 696 BP.
XX
AC   ADG30974;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Liver toxicity predictive DNA 154.
XX
KW   liver toxicity; Norway rat; 24 hour combo All; ds.
XX
OS   Unidentified.
XX
PN   WO2003085083-A2.
XX
PD   16-OCT-2003.
XX
PF   01-APR-2003; 2003WO-US010141.
XX
PR   01-APR-2002; 2002US-0369287P.
XX
PA   (PHAS-) PHASE 1 MOLECULAR TOXICOLOGY.
XX
PI   Kier L,  Nolan TD,  Sankar U,  Derbel M;
XX
DR   WPI; 2003-804300/75.
XX
PT   Predicting the liver toxicity of an agent to an individual by using the
PT   test expression profile with a set of reference expression profiles in a
PT   Predictive Model to determine whether the agent will reduce liver
PT   toxicity in the individual.
XX
PS   Claim 2; Page 178; 379pp; English.
XX
CC   The invention relates to a novel method for predicting the liver toxicity
CC   of an agent to an individual comprising obtaining a biological sample
CC   from an individual treated with the agent, measuring the expression of
CC   one or more liver toxicity predictive genes in the sample and using the
CC   test expression profile with a set of reference expression profiles in a
CC   predictive model to determine whether the agent will reduce liver
CC   toxicity in the individual. The method of the invention may be useful for
CC   predicting the liver toxicity of an agent to an individual. The current
```

```
CC   sequence is that of the liver toxicity predictive DNA of the invention
CC   which represents a 24 hour combo All gene.
XX
SQ   Sequence 696 BP; 216 A; 147 C; 184 G; 146 T; 0 U; 3 Other;

  Query Match             4.8%;  Score 63.2;  DB 10;  Length 696;
  Best Local Similarity   84.5%;  Pred. No. 1e-06;
  Matches   71;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              | ||||| || |    |  || ||
Db         76 GAGAAGCTGCCTTCTCATTGGCTG 99


RESULT 61
ADG45593
ID   ADG45593 standard; DNA; 696 BP.
XX
AC   ADG45593;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Liver inflammatory predictive gene related DNA sequence.
XX
KW   liver; liver toxicity; liver toxicity predictive gene;
KW   liver inflammation predictive gene; inflammation; gene; ds; rat.
XX
OS   Rattus norvegicus.
XX
PN   WO2003095624-A2.
XX
PD   20-NOV-2003.
XX
PF   09-MAY-2003; 2003WO-US014832.
XX
PR   10-MAY-2002; 2002US-0379831P.
XX
PA   (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI   Kier L,  Nolan TD,  Sankar U,  Derbel M;
XX
DR   WPI; 2004-042456/04.
XX
PT   Predicting the liver toxicity in an individual to an agent by measuring
PT   the expression of one or more liver toxicity predictive genes in the
PT   sample from the individual treated with the agent.
XX
PS   Claim 7; Fig 4; 336pp; English.
XX
CC   The present invention describes a method for predicting the liver
CC   toxicity in an individual to an agent. The method comprises: (a)
CC   obtaining a biological sample from the individual treated with the agent;
CC   (b) measuring the expression of one or more liver toxicity predictive
CC   genes in the sample; and (c) using the test expression profile with a set
CC   of reference expression profiles in a Predictive Model to determine
CC   whether the agent will induce liver toxicity in the individual. Also
CC   described: (1) predicting the liver toxicity of an agent using an in
CC   vitro system; (2) predicting the liver toxicity in a biological sample
CC   from an individual, in vitro cell cultures or explants to an agent via a
CC   programmable machine; (3) a computer program product for enabling a
CC   computer to perform Predictive Model analysis for liver toxicity on a
CC   biological sample from an individual, in vitro cell cultures or explants
CC   to an agent; (4) a computer system adopted to predict liver toxicity in a
CC   biological system from an individual, in vitro cell cultures or explants
CC   to an agent; (5) a computer program product for predicting liver toxicity
CC   from a test sample expression profile; (6) mining genes predictive of
CC   liver toxicity; (7) an integrated system for predicting liver toxicity;
CC   and (8) identifying one or more liver inflammation predictive genes. The
CC   method is useful for predicting the liver toxicity in an individual to an
CC   agent. The present sequence represents a liver toxicity predictive gene
CC   sequence that represents a 24 hour combo 3 gene, which is specifically
CC   claimed in the exemplification of the present invention.
XX
SQ   Sequence 696 BP; 216 A; 147 C; 184 G; 146 T; 0 U; 3 Other;

  Query Match             4.8%;  Score 63.2;  DB 12;  Length 696;
  Best Local Similarity   84.5%;  Pred. No. 1e-06;
  Matches   71;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              | ||||| || |    |  || ||
Db         76 GAGAAGCTGCCTTCTCATTGGCTG 99


RESULT 62
ADR91217
ID   ADR91217 standard; cDNA; 696 BP.
XX
AC   ADR91217;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Spleen necrosis predictive cDNA sequence, SEQ ID No 173.
XX
KW   altered expression; toxic response; spleen; toxicity; lymphoid; gene; ss.
XX
OS   Unidentified.
XX
PN   WO2004083402-A2.
XX
PD   30-SEP-2004.
XX
PF   17-MAR-2004; 2004WO-US008371.
XX
PR   17-MAR-2003; 2003US-0455443P.
XX
PA   (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI   Sankar U,  Kier L,  Derbel M,  Nolan T;
XX
DR   WPI; 2004-691048/67.
XX
PT   New composition comprises cDNAs useful for detecting altered expression
PT   of genes in a toxic response of the spleen or for predicting toxic
PT   responses to one or more agents including lymphoid tissue types or other
PT   species.
XX
PS   Claim 1; SEQ ID NO 173; 249pp; English.
XX
CC   The invention relates to a novel composition comprising cDNAs for use in
CC   detecting the altered expression of genes in a toxic response of the
CC   spleen, where the cDNAs comprises 50-816 base pairs (ADR91045-ADR91348)
CC   or their complete complements. The invention further comprises:
CC   monitoring the treatment of compound toxicity in a sample; predicting the
CC   spleen toxicity in an individual to an agent; predicting the spleen
CC   toxicity of an agent using an in vitro system; a computer program product
CC   for predicting spleen toxicity from an expression profile of nucleic
CC   acids from a sample under test, comprising a computer readable medium
CC   bearing an encrypted training data set, encrypted lists of genes selected
CC   from the cDNAs, and a predictive model for causing a general purpose
CC   computer to predict the spleen toxicity of the sample based upon the
CC   training data set, the list of genes selected from the cDNAs, and the
CC   expression profile of nucleic acids from the sample; and an integrated
CC   system for predicting spleen toxicity, comprising means for measuring
CC   gene expression profiles of spleen predictive genes from samples exposed
CC   to the test agent and a computer system operably linked to the means that
CC   is capable of implementing a predictive model. The composition comprising
CC   cDNAs is useful for detecting altered expression of genes in a toxic
```

```
CC   response of the spleen or for predicting toxic responses to one or more
CC   agents including lymphoid tissue types or other species. The predictive
CC   genes and models of the invention are useful for identifying and
CC   evaluating various in vitro systems that can be used to accurately
CC   predict in vivo toxicity. This polynucleotide sequence represents one of
CC   the genes predictive for spleen necrosis of the invention.
XX
SQ   Sequence 696 BP; 216 A; 147 C; 184 G; 146 T; 0 U; 3 Other;

  Query Match             4.8%;  Score 63.2;  DB 13;  Length 696;
  Best Local Similarity   84.5%;  Pred. No. 1e-06;
  Matches   71;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              | ||||| || |    |  || ||
Db         76 GAGAAGCTGCCTTCTCATTGGCTG 99


RESULT 63
ABT09004
ID   ABT09004 standard; DNA; 704 BP.
XX
AC   ABT09004;
XX
DT   05-DEC-2002  (first entry)
XX
DE   Phase-1 Rat CT gene SEQ ID No 92.
XX
KW   Rat; toxicity study; rat toxic response gene; toxicological response;
KW   drug development; phase-1 rat CT gene; ds.
XX
OS   Rattus sp.
XX
PN   WO200266682-A2.
XX
PD   29-AUG-2002.
XX
PF   29-JAN-2002; 2002WO-US002935.
XX
PR   29-JAN-2001; 2001US-0264933P.
PR   26-JUL-2001; 2001US-0308161P.
XX
PA   (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI   Farris G,  Hicken SH,  Farr SB;
XX
DR   WPI; 2002-674961/72.
XX
PT   Evaluating the toxicity of an agent, useful in drug development or in
PT   determining toxicological responses to a new drug, by determining the
PT   expression of rat toxicologically relevant genes in the test animal in
PT   response to the test agent.
XX
PS   Disclosure; Page 128; 388pp; English.
XX
CC   The invention relates to a method used for evaluating the toxicity of an
CC   agent comprising determining the expression of a rat toxic response
CC   gene(s) in the test animal in response to the agent. The method is useful
CC   in drug development, particularly for conducting toxicity studies and
CC   analysis before a new drug or compound is approved for human consumption
CC   or use. The method is also useful in determining toxicological responses
CC   to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC   gene of the invention
XX
SQ   Sequence 704 BP; 170 A; 195 C; 168 G; 167 T; 0 U; 4 Other;

  Query Match             4.8%;  Score 63.2;  DB 6;  Length 704;
  Best Local Similarity   84.5%;  Pred. No. 1e-06;
  Matches   71;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAT 77

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              ||| |   ||| || |||   |||
Db         78 GGGTATGTGCGGGATGAGACAGTG 101


RESULT 64
ADG30925
ID   ADG30925 standard; DNA; 704 BP.
XX
AC   ADG30925;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Liver toxicity predictive DNA 105.
XX
KW   liver toxicity; Norway rat; 24 hour combo All; ds.
XX
OS   Unidentified.
XX
PN   WO2003085083-A2.
XX
PD   16-OCT-2003.
XX
PF   01-APR-2003; 2003WO-US010141.
XX
PR   01-APR-2002; 2002US-0369287P.
XX
PA   (PHAS-) PHASE 1 MOLECULAR TOXICOLOGY.
XX
PI   Kier L,  Nolan TD,  Sankar U,  Derbel M;
XX
DR   WPI; 2003-804300/75.
XX
PT   Predicting the liver toxicity of an agent to an individual by using the
PT   test expression profile with a set of reference expression profiles in a
PT   Predictive Model to determine whether the agent will reduce liver
PT   toxicity in the individual.
XX
PS   Claim 2; Page 168; 379pp; English.
XX
CC   The invention relates to a novel method for predicting the liver toxicity
CC   of an agent to an individual comprising obtaining a biological sample
CC   from an individual treated with the agent, measuring the expression of
CC   one or more liver toxicity predictive genes in the sample and using the
CC   test expression profile with a set of reference expression profiles in a
CC   predictive model to determine whether the agent will reduce liver
CC   toxicity in the individual. The method of the invention may be useful for
CC   predicting the liver toxicity of an agent to an individual. The current
CC   sequence is that of the liver toxicity predictive DNA of the invention
CC   which represents a 24 hour combo All gene.
XX
SQ   Sequence 704 BP; 170 A; 195 C; 168 G; 167 T; 0 U; 4 Other;

  Query Match             4.8%;  Score 63.2;  DB 10;  Length 704;
  Best Local Similarity   84.5%;  Pred. No. 1e-06;
  Matches   71;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAT 77

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              ||| |   ||| || |||   |||
Db         78 GGGTATGTGCGGGATGAGACAGTG 101


RESULT 65
ADR91162
```

ID ADR91162 standard; cDNA; 704 BP.
XX
AC ADR91162;
XX
DT 16-DEC-2004 (first entry)
XX
DE Spleen necrosis predictive cDNA sequence, SEQ ID No 118.
XX
KW altered expression; toxic response; spleen; toxicity; lymphoid; gene; ss.
XX
OS Unidentified.
XX
PN WO2004083402-A2.
XX
PD 30-SEP-2004.
XX
PF 17-MAR-2004; 2004WO-US008371.
XX
PR 17-MAR-2003; 2003US-0455443P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Sankar U, Kier L, Derbel M, Nolan T;
XX
DR WPI; 2004-691048/67.
XX
PT New composition comprises cDNAs useful for detecting altered expression
PT of genes in a toxic response of the spleen or for predicting toxic
PT responses to one or more agents including lymphoid tissue types or other
PT species.
XX
PS Claim 1; SEQ ID NO 118; 249pp; English.
XX
CC The invention relates to a novel composition comprising cDNAs for use in
CC detecting the altered expression of genes in a toxic response of the
CC spleen, where the cDNAs comprises 50-816 base pairs (ADR91045-ADR91348)
CC or their complete complements. The invention further comprises:
CC monitoring the treatment of compound toxicity in a sample; predicting the
CC spleen toxicity in an individual to an agent; predicting the spleen
CC toxicity of an agent using an in vitro system; a computer program product
CC for predicting spleen toxicity from an expression profile of nucleic
CC acids from a sample under test, comprising a computer readable medium
CC bearing an encrypted training data set, encrypted lists of genes selected
CC from the cDNAs, and a predictive model for causing a general purpose
CC computer to predict the spleen toxicity of the sample based upon the
CC training data set, the list of genes selected from the cDNAs, and the
CC expression profile of nucleic acids from the sample; and an integrated
CC system for predicting spleen toxicity, comprising means for measuring
CC gene expression profiles of spleen predictive genes from samples exposed
CC to the test agent and a computer system operably linked to the means that
CC is capable of implementing a predictive model. The composition comprising
CC cDNAs is useful for detecting altered expression of genes in a toxic
CC response of the spleen or for predicting toxic responses to one or more
CC agents including lymphoid tissue types or other species. The predictive
CC genes and models of the invention are useful for identifying and
CC evaluating various in vitro systems that can be used to accurately
CC predict in vivo toxicity. This polynucleotide sequence represents one of
CC the genes predictive for spleen necrosis of the invention.
XX
SQ Sequence 704 BP; 170 A; 195 C; 168 G; 167 T; 0 U; 4 Other;

Query Match 4.8%; Score 63.2; DB 13; Length 704;
Best Local Similarity 84.5%; Pred. No. 1e-06;
Matches 71; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
Db 18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAT 77

Qy 77 GGGAAGCAGCGTGAAGAGCGGGTG 100
Db 78 GGGTATGTGCGGGATGAGACAGTG 101

RESULT 66
ABT09153
ID ABT09153 standard; DNA; 722 BP.
XX
AC ABT09153;
XX
DT 05-DEC-2002 (first entry)
XX
DE Phase-1 Rat CT gene SEQ ID No 241.
XX
KW Rat; toxicity study; rat toxic response gene; toxicological response;
KW drug development; phase-1 rat CT gene; ds.
XX
OS Rattus sp.
XX
PN WO200266682-A2.
XX
PD 29-AUG-2002.
XX
PF 29-JAN-2002; 2002WO-US002935.
XX
PR 29-JAN-2001; 2001US-0264933P.
PR 26-JUL-2001; 2001US-0308161P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Farris G, Hicken SH, Farr SB;
XX
DR WPI; 2002-674961/72.
XX
PT Evaluating the toxicity of an agent, useful in drug development or in
PT determining toxicological responses to a new drug, by determining the
PT expression of rat toxicologically relevant genes in the test animal in
PT response to the test agent.
XX
PS Disclosure; Page 165; 388pp; English.
XX
CC The invention relates to a method used for evaluating the toxicity of an
CC agent comprising determining the expression of a rat toxic response
CC gene(s) in the test animal in response to the agent. The method is useful
CC in drug development, particularly for conducting toxicity studies and
CC analysis before a new drug or compound is approved for human consumption
CC or use. The method is also useful in determining toxicological responses
CC to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC gene of the invention
XX
SQ Sequence 722 BP; 184 A; 180 C; 205 G; 150 T; 0 U; 3 Other;

Query Match 4.8%; Score 62.8; DB 6; Length 722;
Best Local Similarity 69.7%; Pred. No. 1.4e-06;
Matches 85; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
Db 17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGG 76

Qy 77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
Db 77 CCGCCGCCGTCTCTGCAGCCCACGGGTCCCTGGGCCGTTCCGCCGCCCGCGTGCGGCCCC 136

Qy 137 AG 138
Db 137 GG 138

RESULT 67
ADG45729
ID ADG45729 standard; DNA; 722 BP.
XX
AC ADG45729;
XX
DT 26-FEB-2004 (first entry)

```
XX
DE   Liver inflammatory predictive gene related DNA sequence.
XX
KW   liver; liver toxicity; liver toxicity predictive gene;
KW   liver inflammation predictive gene; inflammation; gene; ds; rat.
XX
OS   Rattus norvegicus.
XX
PN   WO2003095624-A2.
XX
PD   20-NOV-2003.
XX
PF   09-MAY-2003; 2003WO-US014832.
XX
PR   10-MAY-2002; 2002US-0379831P.
XX
PA   (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI   Kier L,  Nolan TD,  Sankar U,  Derbel M;
XX
DR   WPI; 2004-042456/04.
XX
PT   Predicting the liver toxicity in an individual to an agent by measuring
PT   the expression of one or more liver toxicity predictive genes in the
PT   sample from the individual treated with the agent.
XX
PS   Claim 7; Fig 4; 336pp; English.
XX
CC   The present invention describes a method for predicting the liver
CC   toxicity in an individual to an agent. The method comprises: (a)
CC   obtaining a biological sample from the individual treated with the agent;
CC   (b) measuring the expression of one or more liver toxicity predictive
CC   genes in the sample; and (c) using the test expression profile with a set
CC   of reference expression profiles in a Predictive Model to determine
CC   whether the agent will induce liver toxicity in the individual. Also
CC   described: (1) predicting the liver toxicity of an agent using an in
CC   vitro system; (2) predicting the liver toxicity in a biological sample
CC   from an individual, in vitro cell cultures or explants to an agent via a
CC   programmable machine; (3) a computer program product for enabling a
CC   computer to perform Predictive Model analysis for liver toxicity on a
CC   biological sample from an individual, in vitro cell cultures or explants
CC   to an agent; (4) a computer system adopted to predict liver toxicity in a
CC   biological system from an individual, in vitro cell cultures or explants
CC   to an agent; (5) a computer program product for predicting liver toxicity
CC   from a test sample expression profile; (6) mining genes predictive of
CC   liver toxicity; (7) an integrated system for predicting liver toxicity;
CC   and (8) identifying one or more liver inflammation predictive genes. The
CC   method is useful for predicting the liver toxicity in an individual to an
CC   agent. The present sequence represents a liver toxicity predictive gene
CC   sequence that represents a 24 hour combo 3 gene, which is specifically
CC   claimed in the exemplification of the present invention.
XX
SQ   Sequence 722 BP; 184 A; 180 C; 205 G; 150 T; 0 U; 3 Other;

  Query Match             4.8%;  Score 62.8;  DB 12;  Length 722;
  Best Local Similarity   69.7%;  Pred. No. 1.4e-06;
  Matches   85;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGG 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
                |  || |  |    |||    |  |    ||| |  | || ||   ||||  |||  |
Db         77 CCGCCGCCGTCTCTGCAGCCCACGGGTCCCTGGGCCGTTCCGCCGCCCGCGTGCGGCCCC 136

Qy        137 AG 138
               |
Db        137 GG 138


RESULT 68
ABT09311
ID   ABT09311 standard; DNA; 630 BP.
XX
AC   ABT09311;
XX
DT   05-DEC-2002  (first entry)
XX
DE   Phase-1 Rat CT gene SEQ ID No 399.
XX
KW   Rat; toxicity study; rat toxic response gene; toxicological response;
KW   drug development; phase-1 rat CT gene; ds.
XX
OS   Rattus sp.
XX
PN   WO200266682-A2.
XX
PD   29-AUG-2002.
XX
PF   29-JAN-2002; 2002WO-US002935.
XX
PR   29-JAN-2001; 2001US-0264933P.
PR   26-JUL-2001; 2001US-0308161P.
XX
PA   (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI   Farris G,  Hicken SH,  Farr SB;
XX
DR   WPI; 2002-674961/72.
XX
PT   Evaluating the toxicity of an agent, useful in drug development or in
PT   determining toxicological responses to a new drug, by determining the
PT   expression of rat toxicologically relevant genes in the test animal in
PT   response to the test agent.
XX
PS   Disclosure; Page 205; 388pp; English.
XX
CC   The invention relates to a method used for evaluating the toxicity of an
CC   agent comprising determining the expression of a rat toxic response
CC   gene(s) in the test animal in response to the agent. The method is useful
CC   in drug development, particularly for conducting toxicity studies and
CC   analysis before a new drug or compound is approved for human consumption
CC   or use. The method is also useful in determining toxicological responses
CC   to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC   gene of the invention
XX
SQ   Sequence 630 BP; 177 A; 150 C; 142 G; 161 T; 0 U; 0 Other;

  Query Match             4.7%;  Score 62.2;  DB 6;  Length 630;
  Best Local Similarity   76.8%;  Pred. No. 1.9e-06;
  Matches   76;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGTG 79

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCT 115
                |  | ||   |  |   |||||||   |||||   ||
Db         80 ATGTTGGAGAAGGCTGCCGGGGTGTTGCTAGGGGCGGCT 118


RESULT 69
ADG45672
ID   ADG45672 standard; DNA; 630 BP.
XX
AC   ADG45672;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Liver inflammatory predictive gene related DNA sequence.
XX
KW   liver; liver toxicity; liver toxicity predictive gene;
KW   liver inflammation predictive gene; inflammation; gene; ds; rat.
XX
OS   Rattus norvegicus.
```

XX
PN WO2003095624-A2.
XX
PD 20-NOV-2003.
XX
PF 09-MAY-2003; 2003WO-US014832.
XX
PR 10-MAY-2002; 2002US-0379831P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Kier L, Nolan TD, Sankar U, Derbel M;
XX
DR WPI; 2004-042456/04.
XX
PT Predicting the liver toxicity in an individual to an agent by measuring
PT the expression of one or more liver toxicity predictive genes in the
PT sample from the individual treated with the agent.
XX
PS Claim 7; Fig 4; 336pp; English.
XX
CC The present invention describes a method for predicting the liver
CC toxicity in an individual to an agent. The method comprises: (a)
CC obtaining a biological sample from the individual treated with the agent;
CC (b) measuring the expression of one or more liver toxicity predictive
CC genes in the sample; and (c) using the test expression profile with a set
CC of reference expression profiles in a Predictive Model to determine
CC whether the agent will induce liver toxicity in the individual. Also
CC described: (1) predicting the liver toxicity of an agent using an in
CC vitro system; (2) predicting the liver toxicity in a biological sample
CC from an individual, in vitro cell cultures or explants to an agent via a
CC programmable machine; (3) a computer program product for enabling a
CC computer to perform Predictive Model analysis for liver toxicity on a
CC biological sample from an individual, in vitro cell cultures or explants
CC to an agent; (4) a computer system adopted to predict liver toxicity in a
CC biological system from an individual, in vitro cell cultures or explants
CC to an agent; (5) a computer program product for predicting liver toxicity
CC from a test sample expression profile; (6) mining genes predictive of
CC liver toxicity; (7) an integrated system for predicting liver toxicity;
CC and (8) identifying one or more liver inflammation predictive genes. The
CC method is useful for predicting the liver toxicity in an individual to an
CC agent. The present sequence represents a liver toxicity predictive gene
CC sequence that represents a 24 hour combo 3 gene, which is specifically
CC claimed in the exemplification of the present invention.
XX
SQ Sequence 630 BP; 177 A; 150 C; 142 G; 161 T; 0 U; 0 Other;

Query Match 4.7%; Score 62.2; DB 12; Length 630;
Best Local Similarity 76.8%; Pred. No. 1.9e-06;
Matches 76; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
Db 20 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGTG 79

Qy 77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCT 115
Db 80 ATGTTGGAGAAGGCTGCCGGGGTGTTGCTAGGGGCGGCT 118

RESULT 70
ADR91268
ID ADR91268 standard; cDNA; 630 BP.
XX
AC ADR91268;
XX
DT 16-DEC-2004 (first entry)
XX
DE Spleen necrosis predictive cDNA sequence, SEQ ID No 224.
XX
KW altered expression; toxic response; spleen; toxicity; lymphoid; gene; ss.
XX
OS Unidentified.

XX
PN WO2004083402-A2.
XX
PD 30-SEP-2004.
XX
PF 17-MAR-2004; 2004WO-US008371.
XX
PR 17-MAR-2003; 2003US-0455443P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Sankar U, Kier L, Derbel M, Nolan T;
XX
DR WPI; 2004-691048/67.
XX
PT New composition comprises cDNAs useful for detecting altered expression
PT of genes in a toxic response of the spleen or for predicting toxic
PT responses to one or more agents including lymphoid tissue types or other
PT species.
XX
PS Claim 1; SEQ ID NO 224; 249pp; English.
XX
CC The invention relates to a novel composition comprising cDNAs for use in
CC detecting the altered expression of genes in a toxic response of the
CC spleen, where the cDNAs comprises 50-816 base pairs (ADR91045-ADR91348)
CC or their complete complements. The invention further comprises:
CC monitoring the treatment of compound toxicity in a sample; predicting the
CC spleen toxicity in an individual to an agent; predicting the spleen
CC toxicity of an agent using an in vitro system; a computer program product
CC for predicting spleen toxicity from an expression profile of nucleic
CC acids from a sample under test, comprising a computer readable medium
CC bearing an encrypted training data set, encrypted lists of genes selected
CC from the cDNAs, and a predictive model for causing a general purpose
CC computer to predict the spleen toxicity of the sample based upon the
CC training data set, the list of genes selected from the cDNAs, and the
CC expression profile of nucleic acids from the sample; and an integrated
CC system for predicting spleen toxicity, comprising means for measuring
CC gene expression profiles of spleen predictive genes from samples exposed
CC to the test agent and a computer system operably linked to the means that
CC is capable of implementing a predictive model. The composition comprising
CC cDNAs is useful for detecting altered expression of genes in a toxic
CC response of the spleen or for predicting toxic responses to one or more
CC agents including lymphoid tissue types or other species. The predictive
CC genes and models of the invention are useful for identifying and
CC evaluating various in vitro systems that can be used to accurately
CC predict in vivo toxicity. This polynucleotide sequence represents one of
CC the genes predictive for spleen necrosis of the invention.
XX
SQ Sequence 630 BP; 177 A; 150 C; 142 G; 161 T; 0 U; 0 Other;

Query Match 4.7%; Score 62.2; DB 13; Length 630;
Best Local Similarity 76.8%; Pred. No. 1.9e-06;
Matches 76; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
Db 20 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGTG 79

Qy 77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCT 115
Db 80 ATGTTGGAGAAGGCTGCCGGGGTGTTGCTAGGGGCGGCT 118

RESULT 71
AAA38745
ID AAA38745 standard; cDNA; 1549 BP.
XX
AC AAA38745;
XX
DT 14-SEP-2000 (first entry)
XX
DE Murine Hsp-binding protein HspBPM coding sequence.
XX

```
KW   Mouse; heat shock protein binding protein; HspBPM; development;
KW   apoptosis; cellular stress; heart disease; cancer; hypoxia; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             160. .1233
FT                   /*tag= a
FT                   /product= "heat shock protein-binding protein"
XX
PN   WO200031114-A1.
XX
PD   02-JUN-2000.
XX
PF   19-NOV-1999;   99WO-US027651.
XX
PR   20-NOV-1998;   98US-0109351P.
XX
PA   (GUER/) GUERRIERO V.
PA   (RAYN/) RAYNES D A.
XX
PI   Guerriero V,  Raynes DA;
XX
DR   WPI; 2000-400030/34.
DR   P-PSDB; AAY97407.
XX
PT   New polynucleotides encoding human heat-shock protein-binding protein,
PT   HspBP-1 and HspBP-2, useful for investigating the effects of heat shock-
PT   protein regulation.
XX
PS   Disclosure; Page 29-30; 38pp; English.
XX
CC   The present sequence is the coding sequence for murine heat shock protein
CC   -binding protein (HspBPM). Its protein is involved in the regulation of
CC   the heat shock proteins, which are components of the development,
CC   apoptosis and cellular stress pathways. The human homologue is useful not
CC   only for research into these areas, but also for treating disorders such
CC   as heart disease, hypoxia and cancer
XX
SQ   Sequence 1549 BP; 315 A; 458 C; 450 G; 326 T; 0 U; 0 Other;

  Query Match             4.7%;  Score 62.2;  DB 3;  Length 1549;
  Best Local Similarity   95.5%;  Pred. No. 3e-06;
  Matches   64;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGG 78
              || |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db          2 TTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCCGG 61

Qy         79 GAAGCAG 85
              |||||||
Db         62 CAAGCAG 68


RESULT 72
ABT08958
ID   ABT08958 standard; DNA; 701 BP.
XX
AC   ABT08958;
XX
DT   05-DEC-2002  (first entry)
XX
DE   Phase-1 Rat CT gene SEQ ID No 46.
XX
KW   Rat; toxicity study; rat toxic response gene; toxicological response;
KW   drug development; phase-1 rat CT gene; ds.
XX
OS   Rattus sp.
XX
PN   WO200266682-A2.
XX
PD   29-AUG-2002.
XX
PF   29-JAN-2002; 2002WO-US002935.
XX
PR   29-JAN-2001; 2001US-0264933P.
PR   26-JUL-2001; 2001US-0308161P.
XX
PA   (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI   Farris G,  Hicken SH,  Farr SB;
XX
DR   WPI; 2002-674961/72.
XX
PT   Evaluating the toxicity of an agent, useful in drug development or in
PT   determining toxicological responses to a new drug, by determining the
PT   expression of rat toxicologically relevant genes in the test animal in
PT   response to the test agent.
XX
PS   Disclosure; Page 117; 388pp; English.
XX
CC   The invention relates to a method used for evaluating the toxicity of an
CC   agent comprising determining the expression of a rat toxic response
CC   gene(s) in the test animal in response to the agent. The method is useful
CC   in drug development, particularly for conducting toxicity studies and
CC   analysis before a new drug or compound is approved for human consumption
CC   or use. The method is also useful in determining toxicological responses
CC   to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC   gene of the invention
XX
SQ   Sequence 701 BP; 165 A; 198 C; 189 G; 147 T; 0 U; 2 Other;

  Query Match             4.7%;  Score 61.6;  DB 6;  Length 701;
  Best Local Similarity   56.4%;  Pred. No. 2.9e-06;
  Matches  115;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
              || |   |          || |   |  |   |   |   |||  |   | | |    |
Db         77 GGAACCATGGCCAGCCCGCTGCGCTCCTTGATGCTACTGCTGGCCGTCCTGGCCGTGGCC 136

Qy        137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
              |  |||   |  |   |||  |   ||        | | ||  ||  || | | || ||
Db        137 TGGGCCGGAACCTCCAGGCCACCCCCGCGATTGTTGGGAGCTCCGCAGGAGGCAGATGCC 196

Qy        197 GCACAGGCGGAGGAGGACCGGGTG 220
                 ||| ||  | | | || | |
Db        197 AGCGAGGAGGGCGTGCAGCGAGCG 220


RESULT 73
ADG30880
ID   ADG30880 standard; DNA; 701 BP.
XX
AC   ADG30880;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Liver toxicity predictive DNA 60.
XX
KW   liver toxicity; Norway rat; 24 hour combo All; ds.
XX
OS   Unidentified.
XX
PN   WO2003085083-A2.
XX
PD   16-OCT-2003.
XX
PF   01-APR-2003; 2003WO-US010141.
XX
PR   01-APR-2002; 2002US-0369287P.
XX
PA   (PHAS-) PHASE 1 MOLECULAR TOXICOLOGY.
```

XX
PI Kier L, Nolan TD, Sankar U, Derbel M;
XX
DR WPI; 2003-804300/75.
XX
PT Predicting the liver toxicity of an agent to an individual by using the
PT test expression profile with a set of reference expression profiles in a
PT Predictive Model to determine whether the agent will reduce liver
PT toxicity in the individual.
XX
PS Claim 2; Page 160; 379pp; English.
XX
CC The invention relates to a novel method for predicting the liver toxicity
CC of an agent to an individual comprising obtaining a biological sample
CC from an individual treated with the agent, measuring the expression of
CC one or more liver toxicity predictive genes in the sample and using the
CC test expression profile with a set of reference expression profiles in a
CC predictive model to determine whether the agent will reduce liver
CC toxicity in the individual. The method of the invention may be useful for
CC predicting the liver toxicity of an agent to an individual. The current
CC sequence is that of the liver toxicity predictive DNA of the invention
CC which represents a 24 hour combo All gene.
XX
SQ Sequence 701 BP; 165 A; 198 C; 189 G; 147 T; 0 U; 2 Other;

Query Match 4.7%; Score 61.6; DB 10; Length 701;
Best Local Similarity 56.4%; Pred. No. 2.9e-06;
Matches 115; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
Db 17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy 77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
Db 77 GGAACCATGGCCAGCCCGCTGCGCTCCTTGATGCTACTGCTGGCCGTCCTGGCCGTGGCC 136

Qy 137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
Db 137 TGGGCCGGAACCTCCAGGCCACCCCCGCGATTGTTGGGAGCTCCGCAGGAGGCAGATGCC 196

Qy 197 GCACAGGCGGAGGAGGACCGGGTG 220
Db 197 AGCGAGGAGGGCGTGCAGCGAGCG 220

RESULT 74
ADG45448
ID ADG45448 standard; DNA; 701 BP.
XX
AC ADG45448;
XX
DT 26-FEB-2004 (first entry)
XX
DE Liver inflammatory predictive gene related DNA sequence.
XX
KW liver; liver toxicity; liver toxicity predictive gene;
KW liver inflammation predictive gene; inflammation; gene; ds; rat.
XX
OS Rattus norvegicus.
XX
PN WO2003095624-A2.
XX
PD 20-NOV-2003.
XX
PF 09-MAY-2003; 2003WO-US014832.
XX
PR 10-MAY-2002; 2002US-0379831P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Kier L, Nolan TD, Sankar U, Derbel M;
XX
DR WPI; 2004-042456/04.
XX
PT Predicting the liver toxicity in an individual to an agent by measuring
PT the expression of one or more liver toxicity predictive genes in the
PT sample from the individual treated with the agent.
XX
PS Claim 7; Fig 4; 336pp; English.
XX
CC The present invention describes a method for predicting the liver
CC toxicity in an individual to an agent. The method comprises: (a)
CC obtaining a biological sample from the individual treated with the agent;
CC (b) measuring the expression of one or more liver toxicity predictive
CC genes in the sample; and (c) using the test expression profile with a set
CC of reference expression profiles in a Predictive Model to determine
CC whether the agent will induce liver toxicity in the individual. Also
CC described: (1) predicting the liver toxicity of an agent using an in
CC vitro system; (2) predicting the liver toxicity in a biological sample
CC from an individual, in vitro cell cultures or explants to an agent via a
CC programmable machine; (3) a computer program product for enabling a
CC computer to perform Predictive Model analysis for liver toxicity on a
CC biological sample from an individual, in vitro cell cultures or explants
CC to an agent; (4) a computer system adopted to predict liver toxicity in a
CC biological system from an individual, in vitro cell cultures or explants
CC to an agent; (5) a computer program product for predicting liver toxicity
CC from a test sample expression profile; (6) mining genes predictive of
CC liver toxicity; (7) an integrated system for predicting liver toxicity;
CC and (8) identifying one or more liver inflammation predictive genes. The
CC method is useful for predicting the liver toxicity in an individual to an
CC agent. The present sequence represents a liver toxicity predictive gene
CC sequence that represents a 24 hour combo 3 gene, which is specifically
CC claimed in the exemplification of the present invention.
XX
SQ Sequence 701 BP; 165 A; 198 C; 189 G; 147 T; 0 U; 2 Other;

Query Match 4.7%; Score 61.6; DB 12; Length 701;
Best Local Similarity 56.4%; Pred. No. 2.9e-06;
Matches 115; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
Db 17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy 77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
Db 77 GGAACCATGGCCAGCCCGCTGCGCTCCTTGATGCTACTGCTGGCCGTCCTGGCCGTGGCC 136

Qy 137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
Db 137 TGGGCCGGAACCTCCAGGCCACCCCCGCGATTGTTGGGAGCTCCGCAGGAGGCAGATGCC 196

Qy 197 GCACAGGCGGAGGAGGACCGGGTG 220
Db 197 AGCGAGGAGGGCGTGCAGCGAGCG 220

RESULT 75
ADH22750
ID ADH22750 standard; DNA; 701 BP.
XX
AC ADH22750;
XX
DT 11-MAR-2004 (first entry)
XX
DE Partial DNA seqeunce of a rat kidney toxicity predictive gene (78).
XX
KW kidney toxicity; toxicology; predictive model; gene expression profile;
KW toxic damage; kidney tubule necrosis; acute renal failure; rat; ds.
XX
OS Rattus sp.
XX
PN WO2003100030-A2.
XX
PD 04-DEC-2003.

XX
PF 27-FEB-2003; 2003WO-US006196.
XX
PR 27-FEB-2002; 2002US-0361128P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Kier L, Nolan TD, Sankar U, Derbel M;
XX
DR WPI; 2004-035137/03.
XX
PT Predicting kidney toxicity in an individual to an agent, useful for
PT predicting toxic responses to one or more agents comprising measuring the
PT expression of one or more kidney toxicity predictive genes.
XX
PS Claim 1; Page 330; 388pp; English.
XX
CC This invention relates to novel isolated kidney toxicity predictive genes
CC and methods of using such genes in the field of toxicology. Specifically,
CC it refers to novel genes that can be used for generating predictive
CC models, which in turn are useful for predicting the in vivo toxic
CC response to one or more agents. The present invention describes obtaining
CC a gene expression profile from a biological sample and using the
CC predictive model to determine whether an agent will induce kidney
CC toxicity in the individual. As such, it can be used to detect any toxic
CC effects that may be manifested as long lasting or chronic consequences
CC including irreversible toxicity or carcinogenesis. Furthermore, the
CC predictive genes can be considered as therapeutic targets for toxic
CC damage or to ameliorate specific disease conditions such as kidney tubule
CC necrosis or acute renal failure, as well as for additional screening
CC assays. This polynucleotide sequence is the partial gene sequence of a
CC rat kidney toxicity predictive gene of the invention.
XX
SQ Sequence 701 BP; 165 A; 198 C; 189 G; 147 T; 0 U; 2 Other;

Query Match 4.7%; Score 61.6; DB 12; Length 701;
Best Local Similarity 56.4%; Pred. No. 2.9e-06;
Matches 115; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
              || |    |        || |   | |    |   |    |||  |   | | |    |
Db         77 GGAACCATGGCCAGCCCGCTGCGCTCCTTGATGCTACTGCTGGCCGTCCTGGCCGTGGCC 136

Qy        137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
                |||   |  |  |||  |   ||      | | ||  ||  || | | || ||
Db        137 TGGGCCGGAACCTCCAGGCCACCCCCGCGATTGTTGGGAGCTCCGCAGGAGGCAGATGCC 196

Qy        197 GCACAGGCGGAGGAGGACCGGGTG 220
                 |||  ||  | | | || | |
Db        197 AGCGAGGAGGGCGTGCAGCGAGCG 220
```

Search completed: March 11, 2006, 06:35:43
Job time : 885 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2006, 06:35:52 ; Search time 6919 Seconds
(without alignments)
10778.813 Million cell updates/sec

Title: US-10-757-745-3
Perfect score: 1312
Sequence: 1 agctattaatgattcgaatt.....tctcaacctttcaggacatc 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1312	100.0	1312	6	BD205491	BD205491 CD40-Inte
	2	1312	100.0	1312	6	AR594295	AR594295 Sequence
	3	1312	100.0	1312	6	AX011601	AX011601 Sequence
	4	1238	94.4	1943	9	MMU251328	AJ251328 Mus muscu
	5	668.6	51.0	1952	8	BC017553	BC017553 Homo sapi
	6	668.2	50.9	1261	8	AF201687	AF201687 Homo sapi
	7	668.2	50.9	1920	6	BD205490	BD205490 CD40-Inte
	8	668.2	50.9	1920	6	AR594294	AR594294 Sequence
	9	668.2	50.9	1920	6	AX011599	AX011599 Sequence
	10	668.2	50.9	1948	6	AR339398	AR339398 Sequence
	11	668.2	50.9	2499	6	E23195	E23195 Topoisomera
	12	666.6	50.8	1921	8	HSA269473	AJ269473 Homo sapi
	13	666.6	50.8	1936	8	AF223469	AF223469 Homo sapi
	14	665	50.7	1898	6	BD157138	BD157138 Primer fo
	15	665	50.7	1898	6	AX878304	AX878304 Sequence
	16	665	50.7	1898	8	AK002168	AK002168 Homo sapi
c	17	648.2	49.4	150344	14	AC152027	AC152027 Dasypus n
	18	585.2	44.6	2151	8	HSA420495	AJ420495 Homo sapi
c	19	490	37.3	1079	6	BD139846	BD139846 Compounds
c	20	490	37.3	1079	6	AR202950	AR202950 Sequence
c	21	490	37.3	1079	6	AR208051	AR208051 Sequence
c	22	490	37.3	1079	6	AX429924	AX429924 Sequence
	23	377.8	28.8	752	6	BD150065	BD150065 Primer fo
	24	377.8	28.8	752	6	AX870003	AX870003 Sequence
	25	357.2	27.2	1236	5	CR761829	CR761829 Xenopus t
	26	354.4	27.0	256608	9	AL589699	AL589699 Mouse DNA
	27	311	23.7	858	6	CQ726600	CQ726600 Sequence
	28	276.6	21.1	1451	5	BC083404	BC083404 Danio rer
	29	274.2	20.9	1455	5	BC097117	BC097117 Danio rer
c	30	273.6	20.9	245394	14	AC125756	AC125756 Rattus no
	31	231.8	17.7	483	6	BD076937	BD076937 5' EST of
	32	227.2	17.3	400	6	CQ663156	CQ663156 Sequence
	33	196	14.9	20938	8	AY613922	AY613922 Homo sapi
	34	196	14.9	102200	8	HS30M3	AL031775 Human DNA
	35	196	14.9	167847	8	CR925830	CR925830 Human DNA
	36	196	14.9	176819	8	CR942205	CR942205 Human DNA
	37	146.2	11.1	201	10	BV203236	BV203236 sqnm21148
c	38	135.6	10.3	179	6	AX339803	AX339803 Sequence
	39	134.4	10.2	176	6	AX351374	AX351374 Sequence
c	40	101.4	7.7	268	6	AR240556	AR240556 Sequence
c	41	95.2	7.3	411	10	HS133N16S	AL110544 H.sapiens
c	42	95	7.2	238494	14	AC095717	AC095717 Rattus no
	43	82.6	6.3	198752	5	BX957279	BX957279 Zebrafish
	44	82.6	6.3	200076	5	BX511258	BX511258 Zebrafish
c	45	76.6	5.8	231985	14	AC156673	AC156673 Bos tauru
	46	75.2	5.7	258	6	BD077628	BD077628 5'EST of
	47	75.2	5.7	258	6	AX939201	AX939201 Sequence
	48	74.6	5.7	444	6	CQ069444	CQ069444 Sequence
	49	74.6	5.7	691	6	CQ078617	CQ078617 Sequence
	50	67.4	5.1	705	6	AX525593	AX525593 Sequence
	51	64.8	4.9	631	6	AX525862	AX525862 Sequence
	52	63.8	4.9	690	6	AX525757	AX525757 Sequence
	53	63.2	4.8	696	6	AX525766	AX525766 Sequence
	54	63.2	4.8	704	6	AX525570	AX525570 Sequence
	55	62.8	4.8	722	6	AX525719	AX525719 Sequence
	56	62.2	4.7	630	6	AX525877	AX525877 Sequence
	57	62.2	4.7	1549	6	AR216294	AR216294 Sequence
	58	62.2	4.7	1549	6	AR680067	AR680067 Sequence
	59	61.6	4.7	701	6	AX525524	AX525524 Sequence
	60	61.6	4.7	718	6	AX525842	AX525842 Sequence
c	61	61.4	4.7	682	6	AX525494	AX525494 Sequence
	62	61.2	4.7	701	6	AX525612	AX525612 Sequence
	63	61.2	4.7	956	6	AX224570	AX224570 Sequence
c	64	60.6	4.6	581	6	AX525594	AX525594 Sequence
	65	60.4	4.6	692	6	AX525588	AX525588 Sequence
	66	60.4	4.6	706	6	AX525704	AX525704 Sequence
	67	60.2	4.6	657	6	AX525660	AX525660 Sequence
	68	60	4.6	669	6	AX525724	AX525724 Sequence
	69	60	4.6	710	6	AX525746	AX525746 Sequence
c	70	59.8	4.6	594	6	AX525583	AX525583 Sequence
	71	59.8	4.6	762	6	AX525652	AX525652 Sequence
	72	59.6	4.5	720	6	AX525581	AX525581 Sequence
	73	59.2	4.5	761	6	AX525860	AX525860 Sequence
	74	59	4.5	686	6	AX525728	AX525728 Sequence
c	75	59	4.5	691	6	AX525495	AX525495 Sequence

ALIGNMENTS

RESULT 1
BD205491
LOCUS BD205491 1312 bp DNA linear PAT 17-JUL-2003
DEFINITION CD40-Interacting and TRAF-interacting protein.
ACCESSION BD205491
VERSION BD205491.1 GI:33015261
KEYWORDS JP 2002512796-A/2.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

```
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting protein
  JOURNAL   Patent: JP 2002512796-A 2 08-MAY-2002;
            VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT     OS   Mus musculus (mouse)
            PN   JP 2002512796-A/2
            PD   08-MAY-2002
            PF   28-APR-1999 JP 2000546003
            PR   29-APR-1998 EP   98201392.2
            PI   STEFAN MARIA CHRISTIAAN PYPE,
            PI   JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI   DANNY
            FRANCOIS EVELINE HUYLEBROECK
            PC   C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
              A61P37/02,
            PC   A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
            PC   G01N33/566,
            PC   C12N15/00,A61K37/02
            CC   CD40-Interacting and TRAF-interacting protein. FH   Key
                 Location/Qualifiers
            FT   CDS             (122). .(1234).
FEATURES              Location/Qualifiers
     source           1. .1312
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:10090"
ORIGIN

  Query Match           100.0%;  Score 1312;  DB 6;  Length 1312;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-266;
  Matches 1312;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60

Qy         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120

Qy        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180

Qy        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240

Qy        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300

Qy        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360

Qy        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420

Qy        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480

Qy        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540

Qy        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600

Qy        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660

Qy        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720

Qy        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780

Qy        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840

Qy        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900

Qy        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960

Qy        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020

Qy       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080

Qy       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140

Qy       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200

Qy       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260

Qy       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312


RESULT 2
AR594295
LOCUS       AR594295               1312 bp    DNA     linear   PAT 15-DEC-2004
DEFINITION  Sequence 3 from patent US 6812203.
ACCESSION   AR594295
VERSION     AR594295.1  GI:56643901
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1312)
  AUTHORS   Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting proteins
  JOURNAL   Patent: US 6812203-A 3 02-NOV-2004;
            Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
            Zwijnaarde;
            WOX;
FEATURES              Location/Qualifiers
     source           1. .1312
                      /organism="unknown"
                      /mol_type="genomic DNA"
ORIGIN

  Query Match           100.0%;  Score 1312;  DB 6;  Length 1312;
```

```
Best Local Similarity   100.0%;  Pred. No. 2.4e-266;
Matches 1312;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60

Qy         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120

Qy        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180

Qy        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240

Qy        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300

Qy        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360

Qy        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420

Qy        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480

Qy        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540

Qy        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600

Qy        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660

Qy        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720

Qy        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780

Qy        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840

Qy        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900

Qy        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960

Qy        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020

Qy       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080

Qy       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140

Qy       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200

Qy       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260

Qy       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
```

RESULT 3
AX011601

```
LOCUS       AX011601                1312 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 3 from Patent WO9955859.
ACCESSION   AX011601
VERSION     AX011601.1  GI:9998125
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 3 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1312
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
     CDS             122. .1234
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07556.1"
                     /db_xref="GI:9998126"
                     /translation="MASGSSSDAAEPAGPAGRAASAPEAAQAEEDRVKRRRLQCLGFA
                     LVGGCDPTMVPSVLRENDWQTQKALSAYFELPENDQGWPRQPPTSFKSEAYVDLTNED
                     ANDTTILEASPSGTPLEDSSTISFITWNIDGLDGCNLPERARGVCSCLALYSPDVVFL
                     QEVIPPYCAYLKKRAASYTIITGNEEGYFTAILLKKGRVKFKSQEIIPFPNTKMMRNL
                     LCVNVSLGGNEFCLMTSHLESTREHSAERIRQLKTVLGKMQEAPDSTTVIFAGDTNLR
                     DQEVIKCGGLPDNVFDAWEFLGKPKHCQYTWDTKANNNLRIPAAYKHRFDRIFFRAEE
                     GHLIPQSLDLVGLEKLDCGRFPSDHWGLLCTLNVVL"
ORIGIN
```

```
  Query Match           100.0%;  Score 1312;  DB 6;  Length 1312;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-266;
  Matches 1312;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCTATTAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60

Qy         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATTCGGCACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGC 120

Qy        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGC 180
```

```
Qy       181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240
Db       181 GTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTG 240

Qy       241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300
Db       241 CCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGA 300

Qy       301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360
Db       301 GAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360

Qy       361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420
Db       361 AGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAA 420

Qy       421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480
Db       421 CGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480

Qy       481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540
Db       481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540

Qy       541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600
Db       541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600

Qy       601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660
Db       601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660

Qy       661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720
Db       661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720

Qy       721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780
Db       721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780

Qy       781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840
Db       781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840

Qy       841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900
Db       841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900

Qy       901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960
Db       901 GGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGT 960

Qy       961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020
Db       961 TATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACC 1020

Qy      1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080
Db      1021 TAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGC 1080

Qy      1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140
Db      1081 TTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCA 1140

Qy      1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200
Db      1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200

Qy      1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260
Db      1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACAC 1260

Qy      1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
Db      1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
```

```
RESULT 4
MMU251328
LOCUS       MMU251328              1943 bp    mRNA    linear   ROD 27-MAY-2000
DEFINITION  Mus musculus mRNA for putative TRAF and TNF receptor associated
            protein (TTRAP gene).
ACCESSION   AJ251328
VERSION     AJ251328.1  GI:8247282
KEYWORDS    ttrap gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van
            Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,
            Huylebroeck,D. and Remacle,J.E.
  TITLE     TTRAP, a novel protein that associates with CD40, tumor necrosis
            factor (TNF) receptor-75 and TNF receptor-associated factors
            (TRAFs), and that inhibits nuclear factor-kappa B activation
  JOURNAL   J. Biol. Chem. 275 (24), 18586-18593 (2000)
   PUBMED   10764746
REFERENCE   2  (bases 1 to 1943)
  AUTHORS   Pype,S.M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-NOV-1999) Pype S.M.C., VIB07, Department Cell Growth,
            Differentiation and Developme, Flanders Interuniversity Institute
            for Biotechnology, K.U.Leuven, Campus Gasthuisberg, Herestraat 49,
            3000 Leuven, BELGIUM
COMMENT     Related sequence: AJ269473.
FEATURES             Location/Qualifiers
     source          1. .1943
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /note="IMAGE cDNA EST clone 1262914"
     gene            1. .1943
                     /gene="TTRAP"
     CDS             48. .1160
                     /gene="TTRAP"
                     /codon_start=1
                     /product="putative TRAF and TNF receptor associated
                     protein"
                     /protein_id="CAB92971.1"
                     /db_xref="GI:8247283"
                     /db_xref="GOA:Q9JJX7"
                     /db_xref="UniProt/TrEMBL:Q9JJX7"
                     /translation="MASGSSSDAAEPAGPAGRAASAPEAAQAEEDRVKRRRLQCLGFA
                     LVGGCDPTMVPSVLRENDWQTQKALSAYFELPENDQGWPRQPPTSFKSEAYVDLTNED
                     ANDTTILEASPSGTPLEDSSTISFITWNIDGLDGCNLPERARGVCSCLALYSPDVVFL
                     QEVIPPYCAYLKKRAASYTIITGNEEGYFTAILLKKGRVKFKSQEIIPFPNTKMMRNL
                     LCVNVSLGGNEFCLMTSHLESTREHSAERIRQLKTVLGKMQEAPDSTTVIFAGDTNLR
                     DQEVIKCGGLPDNVFDAWEFLGKPKHCQYTWDTKANNNLRIPAAYKHRFDRIFFRAEE
                     GHLIPQSLDLVGLEKLDCGRFPSDHWGLLCTLNVVL"
ORIGIN
```

```
  Query Match             94.4%;  Score 1238;  DB 9;  Length 1943;
  Best Local Similarity   100.0%;  Pred. No. 1e-250;
  Matches 1238;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          75 GCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCA 134
Db           1 GCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCA 60

Qy         135 GCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAG 194
Db          61 GCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAG 120

Qy         195 CAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGT 254
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGT 180

Qy        255 TGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGA 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGA 240

Qy        315 CGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCC 300

Qy        375 AGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATG 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATG 360

Qy        435 ATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTT 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTT 420

Qy        495 CTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAG 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAG 480

Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 540

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 600

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 660

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 720

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 840

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 900

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 960

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1020

Qy       1095 TTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTG 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTG 1080

Qy       1155 TTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCA 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCA 1140

Qy       1215 CCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACTA 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACTA 1200

Qy       1275 GTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              ||||||||||||||||||||||||||||||||||||||
Db       1201 GTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1238

RESULT 5
BC017553
LOCUS       BC017553                1952 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein, mRNA (cDNA
            clone MGC:9099 IMAGE:3920790), complete cds.
ACCESSION   BC017553
VERSION     BC017553.2  GI:34782842
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1952)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1952)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Sep 16, 2003 this sequence version replaced gi:17028464.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 15 Row: m Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 23510347.
FEATURES             Location/Qualifiers
     source          1. .1952
                     /organism="Homo sapiens"
```

```
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:9099 IMAGE:3920790"
                     /tissue_type="Skin, melanotic melanoma."
                     /clone_lib="NIH_MGC_72"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1952
                     /gene="TTRAP"
                     /note="synonyms: AD022, EAP2, MGC9099, dJ30M3.3"
                     /db_xref="GeneID:51567"
                     /db_xref="MIM:605764"
     CDS             38. .1126
                     /gene="TTRAP"
                     /codon_start=1
                     /product="TRAF and TNF receptor-associated protein"
                     /protein_id="AAH17553.1"
                     /db_xref="GI:17028465"
                     /db_xref="GeneID:51567"
                     /db_xref="MIM:605764"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

  Query Match           51.0%;  Score 668.6;  DB 8;  Length 1952;
  Best Local Similarity 77.0%;  Pred. No. 1.7e-130;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     |||||| | |||||
Db         54 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 113

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        114 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 173

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        174 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 233

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   |  | |   |||| |||   ||   ||   || || || |||||||||
Db        234 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 293

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||| |   |
Db        294 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 353

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||
Db        354 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 413

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | ||||||||||||||||||||||||| | ||||||||||| || ||||
Db        414 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 473

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        474 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 533

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| ||||||||||||||||||||||| | |||||||
Db        534 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 593

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        594 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 653

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   ||||||||  ||||||||||||||||||
Db        654 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 713

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||  ||| |||||||||||| |  || |||||||  ||| |
Db        714 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 773

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        774 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 833

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || |||||||||||||||  ||||| || | |||| ||||||
Db        834 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 893

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||  ||   ||    || ||
Db        894 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 953

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||| | ||||||||||||||||||||||||      |||||||
Db        954 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1013

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db       1014 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1073

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1074 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1133

Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1134 TTCAAGTGTGGGTTT 1148
```

```
RESULT 6
AF201687
LOCUS       AF201687                1261 bp    mRNA    linear   PRI 16-MAY-2003
DEFINITION  Homo sapiens ETS1-associated protein 2 (EAP2) mRNA, complete cds.
ACCESSION   AF201687
VERSION     AF201687.1  GI:11493668
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1261)
  AUTHORS   Pei,H., Yordy,J.S., Leng,Q., Zhao,Q., Watson,D.K. and Li,R.
  TITLE     EAPII interacts with ETS1 and modulates its transcriptional
            function
  JOURNAL   Oncogene 22 (18), 2699-2709 (2003)
   PUBMED   12743594
REFERENCE   2  (bases 1 to 1261)
  AUTHORS   Li,R., Pei,H. and Papas,T.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-1999) Center for Molecular and Structural
            Biology, Department of Medicine, and Hollings Cancer Center,
            Medical University of South Carolina, 86 Jonathan Lucas St.,
            Charleston, SC 29425, USA
FEATURES             Location/Qualifiers
     source          1. .1261
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22.1-p22.3"
                     /cell_line="Jurkat"
     gene            1. .1261
                     /gene="EAP2"
     CDS             26. .1114
```

```
                     /gene="EAP2"
                     /codon_start=1
                     /product="ETS1-associated protein 2"
                     /protein_id="AAG35600.1"
                     /db_xref="GI:11493669"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

  Query Match            50.9%;  Score 668.2;  DB 8;  Length 1261;
  Best Local Similarity  77.0%;  Pred. No. 2e-130;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    | |   |  || |   | | ||  |||     ||||| | |||||
Db         42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||  |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |   |||| |||   ||   ||  || || || |||||||||
Db        222 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||    ||  ||| ||||||| |   |
Db        282 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 341

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||    | ||| ||||||||||||||||||||||||||||
Db        342 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 401

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||| | ||||||||| || ||||
Db        402 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 461

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||||| ||
Db        462 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 521

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||||| | |||||||
Db        522 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 581

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        582 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 641

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  ||||||| | ||   ||||||||  ||||||||||||||||||
Db        642 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 701

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||||  |  || |||||||  ||| |
Db        702 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 761

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||||||||||||||| ||| | || ||||||||||||||||||||||||||||| |||
Db        762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  |||||  || | ||||  ||||||  |
Db        822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy        1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
               |||| |||||||||||||||||||||||||||||||||||||  ||    ||    || ||
Db         882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy        1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
               | ||  |||||||| ||| | |||||||||||||||||||||||      ||||||||
Db         942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy        1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
               |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||||
Db        1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy        1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
               ||||||| |||||||||||||| ||  |||||| |||   || || |||||| ||| ||||
Db        1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

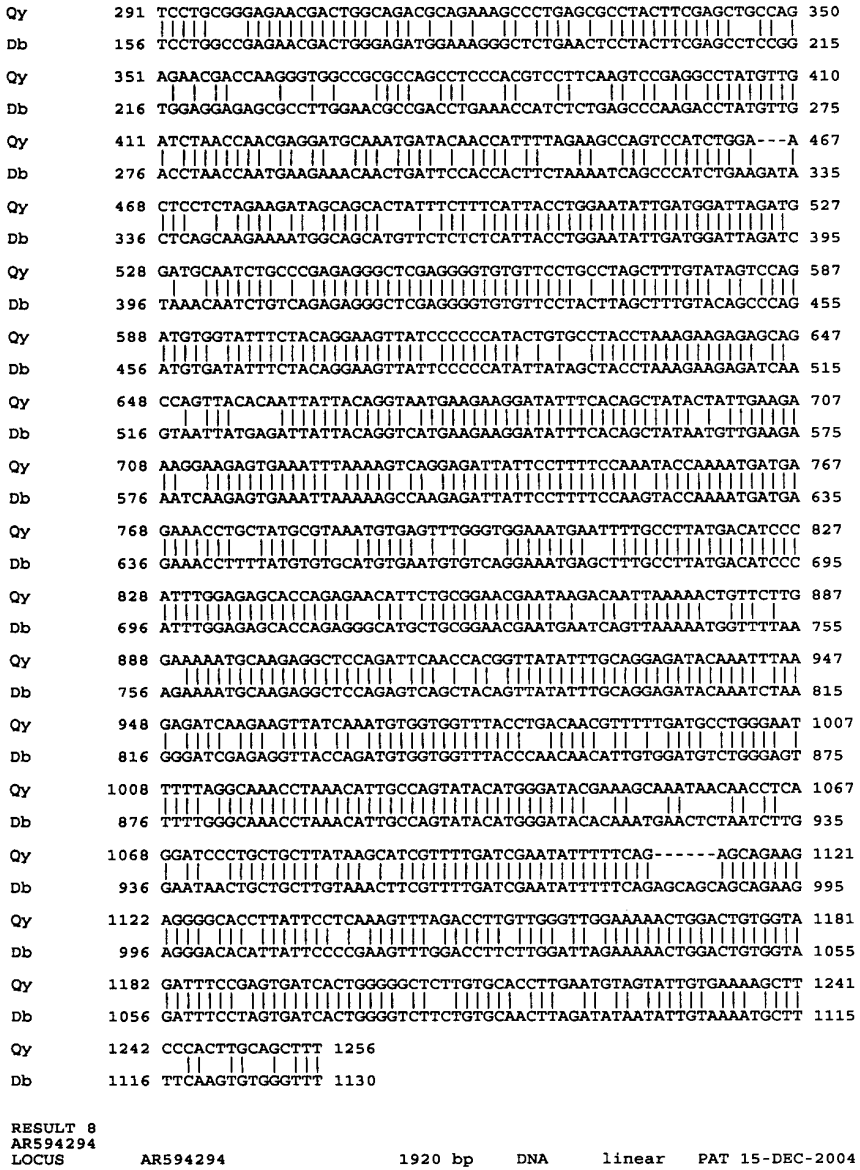
Qy        1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db        1122 TTCAAGTGTGGGTTT 1136


RESULT 7
BD205490
LOCUS       BD205490               1920 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  CD40-Interacting and TRAF-interacting protein.
ACCESSION   BD205490
VERSION     BD205490.1  GI:33015260
KEYWORDS    JP 2002512796-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1920)
  AUTHORS   Pype,S.M.C., Ghislain,J.E.F.J., Remacle and Huylebroeck,D.F.E.
  TITLE     CD40-Interacting and TRAF-interacting protein
  JOURNAL   Patent: JP 2002512796-A 1 08-MAY-2002;
            VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002512796-A/1
            PD   08-MAY-2002
            PF   28-APR-1999 JP 2000546003
            PR   29-APR-1998 EP   98201392.2
            PI   STEFAN MARIA CHRISTIAAN PYPE,
            PI   JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI   DANNY
            FRANCOIS EVELINE HUYLEBROECK
            PC   C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC
              A61P37/02,
            PC   A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,
            PC   G01N33/566,
            PC   C12N15/00,A61K37/02
            CC   CD40-Interacting and TRAF-interacting protein. FH   Key
                 Location/Qualifiers
            FT   CDS             (20). .(1108).
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            50.9%;  Score 668.2;  DB 6;  Length 1920;
  Best Local Similarity  77.0%;  Pred. No. 2.1e-130;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    | |   |  || |   | | ||  |||     ||||| | |||||
Db         36 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 95

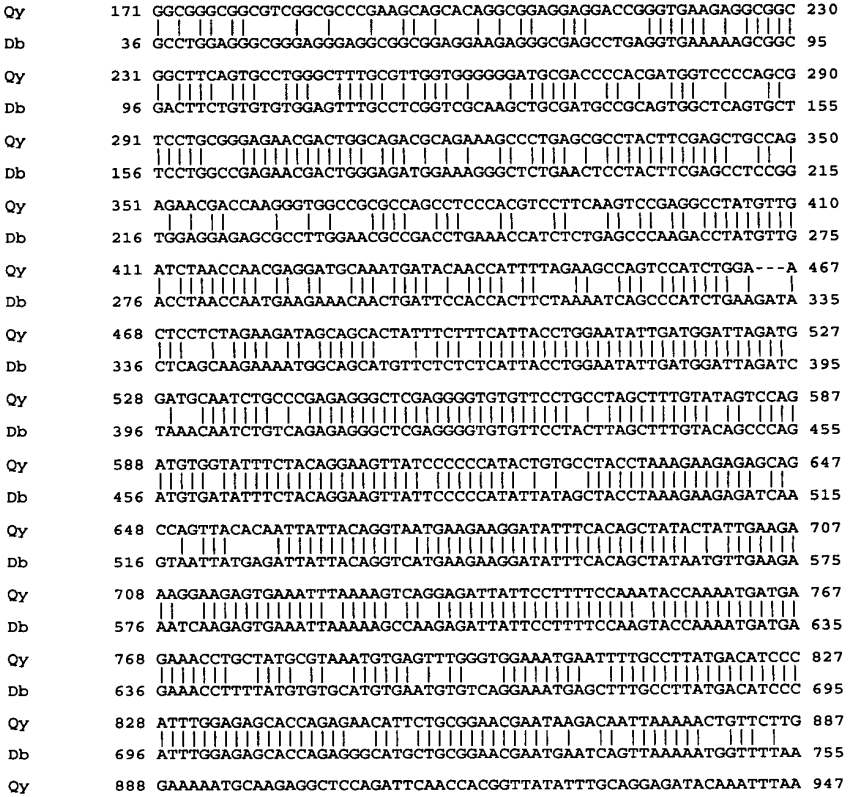
Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db         96 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 155
```



RESULT 8
AR594294
LOCUS AR594294 1920 bp DNA linear PAT 15-DEC-2004

DEFINITION Sequence 1 from patent US 6812203.
ACCESSION AR594294
VERSION AR594294.1 GI:56643900
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
TITLE CD40-Interacting and TRAF-interacting proteins
JOURNAL Patent: US 6812203-A 1 02-NOV-2004;
Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
Zwijnaarde;
WOX;
FEATURES Location/Qualifiers
source 1. .1920
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 50.9%; Score 668.2; DB 6; Length 1920;
Best Local Similarity 77.0%; Pred. No. 2.1e-130;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;



```
Db        756 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 815

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        816 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 875

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        876 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 935

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        936 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 995

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db        996 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1055

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1056 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1115

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1116 TTCAAGTGTGGGTTT 1130


RESULT 9
AX011599
LOCUS       AX011599               1920 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 1 from Patent WO9955859.
ACCESSION   AX011599
VERSION     AX011599.1  GI:9998123
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
  TITLE     Cd40-interacting and traf-interacting proteins
  JOURNAL   Patent: WO 9955859-A 1 04-NOV-1999;
            REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
            (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
            CHRISTIAAN (BE)
FEATURES             Location/Qualifiers
     source          1. .1920
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             20. .1108
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07555.1"
                     /db_xref="GI:9998124"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITTWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

  Query Match           50.9%;  Score 668.2;  DB 6;  Length 1920;
  Best Local Similarity   77.0%;  Pred. No. 2.1e-130;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         36 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 95

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db         96 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 155

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        156 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 215

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        216 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 275

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        276 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 335

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        336 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 395

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        396 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 455

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        456 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 515

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        516 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 575

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        576 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 635

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        636 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 695

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        696 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 755

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        756 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 815

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        816 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 875

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        876 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 935

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        936 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 995

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db        996 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1055

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1056 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1115

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1116 TTCAAGTGTGGGTTT 1130
```

```
RESULT 10
AR339398
LOCUS       AR339398               1948 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 889 from patent US 6569662.
ACCESSION   AR339398
VERSION     AR339398.1  GI:33726255
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1948)
  AUTHORS   Tang,Y.T., Zhou,P. and Drmanac,R.T.
  TITLE     Nucleic acids and polypeptides
  JOURNAL   Patent: US 6569662-A 889 27-MAY-2003;
            Hyseq, Inc.; Sunnyvale, CA
FEATURES             Location/Qualifiers
     source          1. .1948
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           50.9%;  Score 668.2;  DB 6;  Length 1948;
  Best Local Similarity   75.6%;  Pred. No. 2.1e-130;
  Matches  858;  Conservative    0;  Mismatches  268;  Indels    9;  Gaps    2;

Qy        131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
              | ||  | ||| |     | | ||   |   |  |    | | || ||    |  |
Db         15 GCCAAGAATTCGGCACGAGGGAAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAG 74

Qy        191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
              |  || |   | | ||  |||     ||||| | |||||| |||| |||  |||  |||
Db         75 GCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTT 134

Qy        251 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 310
              || | ||| |   | |||||  || |  |||  |   || |||||   ||||||||||
Db        135 GCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGG 194

Qy        311 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 370
              |||  | | |  || |||| | ||||||||||||  || | | | ||    |  | |
Db        195 GAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAA 254

Qy        371 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 430
              ||||  |||   ||   ||  || || || |||||||||| |||||||| || ||  ||
Db        255 CGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 314

Qy        431 AATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGC 487
              | |||| | |||| ||   ||  ||| ||||||| |   ||||  | |||| || |||||
Db        315 ACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGC 374

Qy        488 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 547
              |   | ||| |||||||||||||||||||||||||||||  |  ||||||| | ||||||
Db        375 ATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGG 434

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||||||||||||| | |||||||||| || |||||||| ||||||||||||||||
Db        435 GCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAA 494

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||| |||||||| | |  |||||||||||||||||| ||    | |||     |||||||||
Db        495 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 554

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAA 727
              ||| |||||||||||||||||||||||| | |||||||||  ||||||||||| |||
Db        555 GGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAA 614

Qy        728 AGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT 787
              || || |||||||||||||||||| ||||||||||||||||||||  |||| ||  ||
Db        615 AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 674

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   |||||||| ||||||||||||||||||||||||||||||||||||||
Db        675 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 734

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||| |  || |||||||  ||| |    ||||||||||||||||||
Db        735 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 794

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || ||||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        795 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 854

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||
Db        855 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 914

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||  ||   ||    || ||  | ||  |||||||||| |||
Db        915 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 974

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
              | |||||||||||||||||||||||      |||||||||||| ||| ||||||| | |
Db        975 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 1034

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
              ||||| |||||| |||| || ||||||||||||||||||||||||| ||||||||||||
Db       1035 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 1094

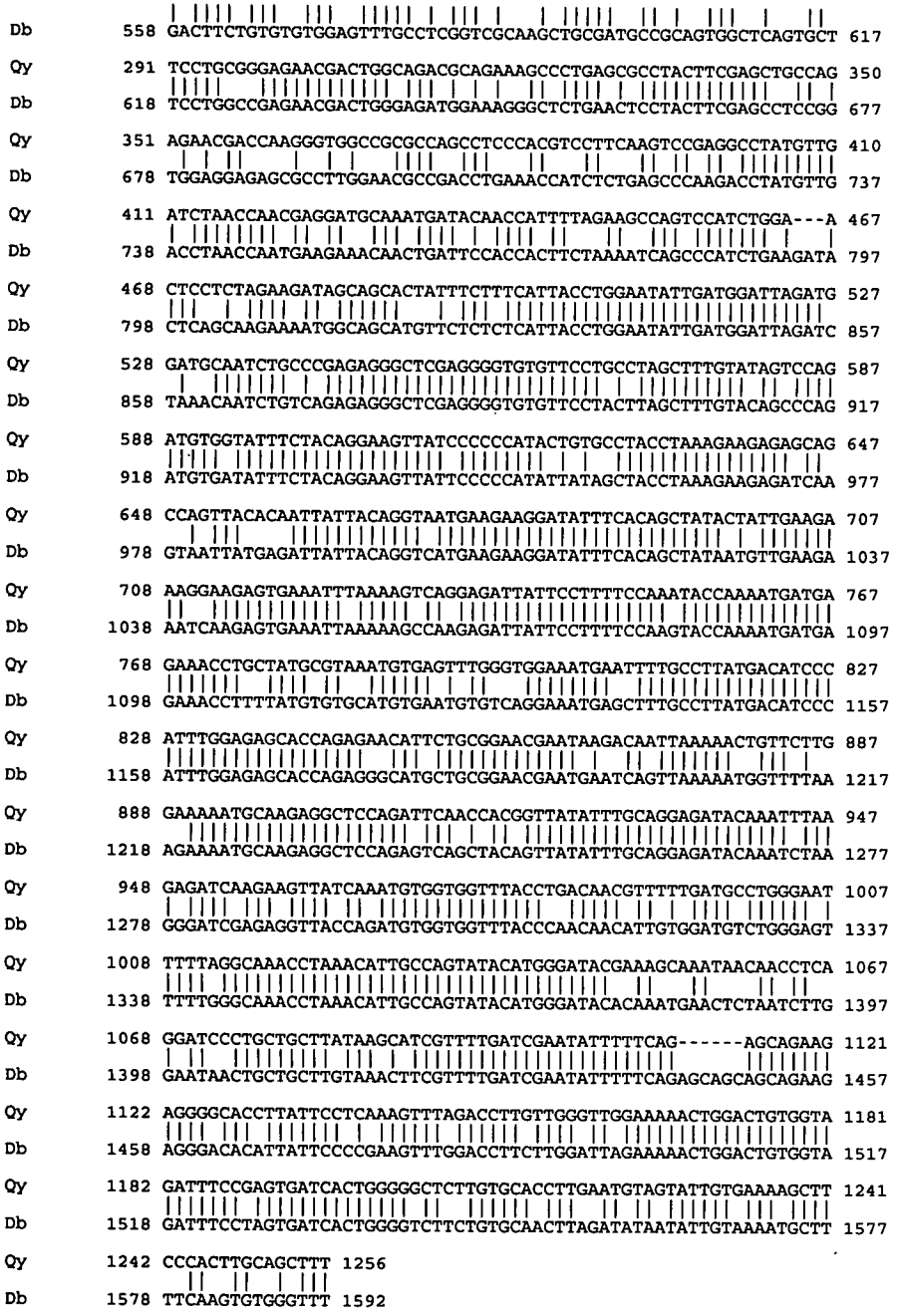
Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              || ||  ||||| ||| || || |||||| ||| |||| ||  ||  | |||
Db       1095 GGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 1149


RESULT 11
E23195
LOCUS       E23195                 2499 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Topoisomerase binding protein.
ACCESSION   E23195
VERSION     E23195.1  GI:13024277
KEYWORDS    JP 1999075856-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2499)
  AUTHORS   Takashi,T. and Kazuhiko,Y.
  TITLE     Topoisomerase binding protein
  JOURNAL   Patent: JP 1999075856-A 1 23-MAR-1999;
            TAKASHI TSURUO,CHUGAI PHARMACEUT CO LTD
COMMENT     OS   Unidentified
            PN   JP 1999075856-A/1
            PD   23-MAR-1999
            PF   17-SEP-1997 JP 1997251544
            PR
            PI   TAKASHI TSURUO,KAZUHIKO YAMANE
            PC   C12N15/09,C07K14/47,C12N9/90,C12N15/00
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .2499
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2499
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match           50.9%;  Score 668.2;  DB 6;  Length 2499;
  Best Local Similarity   77.0%;  Pred. No. 2.2e-130;
  Matches  843;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |  | || |||     ||||| | |||||
Db        498 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 557

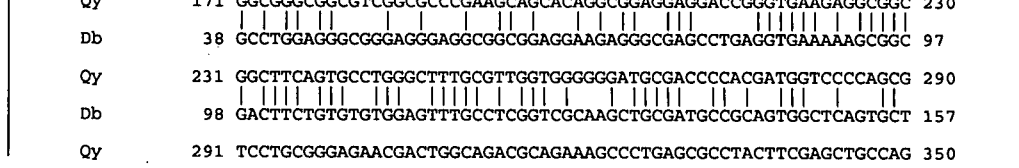
Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
```

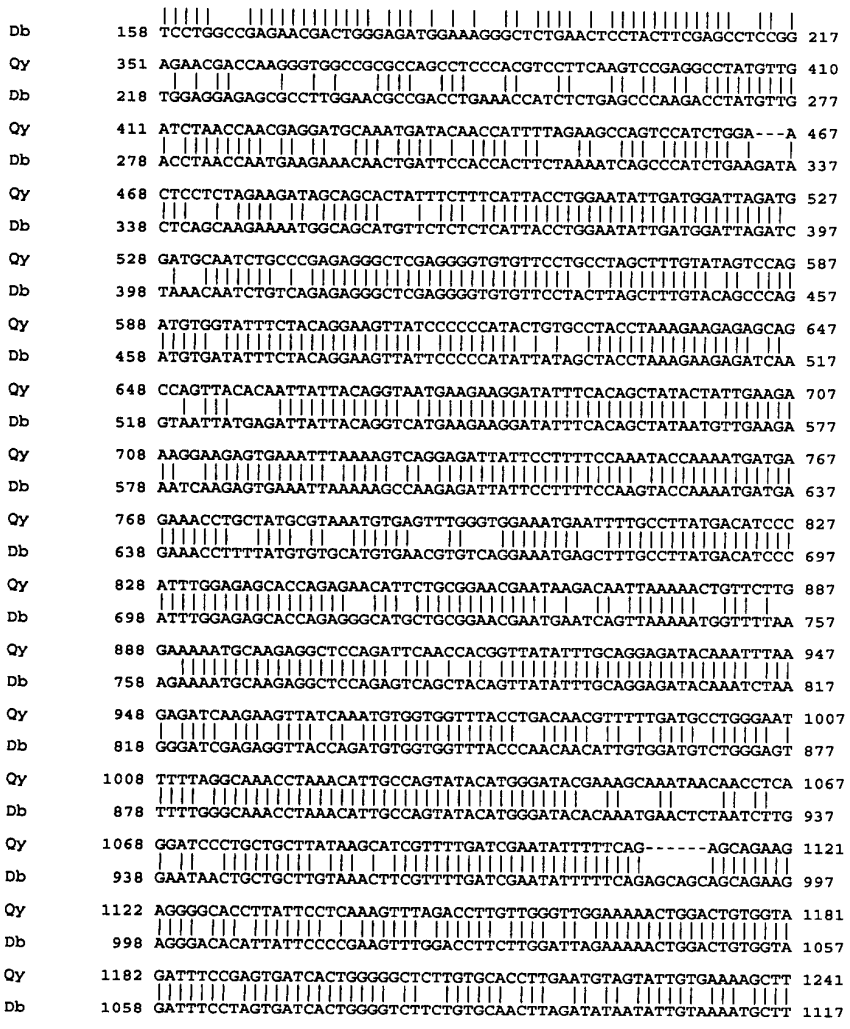


RESULT 12

HSA269473
LOCUS HSA269473 1921 bp mRNA linear PRI 27-MAY-2000
DEFINITION Homo sapiens mRNA for TRAF and TNF receptor associated protein (ttrap gene).
ACCESSION AJ269473
VERSION AJ269473.1 GI:8247253
KEYWORDS TRAF and TNF receptor associated protein; ttrap gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Pype,S., Declercq,W., Ibrahimi,A., Michiels,C., Van Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P., Huylebroeck,D. and Remacle,J.E.
TITLE TTRAP, a novel protein that associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor-associated factors (TRAFs), and that inhibits nuclear factor-kappa B activation
JOURNAL J. Biol. Chem. 275 (24), 18586-18593 (2000)
PUBMED 10764746
REFERENCE 2 (bases 1 to 1921)
AUTHORS Pype,S.
TITLE Direct Submission
JOURNAL Submitted (08-SEP-1999) Pype S., VIB07, Dept. Cell Growth, Differentiation and Development, Flanders Interuniversity Institute for Biotechnology, CELGEN, K.U.Leuven, Gasthuisberg Campus, Herestraat 49, B-3000 Leuven, BELGIUM
COMMENT Related sequence: AL031775.
FEATURES Location/Qualifiers
source 1. .1921
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1. .1921
/gene="ttrap"
5'UTR <1. .21
/gene="ttrap"
CDS 22. .1110
/gene="ttrap"
/function="putative role in signal transduction"
/codon_start=1
/product="TRAF and TNF receptor associated protein"
/protein_id="CAB92966.1"
/db_xref="GI:8247254"
/db_xref="GOA:O95551"
/db_xref="UniProt/TrEMBL:O95551"
/translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
3'UTR 1111. .1921
/gene="ttrap"
polyA_signal 1900. .1905
/gene="ttrap"
ORIGIN

Query Match 50.8%; Score 666.6; DB 8; Length 1921;
Best Local Similarity 76.9%; Pred. No. 4.6e-130;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;

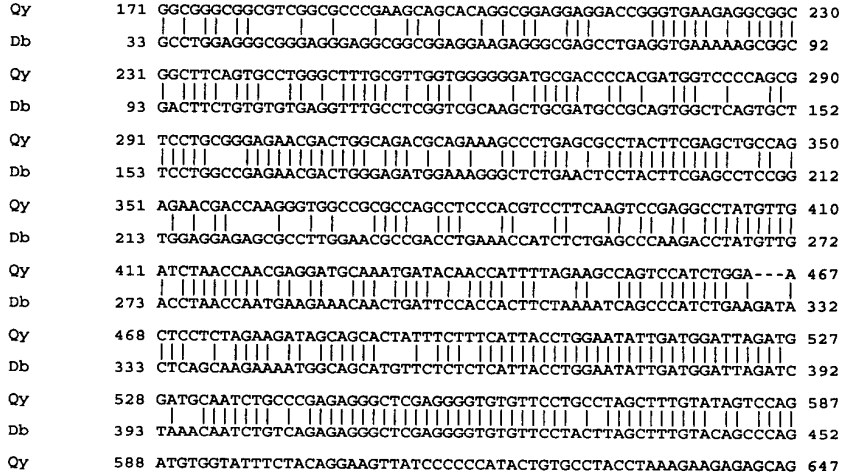




```
Db        158 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 217
Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db        218 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 277
Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db        278 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 337
Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db        338 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 397
Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db        398 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 457
Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db        458 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 517
Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db        518 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 577
Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db        578 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 637
Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTGCCTTATGACATCCC 827
Db        638 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 697
Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
Db        698 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 757
Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
Db        758 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 817
Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
Db        818 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 877
Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
Db        878 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 937
Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
Db        938 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 997
Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
Db        998 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1057
Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
Db       1058 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1117
Qy       1242 CCCACTTGCAGCTTT 1256
Db       1118 TTCAAGTGTGGGTTT 1132
```

RESULT 13
AF223469
LOCUS AF223469 1936 bp mRNA linear PRI 17-APR-2000
DEFINITION Homo sapiens AD022 protein (AD022) mRNA, complete cds.
ACCESSION AF223469
VERSION AF223469.1 GI:7578788
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1936)
AUTHORS Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
TITLE A novel gene expressed in human adrenal gland
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1936)
AUTHORS Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-2000) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, China
FEATURES Location/Qualifiers
source 1. .1936
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
gene 1. .1936
/gene="AD022"
CDS 17. .1105
/gene="AD022"
/codon_start=1
/product="AD022 protein"
/protein_id="AAF64144.1"
/db_xref="GI:7578789"
/translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVRFASVASCDAAVAQ
CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

Query Match 50.8%; Score 666.6; DB 8; Length 1936;
Best Local Similarity 76.9%; Pred. No. 4.6e-130;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;



```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db         33 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 92
Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db         93 GACTTCTGTGTGTGAGGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 152
Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db        153 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 212
Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db        213 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 272
Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db        273 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 332
Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db        333 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 392
Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db        393 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 452
Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
```

```
Db        453 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 512

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        513 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 572

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        573 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 632

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        633 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 692

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        693 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 752

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        753 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 812

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        813 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 872

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        873 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 932

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        933 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 992

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db        993 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1052

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1053 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1112

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1113 TTCAAGTGTGGGTTT 1127
```

RESULT 14
BD157138

```
LOCUS       BD157138               1898 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD157138
VERSION     BD157138.1  GI:27862896
KEYWORDS    JP 2002191363-A/11981.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1898)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 11981 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/11981
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                    Location/Qualifiers
            FT   CDS             (1). .(1086).
FEATURES            Location/Qualifiers
     source         1. .1898
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
ORIGIN
```

Query Match 50.7%; Score 665; DB 6; Length 1898;
Best Local Similarity 76.8%; Pred. No. 9.9e-130;
Matches 841; Conservative 0; Mismatches 245; Indels 9; Gaps 2;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        617 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 676

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        677 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 736

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        737 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 796

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
```

```
Db        797 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 856

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        857 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 916

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        917 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 976

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db        977 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1036

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1037 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1096

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1097 TTCAAGTGTGGGTTT 1111


RESULT 15
AX878304
LOCUS       AX878304               1898 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 13209 from Patent EP1074617.
ACCESSION   AX878304
VERSION     AX878304.1  GI:40033040
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 13209 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .1089
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE90095.1"
                     /db_xref="GI:40033041"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN

  Query Match           50.7%;  Score 665;  DB 6;  Length 1898;
  Best Local Similarity   76.8%;  Pred. No. 9.9e-130;
  Matches  841;  Conservative    0;  Mismatches  245;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        617 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 676

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        677 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 736

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        737 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 796

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        797 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 856

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        857 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 916

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        917 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 976

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db        977 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1036

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1037 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1096

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1097 TTCAAGTGTGGGTTT 1111


RESULT 16
AK002168
LOCUS       AK002168               1898 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ11306 fis, clone PLACE1010031.
ACCESSION   AK002168
```

```
VERSION     AK002168.1  GI:7023882
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
            Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
            Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
            Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
            Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
            Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
            Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
  TITLE     Complete sequencing and characterization of 21,243 full-length
            human cDNAs
  JOURNAL   Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED   14702039
REFERENCE   2
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
            Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
            Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
            Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
            Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1898)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE1010031"
                     /tissue_type="placenta"
                     /clone_lib="PLACE1"
                     /note="cloning vector: pME18SFL3"
     CDS             1. .1089
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAA92119.1"
                     /db_xref="GI:7023883"
                     /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                     CFLAENDWEMERALNSCFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                     SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                     LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                     ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                     PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
ORIGIN
```

```
  Query Match           50.7%;  Score 665;  DB 8;  Length 1898;
  Best Local Similarity   76.8%;  Pred. No. 9.9e-130;
  Matches  841;  Conservative    0;  Mismatches  245;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     ||||||  | ||||| 
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | ||||||  || || |||  |   ||  
Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||| |||||||  || |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   |||| |||   ||   ||   || || ||  |||||||||  
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   | 
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||| 
Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||| | |||||||||||||||||| |||| | |||||||||  || ||||  
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||||| |||||||| | |  ||||||||||||||| || 
Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||     |||||||||||| |||||||||||||||||||||||||| | |||||||  
Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||  
Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  ||||||   ||   ||||||||  |||||||||||||||||  
Db        617 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 676

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||  |  || |||||||  ||| |  
Db        677 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 736

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||||||||||||||||| ||| | || |||||||||||||||||||||||||| ||| 
Db        737 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 796

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  ||||| || | |||| ||||||  | 
Db        797 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 856
```

```
Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||||   ||    ||    || ||
Db        857 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 916

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  ||||||||| ||| | ||||||||||||||||||||||      ||||||||
Db        917 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 976

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db        977 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1036

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1037 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1096

Qy       1242 CCCACTTGCAGCTTT 1256
              ||  ||  | |||
Db       1097 TTCAAGTGTGGGTTT 1111
```

```
RESULT 17
AC152027/c
LOCUS       AC152027              150344 bp    DNA     linear   HTG 11-NOV-2004
DEFINITION  Dasypus novemcinctus clone VMRC5-12D21, WORKING DRAFT SEQUENCE, 5
            ordered pieces.
ACCESSION   AC152027
VERSION     AC152027.2  GI:55700150
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Dasypus novemcinctus (nine-banded armadillo)
  ORGANISM  Dasypus novemcinctus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Xenarthra; Dasypodidae; Dasypus.
REFERENCE   1  (bases 1 to 150344)
  AUTHORS   Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
            Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
            Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Laric,P.,
            Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L.,
            Maduro,V.B., Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
            Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A.,
            Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,
            Shah,K., Sison,C., Stantripop,S., Stephen,E., Thomas,J.W.,
            Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
            Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 150344)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625
            Fishers Lane, Rockville, MD 20852, USA
COMMENT     On Nov 11, 2004 this sequence version replaced gi:54312159.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: ifi
                 Center clone name: 012D21

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
            human), and/or confirmation by PCR testing. In addition,
            the sequence assembly is based on at least 8X average
            coverage in Q20 bases and has been reviewed to rule out
            gross misassemblies, the low-quality ends of sequence
            contigs have been trimmed away, and each base is associated
            with a Phrap-derived quality score.
            -------------- Summary Statistics
                 Sequencing vector: plasmid; n/a; 100% of reads
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 0.990319
                 Consensus quality: 149480 bases at least Q40
                 Consensus quality: 149754 bases at least Q30
                 Consensus quality: 149909 bases at least Q20
                 Insert size: 155000; agarose-fp
                 Insert size: 149944; sum-of-contigs
                 Quality coverage: 9.52x in Q20 bases; agarose-fp
                 Quality coverage: 9.84x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1     36273: contig of 36273 bp in length
            *    36274     36373: gap of unknown length
            *    36374     55049: contig of 18676 bp in length
            *    55050     55149: gap of unknown length
            *    55150    101939: contig of 46790 bp in length
            *   101940    102039: gap of unknown length
            *   102040    114806: contig of 12767 bp in length
            *   114807    114906: gap of unknown length
            *   114907    150344: contig of 35438 bp in length.
FEATURES             Location/Qualifiers
     source          1. .150344
                     /organism="Dasypus novemcinctus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9361"
                     /clone="VMRC5-12D21"
                     /clone_lib="VMRC5"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .36273
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    1. .23436
                     /note="clone overlaps with GenBank Accession Number
                     AC152029 clone VMRC5-168B11 (center project name ifj)"
     gap             36274. .36373
                     /estimated_length=unknown
     misc_feature    36374. .55049
                     /note="assembly_fragment"
     gap             55050. .55149
                     /estimated_length=unknown
     misc_feature    55150. .101939
                     /note="assembly_fragment"
     gap             101940. .102039
                     /estimated_length=unknown
     misc_feature    102040. .114806
                     /note="assembly_fragment"
     gap             114807. .114906
                     /estimated_length=unknown
     misc_feature    114907. .150344
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             49.4%;  Score 648.2;  DB 14;  Length 150344;
```

```
  Best Local Similarity   74.1%;  Pred. No. 5.3e-126;
  Matches  850;  Conservative    0;  Mismatches  288;  Indels    9;  Gaps    2;

Qy        159 CAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGG 218
              | | | |  |  |||  | | ||  |  || |||  |  |   |  | ||||  |||||
Db      53593 CGGCGACCCCGAGGCCAGAGCCGAGGCGGCGCGAGACGACGGGGAGCGGGGAGCCCCGGG 53534

Qy        219 TGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGA 278
              ||||||  ||||||  ||   |||   ||  |||||  | |||  | |   ||||||||  |
Db      53533 TGAAGAAGCGGCGACTCTTGTGTGCGGAGTTTGCCTCGGTCGCGAATTGCGACCCGGCCG 53474

Qy        279 TGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACT 338
              |||  ||   ||  |||||   ||||||||||||  |||  |  |  |||  |  ||  | ||||||
Db      53473 TGGCCCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGAGCGCAGAACTCCTACT 53414

Qy        339 TCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCG 398
              | ||||  |   | | |  ||    |   |  | |||  |||||  ||  |||  | ||  |||
Db      53413 TTGAGCCGTGCGTGGAGGAGAGCGCCCCGGAGAGCCGGCCTCGGACCCCCTGCCAGCCCG 53354

Qy        399 AGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTC 458
              ||||||  ||||||  ||||  ||||||  |||  |  |  ||||  |  |    ||   ||  ||
Db      53353 AGGCCTGTGTTGACCTAAGCAACGAAGATACGACTGATTCCAATTGTTCTAAAATCACCT 53294

Qy        459 CATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTG 515
              ||||||  |   | ||   | ||||||| ||||  |    | |||||  ||||||||||||||||
Db      53293 CATCTGAAAATATTCAGCAAGAAGATGGCAGTATGTTCTCTTTAATTACCTGGAATATTG 53234

Qy        516 ATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTT 575
              |||||||  |||   |  |||  ||||  |||||||||||||||||||||||  |  ||  |||
Db      53233 ATGGATTGGATCTAAACAACCTGCAAGAGAGGGCTCGAGGGGTGTGTTCTTACTTAACTT 53174

Qy        576 TGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAA 635
              ||||  ||||||||||||  ||||||||||||||||||||  ||  ||  ||||      |||||||
Db      53173 TGTACAGTCCAGATGTGATATTTCTACAGGAAGTTATTCCTCCGTACTACAGTTACCTAA 53114

Qy        636 AGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTA 695
              ||||||||||    | ||||   ||||||||||||  ||||||||||||||||||||||
Db      53113 AGAAGAGAGCAAGTAATTACGAGATTATTACAGGTCGTGAAGAAGGATATTTCACAGCTA 53054

Qy        696 TACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATA 755
              ||  | |||||||||  ||||||||||||  |||||  ||  ||||||||||||||||||  ||
Db      53053 TAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTA 52994

Qy        756 CCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCC 815
              ||||||||||||||||||   | ||  ||  ||||||||  ||    |||||||||  |||  ||
Db      52993 CCAAAATGATGAGAAACCTATTGTGTGTGCATGTGAGTGTGTCAGGAAATGAACTTTACC 52934

Qy        816 TTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAA 875
              |||||||  ||||||||  ||||||||||||||  ||||  ||   |||||||||  |   ||  ||||
Db      52933 TTATGACTTCCCATTTAGAGAGCACCAGAGGACATGCTAAGGAACGAATCAATCAGTTAA 52874

Qy        876 AAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAG 935
              ||   |||  |   |||||||||||||||||||||   ||  |  ||||||||||||||||||
Db      52873 AAGTGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGACAGCTACGGTTATATTTGCAGGAG 52814

Qy        936 ATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTG 995
              ||||  ||||||||  ||||||||||||||  ||  ||||||||||||  ||   |||||  ||||  |
Db      52813 ATACGAATTTAAGGGATCAAGAAGTTACCAGATGTGGTGGTTTGCCCAACAACATTTTGG 52754

Qy        996 ATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAA 1055
              |||  ||||||  |||||  |||  | ||||||||||  ||||||||||||||  ||    |
Db      52753 ATGTCTGGGAGTTTTTGGGCAAGCCTGAGCATTGCCAGTTTACATGGGATACGCAAATGA 52694

Qy       1056 ATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG-- 1113
              |    |||||  | ||  ||||||  |||  |||||  |||||||||  ||||  ||||||||
Db      52693 ACTTTAACCTTGGAATACCTGCTACTTGTAAGCTTCGTTTTGATTGAATCTTTTTCAGAG 52634

Qy       1114 ----AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAAC 1169
                  ||||||||||||  ||   | |||||  | ||||  | ||||||  ||||||||||||||||
Db      52633 CAGCAGCAGAAGAGGGCCATATCATTCCCCGAAGTATGGACCTTCTTGGGTTGGAAAAAC 52574

Qy       1170 TGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTAT 1229
              ||||||||||  ||||||||  ||||||||||||||  ||  |||||||  |||   ||  ||  |||

Db      52573 TGGACTGTGGCAGATTTCCTAGTGATCACTGGGGTCTTTTGTGCAGCTTAGATATAATAT 52514

Qy       1230 TGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACTAGTTCTGAATTTGTGT 1289
              |||  |   |||   ||  ||   |||  |   ||  |   |||     |   |  |  |
Db      52513 TGTAACCTGCTATTCAGATGTGAGTTTTAAATCTTCTGCGTTTTAGCAAATATAATTTTT 52454

Qy       1290 AGGTCTC 1296
              |  | ||
Db      52453 ACCTGTC 52447


RESULT 18
HSA420495
LOCUS       HSA420495               2151 bp    mRNA    linear   PRI 23-NOV-2001
DEFINITION  Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1761756.
ACCESSION   AJ420495
VERSION     AJ420495.1  GI:17066359
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Auffray,C., Ansorge,W., Ballabio,A., Estivill,X., Gibson,K.,
            Lehrach,H., Poustka,A. and Lundeberg,J.
  TITLE     The European IMAGE consortium for integrated Molecular analysis of
            human gene transcripts
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2151)
  AUTHORS   Persson,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2001) Persson A., Center for Molecular
            Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagsvagen
            30B, 106 91 Stockholm, SWEDEN
COMMENT     This clone is available royalty-free through IMAGE Consortium
            Distributors. IMPORTANT: This sequence represents the full insert
            of this IMAGE cDNA clone. No attempt has been made to verify
            whether this corresponds to the full-length of the original mRNA
            from which it was derived.
FEATURES             Location/Qualifiers
     source          1. .2151
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="EUROIMAGE 1761756"
                     /clone_lib="NCI_CGAP_Kid3"
     polyA_signal    2105. .2110
ORIGIN

  Query Match           44.6%;  Score 585.2;  DB 8;  Length 2151;
  Best Local Similarity   81.7%;  Pred. No. 7.1e-113;
  Matches  703;  Conservative    0;  Mismatches  148;  Indels    9;  Gaps    2;

Qy        406 TGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGG 465
              ||||||  ||||||||  ||  ||   |||  ||||  |  ||||  ||    ||  |||  |||||||
Db      478 TGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGA 537

Qy        466 A---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATT 522
              |   ||||  |  ||||  ||  ||||||    |  |||  ||||||||||||||||||||||||
Db      538 AGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATT 597

Qy        523 AGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAG 582
              ||||  |  |||||||  |  ||||||||||||||||||||||||  |  ||||||||||  ||
Db      598 AGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAG 657

Qy        583 TCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAG 642
              |||||||||  ||||||||||||||||||||  ||||||||  |  |  ||||||||||||||||
Db      658 CCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAG 717

Qy        643 AGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATT 702
              | ||   | |||    ||||||||||||  ||||||||||||||||||||||||||  | ||
```

```
Db        718 ATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTT 777

Qy        703 GAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAAT 762

Db        778 GAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAAT 837

Qy        763 GATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGAC 822

Db        838 GATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGAC 897

Qy        823 ATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGT 882

Db        898 ATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGT 957

Qy        883 TCTTGGAAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAA 942

Db        958 TTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAA 1017

Qy        943 TTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTG 1002

Db       1018 TCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTG 1077

Qy       1003 GGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAA 1062

Db       1078 GGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAA 1137

Qy       1063 CCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGC 1116

Db       1138 TCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGC 1197

Qy       1117 AGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTG 1176

Db       1198 AGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTG 1257

Qy       1177 TGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAA 1236

Db       1258 TGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAA 1317

Qy       1237 AGCTTCCCACTTGCAGCTTT 1256

Db       1318 TGCTTTTCAAGTGTGGGTTT 1337


RESULT 19
BD139846/c
LOCUS       BD139846               1079 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Compounds for immunotherapy and diagnosis of breast cancer and
            methods for their use.
ACCESSION   BD139846
VERSION     BD139846.1  GI:23234791
KEYWORDS    JP 2002507387-A/63.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for immunotherapy and diagnosis of breast cancer and
            methods for their use
  JOURNAL   Patent: JP 2002507387-A 63 12-MAR-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002507387-A/63
            PD   12-MAR-2002
            PF   22-DEC-1998 JP 2000526543
            PR   24-DEC-1997 US    08/998253,24-DEC-1997 US    08/998255 PR
            17-JUL-1998 US    09/118627,17-JUL-1998 US    09/118554 PI   STEVEN
            G REED,JIANGCHUN XU
            PC   C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,
            PC   C07K14/47,
            PC
            C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/ PC
            53,
            PC   G01N33/577,C12N15/00,A61K37/02,C12N5/00
            CC   Compounds for immunotherapy and diagnosis
            of breast cancer and
            CC             methods for
            CC   their use
            FH   Key             Location/Qualifiers
            FT   source          1. .1079
            FT                   /organism='Homo sapiens (human)'.
FEATURES                Location/Qualifiers
     source             1. .1079
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN

  Query Match             37.3%;  Score 490;  DB 6;  Length 1079;
  Best Local Similarity   82.8%;  Pred. No. 8.4e-93;
  Matches  586;  Conservative    0;  Mismatches  115;  Indels    7;  Gaps    2;

Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614

Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674

Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734

Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794

Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854

Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914

Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974

Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034

Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094

Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148

Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208

Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256

Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373


RESULT 20
AR202950/c
LOCUS       AR202950               1079 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 63 from patent US 6365348.
```

```
ACCESSION   AR202950
VERSION     AR202950.1  GI:21499210
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for diagnosis of Breast cancer and methods for their use
  JOURNAL   Patent: US 6365348-A 63 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

Query Match 37.3%; Score 490; DB 6; Length 1079;
Best Local Similarity 82.8%; Pred. No. 8.4e-93;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              ||||||||||| | |||||||||| || ||||||||| |||||||||||||||||||| |
Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              ||||||| | |  |||||||||||||||| ||   | |||    |||||||||||| |||
Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              |||||||||||||||||||||| | |||||||||  |||||||||||| ||||| || |
Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              |||||||||||||||||| |||||||||||||||||||  |||| ||  |||||| |
Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||   ||||||||  |||||||||||||||||||||||||||||||||||||  ||| |||
Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              |||||||||| |  || |||||||  ||| |    |||||||||||||||||||| |||
Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || |||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              |||||||  |||| || | |||| |||||| ||||| ||||||||||||||||||||||
Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              |||||||||||||  ||   ||    || ||  | ||  ||||||||| ||| | |||||
Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
              |||||||||||||||||||      |||||||||||| ||| |||||| | |||||| |
Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
              ||||| |||| || |||||||||||||||||||||||| ||||||||||||||  ||
Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              |||||| |||  || || |||||| ||| ||||  ||  ||  | |||
Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

```
RESULT 21
AR208051/c
LOCUS       AR208051               1079 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 63 from patent US 6379951.
ACCESSION   AR208051
VERSION     AR208051.1  GI:21507966
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   Reed,S.G. and Xu,J.
  TITLE     Compounds for immunotherapy of breast cancer and methods for their
            use
  JOURNAL   Patent: US 6379951-A 63 30-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

Query Match 37.3%; Score 490; DB 6; Length 1079;
Best Local Similarity 82.8%; Pred. No. 8.4e-93;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

```
Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              ||||||||||| | |||||||||| || ||||||||| |||||||||||||||||||| |
Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              ||||||| | |  |||||||||||||||| ||   | |||    |||||||||||| |||
Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              |||||||||||||||||||||| | |||||||||  |||||||||||| ||||| || |
Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              |||||||||||||||||| |||||||||||||||||||  |||| ||  |||||| |
Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||   ||||||||  |||||||||||||||||||||||||||||||||||||  ||| |||
Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              |||||||||| |  || |||||||  ||| |    |||||||||||||||||||| |||
Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || |||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              |||||||  |||| || | |||| |||||| ||||| ||||||||||||||||||||||
Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              |||||||||||||  ||   ||    || ||  | ||  ||||||||| ||| | |||||
Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
              |||||||||||||||||||      |||||||||||| ||| |||||| | |||||| |
Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
              ||||| |||| || |||||||||||||||||||||||| ||||||||||||||  ||
Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              |||||| |||  || || |||||| ||| ||||  ||  ||  | |||
Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

```
RESULT 22
AX429924/c
LOCUS       AX429924               1079 bp    DNA     linear   PAT 21-JUN-2002
```

```
DEFINITION  Sequence 63 from Patent WO0198339.
ACCESSION   AX429924
VERSION     AX429924.1  GI:21541088
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.
  TITLE     Compositions and methods for the therapy and diagnosis of breast
            cancer
  JOURNAL   Patent: WO 0198339-A 63 27-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           37.3%;  Score 490;  DB 6;  Length 1079;
  Best Local Similarity   82.8%;  Pred. No. 8.4e-93;
  Matches  586;  Conservative    0;  Mismatches  115;  Indels    7;  Gaps    2;

Qy        555 GGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCC 614
              ||||||||||||  | ||||||||||  || |||||||| |||||||||||||||||| |
Db       1079 GGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

Qy        615 CCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATG 674
              |||||||  | |  ||||||||||||||||  ||    | |||    ||||||||||||  |||
Db       1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

Qy        675 AAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGG 734
              |||||||||||||||||||||||| | |||||||||   ||||||||||| ||||| || |
Db        959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAG 900

Qy        735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
              |||||||||||||||||| ||||||||||||||||||||||  |||| ||  |||||| |
Db        899 AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATG 840

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||   ||||||||  |||||||||||||||||||||||||||||||||||||  ||| |||
Db        839 TGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              ||||||||||  ||  |||||||  |||  |    |||||||||||||||||||||| |||
Db        779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 720

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || |||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db        719 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 660

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              |||||||  |||| || | |||| |||||| ||||| ||||||||||||||||||||||
Db        659 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 600

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              ||||||||||||  ||   ||    || ||  | ||  ||||||||| ||| | ||||||
Db        599 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 540

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAG 1148
              |||||||||||||||||      |||||||||| ||| ||||||| | |||||| |
Db        539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT-G 481

Qy       1149 ACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
              ||||| |||| || |||||||||||||||||||||||||| ||||||||||||||| ||
Db        480 ACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTC 421

Qy       1209 TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              |||||| ||| || || |||||| ||| |||| ||  ||  | |||
Db        420 TGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 373
```

```
RESULT 23
BD150065
LOCUS       BD150065                752 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD150065
VERSION     BD150065.1  GI:27855823
KEYWORDS    JP 2002191363-A/4908.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 4908 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/4908
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                     Location/Qualifiers
            FT   source          1. .752
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           28.8%;  Score 377.8;  DB 6;  Length 752;
  Best Local Similarity   72.5%;  Pred. No. 4e-69;
  Matches  528;  Conservative    0;  Mismatches  195;  Indels    5;  Gaps    3;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||      |||||| | |||||
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || || |||  |   ||
Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||| ||||||||  || |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   |||| |||   ||   ||   || || || ||||||||||
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||    | ||| ||||||||||||||||||||||||||||||
Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||| |||| | |||||||||| || ||||
```

```
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
                | |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy        768 G-AAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 826
              | ||||||  |||| ||  ||||||   ||   ||||||||  |||||||| |||||||
Db        617 GAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTAAGACATCC 676

Qy        827 C-ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCT 885
              | ||| || || ||||| ||  |||     |||||||| |  || || ||||  | | |
Db        677 CAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATCAGTTTAAAATGGGTTT 736

Qy        886 TGGAAAAA 893
                 |||
Db        737 AAANGAAA 744


RESULT 24
AX870003
LOCUS       AX870003                752 bp    DNA     linear   PAT 17-DEC-2003
DEFINITION  Sequence 4908 from Patent EP1074617.
ACCESSION   AX870003
VERSION     AX870003.1  GI:40024866
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 4908 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             28.8%;  Score 377.8;  DB 6;  Length 752;
  Best Local Similarity   72.5%;  Pred. No. 4e-69;
  Matches  528;  Conservative    0;  Mismatches  195;  Indels    5;  Gaps    3;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     ||||||| | |||||
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || || |||  |   ||
Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCGGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||| ||||||||   || |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTGCTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
                | | ||   |  | |  ||||  |||   ||   ||  || || || |||||||||
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
```

```
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |   |
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| |||||||||||||||||||||||||||||
Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||| |||| | |||||||||| || ||||
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGCTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy        768 G-AAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 826
              | ||||||  |||| ||  ||||||   ||   ||||||||  |||||||| |||||||
Db        617 GAAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTAAGACATCC 676

Qy        827 C-ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCT 885
              | ||| || || ||||| ||  |||     |||||||| |  || || ||||  | | |
Db        677 CAATTGGGNGAACACCANAGGGCATCTGCGGGAACGAATGAATCAGTTTAAAATGGGTTT 736

Qy        886 TGGAAAAA 893
                 |||
Db        737 AAANGAAA 744


RESULT 25
CR761829
LOCUS       CR761829                1236 bp    mRNA    linear   VRT 15-SEP-2004
DEFINITION  Xenopus tropicalis finished cDNA, clone TGas102g02.
ACCESSION   CR761829
VERSION     CR761829.1  GI:51967738
KEYWORDS    .
SOURCE      Xenopus tropicalis (Silurana tropicalis)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1236)
  AUTHORS   Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
            Francis,M.D., Garrett,N., Gilchrist,M.J., Grafham,D.V.,
            McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
            Voigt,J. and Zorn,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
COMMENT     Sanger Xenopus tropicalis EST/cDNA project.
            This sequence is from a Xenopus Gene Collection (XGC) library, from
            a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
            extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
            ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'
            end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1-blue.
FEATURES             Location/Qualifiers
     source          1. .1236
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="TGas102g02"
                     /clone_lib="XGC-gastrula"
                     /dev_stage="gastrula (stage 10.5-13 mixed)"
ORIGIN
```

```
  Query Match             27.2%;  Score 357.2;  DB 5;  Length 1236;
  Best Local Similarity   67.6%;  Pred. No. 9.3e-65;
  Matches  519;  Conservative    0;  Mismatches  243;  Indels    6;  Gaps    1;

Qy        473 CTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGC 532
              | |||||||   |||  | |||| ||  | || ||||||| |||||  |  |||| ||
Db        403 CAAGAAGATGAGAGCCATTTTTCCTTTCTGACTTGGAATATAGATGGCCTTGATGAATCA 462

Qy        533 AATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTG 592
              |||  |  || |||||||| || ||||||||||    | ||  |||| ||||||||||
Db        463 AATGTCGCAGAAAGGGCTCGTGGTGTGTGTTCCTATTTGGCCCTGTACAGTCCAGATGTA 522

Qy        593 GTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGT 652
              |||||  ||||||| ||||| || ||||||| ||  ||||| |||||||||||| | |||
Db        523 GTATTCTTACAGGAGGTTATTCCACCATACTATGAGTACCTGAAGAAGAGAGCTGTCAGC 582

Qy        653 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 712
              ||||||||||||||||||||| || | ||| ||||| |||||  |  |||| ||  |
Db        583 TACACAATTATTACAGGTAATGAGGATGAATACTTCACTGCTATGATGCTGAAAAAGAGC 642

Qy        713 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAAC 772
              || || || || | ||| || || ||| | || |  |||   ||    ||||||||||
Db        643 AGGGTAAAGTTAATAAGCCAAGAAATTGTACCATACCCATCCACGGTGATGATGAGAAAT 702

Qy        773 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 832
              ||||    | |||||||| | |   ||| ||     | ||| | ||||| || ||| |
Db        703 TTGCTTGTAGCAAATGTGAATATCTCTGGCAACAGTATATGCTTAATGACTTCTCATCTA 762

Qy        833 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 892
              |||||||| | ||| |||||   ||| ||| |||  ||  | | | ||| ||  |||
Db        763 GAGAGCACAAAAGATCATTCAAAGGAGCGATTAAAGCAGCTGGACATTGTGCTAAAAAAG 822

Qy        893 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 952
              |||   || |||||   || | | || || ||||||| ||| ||||||||   |  |||||
Db        823 ATGATGGATGCTCCCCCTTTAGCAACTGTAATATTTGGAGGGGATACAAACCTGCGAGAT 882

Qy        953 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 1012
              || || |||   |||  ||| ||  | ||  | ||  | || ||||  ||||| || ||
Db        883 CAGGAGGTTGCTAAAATTGGAGGCATGCCAAATAATATATTGGATGTTTGGGACTTCTTG 942

Qy       1013 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 1072
              || || ||| | || || |  ||||| ||||| || ||||  || || |||||  |
Db        943 GGTAAGCCTGAGCACTGTCGCTATACCTGGGACACAAAAGTGAACAAGAACCTGCGTGCT 1002

Qy       1073 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGC------AGAAGAGGGG 1126
              ||      ||  | |   || ||||| || |||||||| |||||       || |   |
Db       1003 CCCTACATTTGCAGGTTGCGCTTTGACCGCATATTTTTTAGAGCTTCGCAGGAGGGAAGT 1062

Qy       1127 CACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTT 1186
              ||  | ||  ||  ||   | |  | |||| |||    ||||| |||||||||| ||||||
Db       1063 CAAGTGATCCCCCAGTCTCTTTATCTTGCTGGCACTGAAAAGCTGGACTGTGGCAGATTT 1122

Qy       1187 CCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGA 1234
              || ||||||||||||||  |  ||||   |||   | |  |||| |||
Db       1123 CCAAGTGATCACTGGGGTTTGCTGTGTGACTTTGCTATTATATTATGA 1170


RESULT 26
AL589699
LOCUS       AL589699              256608 bp    DNA     linear   ROD 17-DEC-2004
DEFINITION  Mouse DNA sequence from clone RP23-92G13 on chromosome 13 Contains
            the 3' end of the gene for a novel protein similar to KIAA0386,
            five novel genes, the Gmnn gene for geminin, gene PNAS-27, the
            Ttrap gene for Traf and Tnf receptor associated protein, the gene
            for a novel protein similar to KIAA0319, the gene for the ortholog
            of human and rat aldehyde dehydrogenase 5A1 (succinate-semialdehyde
            dehydrogenase) (ALDH5A1) (SSADH), the 5' end of the Gpld1 gene for
            glycosylphosphatidylinositol specific phospholipase D1 and four CpG
            islands, complete sequence.
ACCESSION   AL589699
VERSION     AL589699.4  GI:13810107
KEYWORDS    HTG; Aldh5a1; CpG island; geminin; Gmnn; Gpld1; KIAA0319; KIAA0386;
            phospholipase; SSADH; Ttrap.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 256608)
  AUTHORS   Pelan,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 26, 2001 this sequence version replaced gi:13751568.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep --------------
            Genome Center
            Center: UK Medical Research Council
            Center code: UK-MRC
            Web site: http://mrcseq.har.mrc.ac.uk
            Contact: mouseq@har.mrc.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence.  Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
            RP23-92G13 is from the RPCI-23 Mouse BAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .256608
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="13"
                     /clone="RP23-92G13"
                     /clone_lib="RPCI-23"
     gene            join(AL513014.9:11808. .79376,4770. .28457)
                     /locus_tag="RP23-461O24.1-003"
     mRNA            join(AL513014.9:11808. .12037,AL513014.9:43061. .43187,
                     AL513014.9:44856. .45011,AL513014.9:46133. .46211,
                     AL513014.9:47748. .47771,AL513014.9:48876. .48929,
                     AL513014.9:50825. .50974,AL513014.9:53729. .53792,
                     AL513014.9:62567. .62610,AL513014.9:66867. .66992,
                     AL513014.9:67218. .67366,AL513014.9:68674. .68803,
                     AL513014.9:74163. .74831,AL513014.9:79192. .79376,
                     4770. .4932,8263. .8398,9931. .10092,12188. .12346,
                     16272. .16474,18350. .18433,19587. .19677,26226. .28457)
                     /locus_tag="RP23-461O24.1-003"
                     /note="match: cDNAs: AK007145.1 AB002384.1
                     match: ESTs: AW822704.1 BB674120.1 AA041998.1 BM144846.1
                     Z42557.1 R20596.1"
     CDS             join(AL513014.9:11989. .12037,AL513014.9:43061. .43187,
                     AL513014.9:44856. .45011,AL513014.9:46133. .46211,
                     AL513014.9:47748. .47771,AL513014.9:48876. .48929,
                     AL513014.9:50825. .50974,AL513014.9:53729. .53792,
                     AL513014.9:62567. .62610,AL513014.9:66867. .66992,
                     AL513014.9:67218. .67366,AL513014.9:68674. .68803,
                     AL513014.9:74163. .74831,AL513014.9:79192. .79376,
                     4770. .4932,8263. .8398,9931. .10092,12188. .12346,
                     16272. .16474,18350. .18433,19587. .19677,26226. .26326)
                     /locus_tag="RP23-461O24.1-003"
                     /standard_name="OTTMUSP00000000566"
```

```
                /note="match: proteins: Q9D9C2 Q9Y4F9"
                /codon_start=1
                /protein_id="CAI26080.1"
                /db_xref="GI:56237839"
                /translation="MPPGTKRLRALGAFSAGLPTRLPEIMLVGSQSFSPGGPNGIIRS
                QSFAGPSGLQERRSRCNSFIENASALKKPQAKLKKMHNLGHKNNNTPKEPQPKRVEEV
                YRALKNGLDEYLEFHQTELDKLTAQLKDMKRNSRLGVLYDLDKQIKTIERYMRRLEFH
                ISKVDELYEAYCIQRRLQDGASKMKQAFATSPASKAARESLSEINRSYKEYTENMCAI
                EAELESLLGEFSIKMKGLAGFARLCPGDQYEIFMKYGRQRWKLKGKIEVNGKQSWDGA
                ETVFLPLIVGFISIKVTELKGLATHILVGSVTCETKELFAARPQVVAVDINDLGTIKL
                NLEITWYPFDVEDTTPSSGPGNKTAALQRRMSMYSQGTPETPTFKDQSFFSNLPDDIF
                ESGKAAEEKRPLSLSFSDLQDGDCVFTSSSATSPSSSHSAHPEITITPAELTHSSLSS
                QNEGTEDSSSASSRNSLGEDHEPKSHPKSDTVEPGKPGVATRSGTESLFLESSVAEAL
                LQESDEASELKPVELDTFEGNITKQLVKRLTSAEGPVTTDKLFFEGSVGSESEAGRSF
                LDGSLEDAFNGLFLALDPHKKQYKEFQDLNQEVTHLDDVLKCKPAGSRSRSSSLSLTV
                ESALESFDFLNTSDFDEEEEDGDDVCHVGGGADSVFSDTETEKSGSVHPEARGHLSEA
                LTEDTGVGTSVAGSPLPLTTGNESLDITIVKHLQYCTQLIQQIVFSSKTPFVARSLLE
                KLSRQVLVMQKLAAVSDENLGNITSVVEAIPEFHKKLSLLAFWTKCCSPSGVYHSSAA
                RLIKQLEASFARSINKDYPGLAEPVFRTLVSQILDRAEPLLSSSLSSEVITVFQYYSF
                FTSHGVSDLETYLGQLTRQVAMVQTLQSLRDEKLLQTMSDLAPSNLPAQQEVLRTLAL
                LLTRDDNEVSEAVTLYLAAASKNEHFREKALLYYCEALTKANLQLQKAACLALKSLEA
                TESIKMLVTLCQSDTEEIRTVASETLLSLGEDGRLAYEQLDKFPRDCVKVGGRHGTEV
                ATAF"
polyA_signal    26424. .26429
                /locus_tag="RP23-461O24.1-003"
polyA_site      26446
                /locus_tag="RP23-461O24.1-003"
polyA_site      26447
                /locus_tag="RP23-461O24.1-003"
polyA_site      26457
                /locus_tag="RP23-461O24.1-003"
gene            join(35873. .36201,40700. .41233)
                /locus_tag="RP23-92G13.3-001"
mRNA            join(35873. .36201,40700. .41233)
                /locus_tag="RP23-92G13.3-001"
                /note="match: cDNAs: AK006029.1
                match: ESTs: AV045211 BF018417 BB012811 BB014319 BB020000
                AV263251 BF319523 AV206925 AV041162 AV259155 AV041846"
polyA_signal    39211. .39216
                /locus_tag="RP23-92G13.3-001"
polyA_site      39233
                /locus_tag="RP23-92G13.3-001"
polyA_site      complement(44486)
polyA_site      complement(44487)
polyA_site      complement(44488)
polyA_signal    complement(44511. .44516)
polyA_signal    complement(44515. .44520)
gene            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56454. .56564))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-001"
mRNA            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56454. .56564))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-001"
                /note="match: ESTs: BI789536.1 BQ178448.1 BF722692.1
                AA510015.1 BG797776.1 BB868504.1 AA014584.1 AA032471.1
                BI454155.1 BE448331.1 W13399.1 BF556131.1 BG085437.1
                AI528457.1 AA277821.1 AA681520.1 W99871.1 AA004051.1"
gene            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56228. .56502))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-002"
mRNA            complement(join(46486. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54858,56228. .56502))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-002"
                /note="match: cDNAs: AF068780.1 BC020061.1
                match: ESTs: BF160648.1 AA437903.1"
CDS             complement(join(46710. .46871,47989. .48099,48299. .48381,
                51241. .51376,52237. .52314,54801. .54851))
                /gene="Gmnn"
                /locus_tag="RP23-92G13.2-001"
                /standard_name="OTTMUSP00000000569"
                /note="match: proteins: O88513 O75496"
                /codon_start=1
                /protein_id="CAI26081.1"
                /db_xref="GI:56237840"
                /translation="MNLSMKQKQEGAQENVKNSPVPRRTLKMIQPSADGSLVGRENEL
                PKGLFKRKLWDDQLASQTSSCGPEANENKDVGDLTQEAFDLISKENPSSQYWKEVAEQ
                RRKALYEALKENEKLHKEIEQKDSEIARLRKENKDLAEVAEHVQYMAEVIERLSNEPL
                DNFESPDSQEFDSEEEAVEYSELEDSGAGTCAEETVSSSTDARPCT"
gene            96664. .108654
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
mRNA            join(96664. .98150,99423. .99599,100503. .100625,
                104527. .104661,107278. .108654)
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
                /note="match: cDNAs: BC002217.1 BC005537.1 AL136632.1
                AK022681.1 AJ420475.1
                match: ESTs: BI078093.1 BG295445.1 BI696541.1 BG717389.1
                BI108344.1 BG975946.1 BI452750.1 BG168325.1 BG295139.1
                BG914583.1 BI753721.1"
CDS             join(98022. .98150,99423. .99599,100503. .100625,
                104527. .104661,107278. .107403)
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
                /standard_name="OTTMUSP00000000563"
                /note="match: proteins: Q9H1Z2 Q99LU8 Q9GZU0 O95550"
                /codon_start=1
                /protein_id="CAI26082.1"
                /db_xref="GI:56237841"
                /translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
                KKSALFEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
                DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEFHHGDYEKQFLHV
                LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
                IEDHLRPYMPE"
polyA_signal    106607. .106612
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
polyA_site      106654
                /gene="RP23-92G13.4"
                /locus_tag="RP23-92G13.4-001"
```

```
  Query Match           27.0%;  Score 354.4;  DB 9;  Length 256608;
  Best Local Similarity   99.7%;  Pred. No. 6.1e-64;
  Matches  355;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135726 AGGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 135785

Qy       1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135786 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 135845

Qy       1077 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTC 1136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135846 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTC 135905

Qy       1137 CTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATC 1196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135906 CTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATC 135965

Qy       1197 ACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135966 ACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 136025

Qy       1257 ACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     136026 ACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 136081


RESULT 27
CQ726600
```

```
LOCUS       CQ726600                858 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 12534 from Patent WO02068579.
ACCESSION   CQ726600
VERSION     CQ726600.1  GI:42290170
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 12534 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .858
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             23.7%;  Score 311;  DB 6;  Length 858;
  Best Local Similarity   82.1%;  Pred. No. 5.1e-55;
  Matches  372;  Conservative    0;  Mismatches   75;  Indels    6;  Gaps    1;

Qy        788 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
              |||| | ||   ||||||||  ||||||||||||||||||||||||||||||||||||
Db        406 GTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 465

Qy        848 CATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
              ||| |||||||||||||  | || |||||||  ||| |    ||||||||||||||||||
Db        466 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCA 525

Qy        908 GATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967
              || ||| | || ||||||||||||||||||||||||| |||| |||| ||| |||| || |
Db        526 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 585

Qy        968 TGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACAT 1027
              |||||||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||
Db        586 TGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACAT 645

Qy       1028 TGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087
              |||||||||||||||||||  ||   ||    || ||  | ||  ||||||||| |||
Db        646 TGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAA 705

Qy       1088 CATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAA 1141
              | |||||||||||||||||||||||      |||||||||||| ||| ||||||| | |
Db        706 CTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGA 765

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201
              ||||| |||||| |||| || ||||||||||||||||||||||||| ||||||||||||
Db        766 AGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGG 825

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGA 1234
              || ||  |||||| |||  || || |||||| |
Db        826 GGTCTTCTGTGCAACTTAGATATAATATTGTAA 858


RESULT 28
BC083404
LOCUS       BC083404                1451 bp    mRNA    linear   VRT 27-MAY-2005
DEFINITION  Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7238334), partial
            cds.
ACCESSION   BC083404
VERSION     BC083404.1  GI:53734140
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 1451)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1451)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: John Postlethwait, University of Oregon
            cDNA Library Preparation: Dr. Sumio Sugano
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 192 Row: c Column: 18.
FEATURES             Location/Qualifiers
     source          1. .1451
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:7238334"
                     /tissue_type="Ovary"
                     /clone_lib="NIH_ZGC_5"
                     /lab_host="DH10B"
                     /note="Vector: pME18S-FL3"
     gene            <1. .1451
                     /gene="TTRAP"
                     /db_xref="GeneID:553516"
     CDS             <1. .1141
                     /gene="TTRAP"
                     /codon_start=2
                     /product="TTRAP protein"
                     /protein_id="AAH83404.1"
                     /db_xref="GI:53734141"
                     /db_xref="GeneID:553516"
                     /translation="FLSGNQTDQQMSALEESRTRLCDQFAFVSGSDSAVAQCYLAENE
                     WDMERALNSFFEAHMDSVFDEEAEKTEVTGNKRKDDTAEASGTKKKLKTDNADCIDLT
                     AEEPTCSITVNSKENQAENGTAKSEVEDSKLSIISWNVDGLDTLNLADRARGLCSYLA
                     LYTPDVVFLQELIPAYVQYLKKRAVSYLFFEGSDDGYFTGIMLRKSRVKFLESEIICF
                     PTTQMMRNLLIAQVTFSGQKLYLMTSHLESCKNQSQERTKQLRVVLQKIKEAPEDAIV
                     IFAGDTNLRDAEVANVGGLPAGVCDVWEQLGKQEHCRYTWDTKANSNKTVPYVSRCRF
```

DRIFLRSAKTGPPVTPDHMALIGMEKLDCGRYTSDHWGIYCTFNT"

ORIGIN

Query Match 21.1%; Score 276.6; DB 5; Length 1451;
Best Local Similarity 60.0%; Pred. No. 9.8e-48;
Matches 481; Conservative 0; Mismatches 314; Indels 6; Gaps 1;

```
Qy        427 TGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAG 486
              || |||   || |  | |  |  | | |    | |  |   ||  ||  ||   |||
Db        334 TGTAAACTCTAAAGAAAATCAGGCTGAAAATGGCACAGCCAAATCTGAAGTGGAGGACAG 393

Qy        487 CACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAG 546
              ||   | ||  |||| | |||||||  | || || ||  ||   ||  |||||  | ||  |
Db        394 CAAGCTCTCCATCATCAGCTGGAATGTGGACGGTTTGGACACATTAAATCTTGCAGATCG 453

Qy        547 GGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGA 606
              ||  | ||  |||| || |  |||||| | || |  |||||||||| ||||| || ||
Db        454 TGCCAGGGGCTTGTGCTCATATCTAGCTCTATACACACCAGATGTGGTTTTTCTTCAAGA 513

Qy        607 AGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTAC 666
              | || ||  | ||   |  || || |||||  | || | ||| |||    || ||
Db        514 GCTCATTCCAGCTTATGTTCAGTATCTTAAGAAACGTGCCGTCAGCTACTTGTTTTTTGA 573

Qy        667 AGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAA 726
              ||| |  || || ||||| || ||  ||| | |||| ||||  |||||| ||||||
Db        574 AGGAAGCGATGATGGATACTTTACCGGGATAATGTTGAGGAAATCAAGAGTCAAATTTTT 633

Qy        727 AAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAA 786
                  ||||| ||    |||||||  ||  |||||||||| || |||||    |   |
Db        634 GGAAAGTGAGATCATCTGCTTTCCAACAACACAAATGATGAGGAATCTGCTCATTGCTCA 693

Qy        787 TGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGA 846
              ||||  |||    ||  |  |  | | | |||||||||||||||  || |||    | | |
Db        694 GGTGACTTTCTCAGGTCAGAAGCTATACTTGATGACATCCCATTTAGAAAGCTGTAAAAA 753

Qy        847 ACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCC 906
              || ||   |||  |||  | ||||||       ||| |||  ||| ||  | || || ||
Db        754 CCAATCCCAGGAGAGAACGAAACAATTGCGGGTTGTGCTTCAAAAGATTAAGGAAGCGCC 813

Qy        907 AGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAA 966
              |||     |||  || || |||||||| || |||||| | |||||   |||  ||   |||
Db        814 AGAAGATGCCATTGTCATTTTTGCAGGTGACACAAATCTCAGAGACGCAGAGGTGGCCAA 873

Qy        967 ATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACA 1026
              |||  |||||  ||||      || | |||||  ||||||     | |||||||   ||||
Db        874 TGTTGGAGGTTTGCCTGCAGGTGTGTGTGATGTGTGGGAACAGCTGGGCAAACAAGAACA 933

Qy       1027 TTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAA 1086
              |||  ||| || |||||  || ||||| || | |||     |    | |||   |    |
Db        934 CTGCAGGTACACCTGGGACACCAAAGCTAACAGCAATAAGACTGTGCCTTACGTCAGCAG 993

Qy       1087 GCATCGTTTTGATCGAATATTTTTCAGAGC------AGAAGAGGGGCACCTTATTCCTCA 1140
              |||  || || || |||||  | ||  |       |    |    ||   | | |||| |
Db        994 ATGTCGCTTCGACCGCATATTCCTGAGGTCTGCTAAAACCGGTCCGCCGGTCACTCCTGA 1053

Qy       1141 AAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTG 1200
                | | | |  |||| |||| |||||||||||| ||| | | | ||||||||||||
Db       1054 CCACATGGCCTTGATTGGAATGGAGAAACTGGACTGCGGTCGTTATACCAGTGATCACTG 1113

Qy       1201 GGGGCTCTTGTGCACCTTGAA 1221
              |||  | |  ||||| || ||
Db       1114 GGGTATTTACTGCACTTTCAA 1134
```

RESULT 29
BC097117
LOCUS BC097117 1455 bp mRNA linear VRT 23-JUN-2005
DEFINITION Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7431816), partial cds.
ACCESSION BC097117
VERSION BC097117.1 GI:66910444
KEYWORDS .
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 1455)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
CONSRTM Mammalian Gene Collection Program Team
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 1455)
AUTHORS .
CONSRTM NIH MGC Project
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2005) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Will Talbot
cDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 226 Row: g Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 53734140.

FEATURES Location/Qualifiers
source 1. .1455
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7431816"
/tissue_type="Embryo, whole, 2-8 hr, mixed wild type, pool of 2500 embryos"
/clone_lib="NIH_ZGC_15"
/lab_host="DH10B"
/note="Vector: pME18S-FL3"
gene <1. .1455
/gene="TTRAP"
/db_xref="GeneID:553516"
CDS <1. .438
/gene="TTRAP"
/codon_start=1
/product="TTRAP protein"
/protein_id="AAH97117.1"
/db_xref="GI:66910445"
/db_xref="GeneID:553516"

```
                     /translation="VFSGNQTDQQMSALGESRTRLCDQFAFVSGSDSAVAQCYLAENE
                     WDMERALNSFFEAHMDSVFDEEAEKTEVTGNKRKDDTAEASGTKKKLKTDNADCIDLT
                     AEEPTCSITVNSKKIRLKMAQPNLKWRTASSPSSAGMWTVWTH"
ORIGIN

  Query Match             20.9%;  Score 274.2;  DB 5;  Length 1455;
  Best Local Similarity   61.2%;  Pred. No. 3.2e-47;
  Matches  463;  Conservative    0;  Mismatches  288;  Indels    6;  Gaps    1;

Qy        471 CTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGAT 530
              ||  ||  ||   |||||  | ||  |||| | ||||||| | || || || ||    ||
Db        376 CTGAAGTGGAGGACAGCAAGCTCTCCATCATCAGCTGGAATGTGGACGGTTTGGACACAT 435

Qy        531 GCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATG 590
              |||||  | ||  | ||  | ||  |||| || |  |||||| | || |  |||||||
Db        436 TAAATCTTGCAGATCGTGCCAGGGGCTTGTGCTCATATCTAGCTCTATACACACCAGATG 495

Qy        591 TGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGCAGCCA 650
              |||| ||||| || ||  | || ||  | ||   |   || || |||||  | || | ||
Db        496 TGGTTTTTCTTCAAGAGCTCATTCCAGCTTATGTTCAGTATCTTAAGAAACGTGCCGTCA 555

Qy        651 GTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAG 710
              | |||    || ||  ||| |  || || ||||| || || |  ||| | || | ||||
Db        556 GCTACTTGTTTTTTGAAGGAAGCGATGATGGATACTTTACCGGGATAATGTTAAGGAAAT 615

Qy        711 GAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAA 770
              ||||| |||||          ||||| ||  ||||||||  ||  ||||||||||  |
Db        616 CAAGAGTAAAATTTTTGGAAAGCGAGATCATCTGTTTTCCAACAACACAAATGATGAGGA 675

Qy        771 ACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATT 830
              | |||||    |   | |||| |||    ||  |  |  | | | |||||||||||||
Db        676 ATCTGCTCATTGCTCAGGTGACTTTCTCAGGTCAGAAGCTATACTTGATGACATCCCATT 735

Qy        831 TGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAA 890
              | || |||   | | | || ||   |||  |||  | ||||||     ||| |||  ||
Db        736 TAGAAAGCTGTAAAAACCAATCCCAGGAGAGAACGAAACAATTGCGGGTTGTGCTTCAAA 795

Qy        891 AAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAG 950
              | ||  | || || |||||     |||  || || |||||||| || |||||| | ||||
Db        796 AGATTAAGGAAGCGCCAGAAGATGCCATTGTCATTTTTGCAGGCGACACAAATCTCAGAG 855

Qy        951 ATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTT 1010
              |   ||| ||   |||   ||| ||||| ||||      || | |||||  ||||||
Db        856 ACGCAGAGGTGGCCAATGTTGGAGGTTTGCCTGCAGGTGTGTGTGATGTGTGGGAACAGC 915

Qy       1011 TAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGA 1070
              | |||||||   |||| |||  ||| || ||||| || ||||| || | |||    |
Db        916 TGGGCAAACAAGAACACTGCAGGTACACCTGGGACACCAAAGCTAACAGCAATAAGACTG 975

Qy       1071 TCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGC------AGAAGAGG 1124
              | |||    |     |   ||| || || || ||||| | ||  |      |   |
Db        976 TGCCTTACGTCAGCAGATGTCGCTTCGACCGCATATTCCTGAGGTCTGCTAAAACCGCTC 1035

Qy       1125 GGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGAT 1184
              ||   | | |||| |      | | |  ||||  |||| |||||||||||| ||| | |
Db       1036 CGCCGGTCACTCCTGACCACAAGGCCTTGATTGGAATGGAGAAACTGGACTGCGGTCGTT 1095

Qy       1185 TTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAA 1221
                | | |||||||||||||||  | |  |||| || ||
Db       1096 ATACCAGTGATCACTGGGGTATTTACTGCACTTTCAA 1132


RESULT 30
AC125756/c
LOCUS       AC125756               245394 bp    DNA     linear   HTG 13-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-42H3, WORKING DRAFT SEQUENCE, 2
            unordered pieces.
ACCESSION   AC125756
VERSION     AC125756.5  GI:30578987
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 245394)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 245394)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 245394)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 13, 2003 this sequence version replaced gi:23265926.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
```

```
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GJVU
                Center clone name: CH230-42H3
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 238380 bases at least Q40
                Consensus quality: 239660 bases at least Q30
                Consensus quality: 240760 bases at least Q20
                Estimated insert size: 255581; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
            * NOTE: This sequence may represent more than one clone.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   242009: contig of 242009 bp in length
            *  242010   242109: gap of unknown length
            *  242110   245394: contig of 3285 bp in length.
FEATURES             Location/Qualifiers
     source          1. .245394
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-42H3"
     misc_feature    1. .1025
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    1076. .2080
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    3723. .4536
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH348771"
     misc_feature    128656. .128715
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH260088"
     misc_feature    236171. .238412
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    239605. .242009
                     /note="wgs_end_extension
                     clone_end:Sp6"
     gap             242010. .242109
                     /estimated_length=unknown
ORIGIN

  Query Match             20.9%;  Score 273.6;  DB 14;  Length 245394;
  Best Local Similarity   88.4%;  Pred. No. 6.9e-47;
  Matches  311;  Conservative    0;  Mismatches   34;  Indels    7;  Gaps    1;

Qy        957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016
              | |||||||||||||||||||||||||||||| |||||||||||||||||||||||| ||||
Db     140207 AGGTTATCAAATGTGGTGGTTTACCTGACAATGTTTTTGATGCCTGGGAATTTTTGGGCA 140148

Qy       1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076
              ||||||| ||||||| |||||||||||||||||||||||  | || ||||||||| ||||
Db     140147 AACCTAAGCATTGCCGGTATACATGGGATACGAAAGCAAACGATAATCTCAGGATTCCTG 140088

Qy       1077 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTC 1136
              | |||| |||||||||||||||||||||||| |||||||||||||||| ||||||||||
Db     140087 CCGCTTGTAAGCATCGTTTTGATCGAATATTCTTCAGAGCAGAAGAGGGACACCTTATTC 140028

Qy       1137 CTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATC 1196
              | |||||||||||||  |||||||||||| |||||||||||||||||||| |||||||
Db     140027 CCCAAAGTTTAGACCTAATTGGGTTGGAAAGACTGGACTGTGGTAGATTTCCTAGTGATC 139968

Qy       1197 ACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              |||||||||| |||| ||||||||||||||||||||||||||||||||        ||
Db     139967 ACTGGGGGCTCCTGTGTACCTTGAATGTAGTATTGTGAAAAGCTTCCCA-------AATT 139915

Qy       1257 ACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGA 1308
              ||  | |||| |||||||||||||||||||  || ||||||||||  | |
Db     139914 CCAGCTCTGTTTGCACTAGTTCTGAATTTGTACAGATCTCAACCTTTAGGAA 139863


RESULT 31
BD076937
LOCUS       BD076937                483 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  5' EST of secretory protein expressed in prostate.
ACCESSION   BD076937
VERSION     BD076937.1  GI:22622540
KEYWORDS    JP 2001512013-A/184.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
  TITLE     5' EST of secretory protein expressed in prostate
  JOURNAL   Patent: JP 2001512013-A 184 21-AUG-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001512013-A/184
            PD   21-AUG-2001
            PF   31-JUL-1998 JP 2000505291
            PR   01-AUG-1997 US    08/905144
            PI   JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO  PI
            LACROIX
            PC   C12N15/09,C07K14/47,C12P21/02,C12Q1/02,C12Q1/68,C12N15/00 CC
            blastn
            CC   identity 99
            CC   region 1. .310
            CC   id T26956
            CC   est
            CC   blastn
            CC   identity 98
            CC   region 1. .315
            CC   id T31666
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 137. .267
            CC   id R14990
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 63. .137
            CC   id R14990
            CC   est
            CC   blastn
            CC   identity 100
            CC   region 1. .50
            CC   id R14990
```

```
CC   est
CC   Von Heijne matrix
CC   score 5.2
CC   seq LCVEFASVASCDA/AV
FH   Key             Location/Qualifiers
FT   misc_feature    105. .414
FT   misc_feature    45. .359
FT   misc_feature    202. .332
FT   misc_feature    127. .201
FT   misc_feature    65. .114
FT   sig_peptide     1. .120.
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             17.7%;  Score 231.8;  DB 6;  Length 483;
  Best Local Similarity   70.0%;  Pred. No. 2.5e-38;
  Matches  327;  Conservative    0;  Mismatches  137;  Indels    3;  Gaps    1;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   | || |   | | ||  |||     ||||||| | ||||| 
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |  | |||||  || |  |||  |   ||  
Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||   || |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   ||||  |||   ||   ||   || || || |||||||||
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |   |
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||| | |||||||||| || ||||
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTA 634
              ||||| |||||||||||||||| |||||||| | |  |||||||
Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTA 483


RESULT 32
CQ663156
LOCUS       CQ663156                400 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 8082 from Patent WO02070737.
ACCESSION   CQ663156
VERSION     CQ663156.1  GI:42140949
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 8082 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             17.3%;  Score 227.2;  DB 6;  Length 400;
  Best Local Similarity   79.3%;  Pred. No. 2.3e-37;
  Matches  284;  Conservative    0;  Mismatches   68;  Indels    6;  Gaps    1;

Qy        795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTG 854
              ||| |  |||  | |     |  |||||||||||||||||||||||||  ||| |||
Db         42 TGGATCCAAAGAATTCGGCACGAGTGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 101

Qy        855 CGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
              |||||||||| |  || |||||||  ||| |    |||||||||||||||||||| |||
Db        102 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAG 161

Qy        915 CCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTG 974
              | || |||||||||||||||||||||||| |||| |||| ||| |||| || |||||||||
Db        162 CTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTG 221

Qy        975 GTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGT 1034
              |||||||  ||||| || | |||| ||||| ||||| ||||||||||||||||||||||
Db        222 GTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGT 281

Qy       1035 ATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094
              |||||||||||||  ||    ||     || ||  | ||  |||||||| ||| | |||||
Db        282 ATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTT 341

Qy       1095 TTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTT 1146
              |||||||||||||||||      ||||||||||| ||| ||||||| | ||||||
Db        342 TTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT 399


RESULT 33
AY613922
LOCUS       AY613922                20938 bp    DNA     linear   PRI 11-MAY-2004
DEFINITION  Homo sapiens TRAF and TNF receptor associated protein (TTRAP) gene,
            complete cds.
ACCESSION   AY613922
VERSION     AY613922.1  GI:47060296
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 20938)
  AUTHORS   Rieder,M.J., Daniels,R.L., da Ponte,S.H., Hastings,N.C.,
            Ahearn,M.O., Rajkumar,N., Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2004) Genome Sciences, University of Washington,
            1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  SeattleSNPs. NHLBI HL66682 Program
            for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.gs.washington.edu).
FEATURES             Location/Qualifiers
     source          1. .20938
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     repeat_region   1. .32
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     variation       81
                     /frequency="0.67"
                     /replace=""
     repeat_region   87. .368
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     repeat_region   369. .673
                     /rpt_family="Alu"
```

```
                /rpt_type=dispersed
variation       449
                /frequency="0.96"
                /replace=""
variation       540
                /frequency="0.01"
                /replace="t"
variation       586
                /frequency="0.01"
                /replace="a"
variation       650
                /frequency="0.03"
                /replace="g"
variation       658. .660
                /frequency="0.01"
                /replace=""
repeat_region   683. .989
                /rpt_family="Alu"
                /rpt_type=dispersed
variation       695
                /frequency="0.18"
                /replace="t"
variation       922
                /frequency="0.02"
                /replace="t"
variation       1319
                /frequency="0.15"
                /replace="t"
variation       1456
                /frequency="0.19"
                /replace="a"
variation       1532
                /frequency="0.01"
                /replace="c"
variation       1711. .1713
                /frequency="0.04"
                /replace=""
variation       1795
                /frequency="0.10"
                /replace="a"
variation       1821
                /frequency="0.02"
                /replace="t"
variation       1830. .1862
                /frequency="0.04"
                /replace=""
variation       1913. .1914
                /frequency="0.13"
                /replace=""
gene            <2005. .18938
                /gene="TTRAP"
mRNA            join(<2005. .2194,2281. .2366,10172. .10345,11003. .11094,
                14376. .14494,15753. .15923,17846. .18938)
                /gene="TTRAP"
                /product="TRAF and TNF receptor associated protein"
CDS             join(2030. .2194,2281. .2366,10172. .10345,11003. .11094,
                14376. .14494,15753. .15923,17846. .18127)
                /gene="TTRAP"
                /codon_start=1
                /product="TRAF and TNF receptor associated protein"
                /protein_id="AAT09764.1"
                /db_xref="GI:47060297"
                /translation="MELGSCLEGGREAAEEEGEPEVKKRRLLCVEFASVASCDAAVAQ
                CFLAENDWEMERALNSYFEPPVEESALERRPETISEPKTYVDLTNEETTDSTTSKISP
                SEDTQQENGSMFSLITWNIDGLDLNNLSERARGVCSYLALYSPDVIFLQEVIPPYYSY
                LKKRSSNYEIITGHEEGYFTAIMLKKSRVKLKSQEIIPFPSTKMMRNLLCVHVNVSGN
                ELCLMTSHLESTRGHAAERMNQLKMVLKKMQEAPESATVIFAGDTNLRDREVTRCGGL
                PNNIVDVWEFLGKPKHCQYTWDTQMNSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL
                DLLGLEKLDCGRFPSDHWGLLCNLDIIL"
variation       2381
                /gene="TTRAP"
                /frequency="0.08"
                /replace="a"
variation       2767
                /gene="TTRAP"
                /frequency="0.08"
                /replace="g"
variation       2800
                /gene="TTRAP"
                /frequency="0.79"
                /replace=""
repeat_region   2919. .3323
                /rpt_family="L2"
                /rpt_type=dispersed
variation       2927
                /gene="TTRAP"
                /frequency="0.13"
                /replace="g"
variation       3004
                /gene="TTRAP"
                /frequency="0.04"
                /replace="t"
variation       3061
                /gene="TTRAP"
                /frequency="0.13"
                /replace="t"
variation       3062
                /gene="TTRAP"
                /frequency="0.13"
                /replace="t"
variation       3114
                /gene="TTRAP"
                /frequency="0.21"
                /replace="c"
variation       3418
                /gene="TTRAP"
                /frequency="0.03"
                /replace="c"
variation       3781
                /gene="TTRAP"
                /frequency="0.15"
                /replace="g"
variation       3988
                /gene="TTRAP"
                /frequency="0.01"
                /replace="a"
variation       4191
                /gene="TTRAP"
                /frequency="0.01"
                /replace="c"
variation       4393
                /gene="TTRAP"
                /frequency="0.97"
                /replace=""
variation       4414
                /gene="TTRAP"
                /frequency="0.03"
                /replace="c"
variation       4437
                /gene="TTRAP"
                /frequency="0.07"
                /replace="g"
repeat_region   4446. .4530
                /rpt_family="MER1_type"
                /rpt_type=dispersed
repeat_region   4559. .4609
                /rpt_family="MER1_type"
                /rpt_type=dispersed
variation       4565
                /gene="TTRAP"
                /frequency="0.52"
                /replace=""
repeat_region   4625. .4760
                /rpt_family="MER1_type"
                /rpt_type=dispersed
variation       4658
```

```
                     /gene="TTRAP"
                     /frequency="0.01"
                     /replace="a"
     variation       4861
                     /gene="TTRAP"
                     /frequency="0.02"
                     /replace="a"
     variation       4941
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="g"
     variation       4983
                     /gene="TTRAP"
                     /frequency="0.07"
                     /replace="t"
     variation       5024
                     /gene="TTRAP"
                     /frequency="0.07"

  Query Match             14.9%;  Score 196;  DB 8;  Length 20938;
  Best Local Similarity   80.1%;  Pred. No. 1.3e-30;
  Matches  245;  Conservative    0;  Mismatches   55;  Indels    6;  Gaps    1;

Qy        957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016
              | |||| || |||||||||||||||  ||||| || | |||| |||||| ||||| ||||
Db      17844 AGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCA 17903

Qy       1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076
              |||||||||||||||||||||||||||||||  ||    ||    || ||  | ||  |||
Db      17904 AACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTG 17963

Qy       1077 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACC 1130
              |||||| ||| | |||||||||||||||||||||||      |||||||||||| |||
Db      17964 CTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACA 18023

Qy       1131 TTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGA 1190
              ||||||| | |||||| |||||| |||| || ||||||||||||||||||||||||||| |
Db      18024 TTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTA 18083

Qy       1191 GTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGC 1250
              |||||||||||| ||  |||||| |||  || || |||||| ||| ||||  ||  ||
Db      18084 GTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGT 18143

Qy       1251 AGCTTT 1256
                | |||
Db      18144 GGGTTT 18149


RESULT 34
HS30M3
LOCUS       HS30M3                102200 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP1-30M3 on chromosome 6p22.1-22.3
            Contains the 5' end of gene KIAA0319, the gene for TRAF and TNF
            receptor-associated protein (TTRAP)(EAP2, AD022, MGC9099), gene
            FLJ20501, the 3' end of gene FLJ12619 (DKFZP564G182) and two CpG
            islands, complete sequence.
ACCESSION   AL031775
VERSION     AL031775.1  GI:4071041
KEYWORDS    HTG; AD022; CpG island; DKFZP564G182; FLJ12619; FLJ20501; KIAA0319;
            TNF; TRAF; TTRAP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 102200)
  AUTHORS   Phillips,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 29, 1998 this sequence version replaced gi:4007165.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP1-30M3 is from the library RPCI-1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            -------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .102200
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="RZPD:RPCIP704M0330"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="p22.1-22.3"
                     /clone="RP1-30M3"
                     /clone_lib="RPCI-1"
     gene            join(complement(AL133264.10:4611. .6907),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-001"
     mRNA            join(complement(AL133264.10:4611. .6907),
                     complement(AL133264.10:2219. .2395),
                     complement(AL133264.10:389. .511),5119. .5253,7768. .9169)
                     /gene="C6orf62"
                     /locus_tag="RP1-30M3.4-001"
                     /product="chromosome 6 open reading frame 62"
                     /note="match: ESTs: AA084220 AA093151 AA168907 AA197971
                     AA200765 AA237701 AA292313 AA295166 AA402388 AA478637
                     AA591228 AA612170 AA638148 AA647423 AA647435 AA647742
                     AA672357 AA798123 AA798467 AA815944 AA833812 AA863774
                     AI053561 AI054330 AI122327 AI156738 AI191279 AI194277
                     AI284633 AI316639 AI390324 AI458673 AI688647 AI694146
                     AI694199 AI788240 AI802319 AI807413 AI831412 AI932363
                     AI935446 AI982078 AJ396992 AJ398211 AL529211 AL563388.1
                     AL601897.1 AU051650 AU125496 AU135776 AU137791 AU138464
                     AV656408 AV686324 AV698092 AV703346 AV708342 AV708873
                     AV709145 AV709526 AV739602 AV742411 AV743612 AV744587
                     AV745018 AW050065 AW063931 AW270998 AW300107 AW300121
                     AW301984 AW302318 AW302718 AW302841 AW302861 AW303029
                     AW355281 AW822881 AW859606 AW959243 BB067326 BB123906
                     BB609972 BB624305 BB628344 BB654156 BB823708 BE007675
                     BE049029 BE049182 BE120619 BE178692 BE178693 BE244058
                     BE246922 BE287246 BE289710 BE311050 BE369398 BE377181
                     BE386264 BE465032 BE484573 BE554360 BE571717 BE625094
                     BE766876 BE766886 BE770856 BE785982.1 BE813002 BE817780
                     BE829341 BE829344 BE829355 BE829359 BE893564 BE895461
                     BE896141 BE908143 BE913030 BE945891 BE991271 BE992722
                     BF015474 BF021751 BF022025 BF029248 BF037603 BF060909
                     BF064161 BF109576 BF111215 BF122334 BF137583 BF139381
                     BF140373 BF145038 BF148549 BF150174 BF165619 BF207643
                     BF326037 BF326041 BF326042 BF326045 BF334390 BF341149
                     BF376614 BF378829 BF380539 BF380550 BF465841 BF535383
                     BF578793 BF582307 BF748822 BF748938 BF748963 BF765346
```

BF766403 BF784199 BF789335 BF789696 BF885392 BF896088
BF940327 BF984391 BG030354 BG035327 BG088552 BG116557
BG148023 BG167018 BG168325 BG247693 BG252842 BG295272
BG295445 BG342500 BG610013 BG655426 BG667600 BG677384.1
BG678705 BG692520 BG717389 BG801183 BG802707 BG803992
BG863302 BG914583 BG917993 BG940215 BG943615 BG975946
BI080618 BI092275 BI106992 BI107793 BI108344 BI222784
BI258149 BI326830 BI651746 BI659175 BI666111 BI688353
BI689042 BI694119 BI696541 BI696575 BI696638 BI733610
BI753721 BI754669 BI851772 BI856689.1 BI872006 BI872428
BI903238 BI903250 BI960823 BI986171 BM016286 BM069622
BM069876 BM147734 BM237580 BM238719 BM238848 BM423047
BM423107 F11344 F20959 H17162 H98057 N56057 R25494 T11219
T51727 W47233
match: cDNAs: AB046018.1 AJ420475.1 AK022681.1 AL136632.1
BC002217.1 BC005537.1"

gene join(complement(AL133264.10:101. .211),5119. .5253,
7768. .8194)
/gene="C6orf62"
/locus_tag="RP1-30M3.4-002"

mRNA join(complement(AL133264.10:101. .211),5119. .5253,
7768. .8194)
/gene="C6orf62"
/locus_tag="RP1-30M3.4-002"
/product="chromosome 6 open reading frame 62"
/note="match: ESTs: BG772628 BG772628.1"

gene join(complement(AL133264.10:5931. .6069),
complement(AL133264.10:2219. .2395),
complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
/gene="C6orf62"
/locus_tag="RP1-30M3.4-003"

mRNA join(complement(AL133264.10:5931. .6069),
complement(AL133264.10:2219. .2395),
complement(AL133264.10:389. .511),5119. .5253,7768. .7934)
/gene="C6orf62"
/locus_tag="RP1-30M3.4-003"
/product="chromosome 6 open reading frame 62"
/note="match: ESTs: BG342500.1 BG692520.1 BG803992.1"

CDS join(complement(AL133264.10:4611. .4739),
complement(AL133264.10:2219. .2395),
complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
/gene="C6orf62"
/locus_tag="RP1-30M3.4-003"
/standard_name="OTTHUMP00000016093"
/note="match: proteins: O95550 Q961D0 Q99K03 Q99LU8 Q9GZU0
Q9H1Z2 Q9N0B2 Q9QEJ3 Q9W384 Q9XWI7"
/codon_start=1
/product="chromosome 6 open reading frame 62"
/protein_id="CAI20135.1"
/db_xref="GI:56202469"
/translation="MGDPNSRKKQALNRLRAQLRKKKESLADQFDFKMYIAFVFKEKK
KKSALFEVSEVIPVMTNNYEENILKGVRDSSYSLESSLELLQKDVVQLHAPRYQSMRR
DVIGCTQEMDFILWPRNDIEKIVCLLFSRWKESDEPFRPVQAKFEFHHGDYEKQFLHV
LSRKDKTGIVVNNPNQSVFLFIDRQHLQTPKNKATIFKLCSICLYLPQEQLTHWAVGT
IEDHLRPYMPE"

CDS join(complement(AL133264.10:5931. .5972),
complement(AL133264.10:2219. .2395),
complement(AL133264.10:389. .511),5119. .5253,7768. .7893)
/gene="C6orf62"
/locus_tag="RP1-30M3.4-003"
/standard_name="OTTHUMP00000016094"
/codon_start=1
/product="chromosome 6 open reading frame 62"
/protein_id="CAI20136.1"
/db_xref="GI:56202470"
/db_xref="UniProt/TrEMBL:Q5JVZ2"
/translation="MSENLSDPVSPVVRKKKSALFEVSEVIPVMTNNYEENILKGVRD
SSYSLESSLELLQKDVVQLHAPRYQSMRRDVIGCTQEMDFILWPRNDIEKIVCLLFSR
WKESDEPFRPVQAKFEFHHGDYEKQFLHVLSRKDKTGIVVNNPNQSVFLFIDRQHLQT
PKNKATIFKLCSICLYLPQEQLTHWAVGTIEDHLRPYMPE"

polyA_signal 9124. .9129
/gene="C6orf62"
/locus_tag="RP1-30M3.4-001"

gene complement(join(12316. .12571,15963. .16147,46686. .46959))
/gene="THEM2"
/locus_tag="RP1-30M3.2-001"

mRNA complement(join(12316. .12571,15963. .16147,46686. .46959))
/gene="THEM2"
/locus_tag="RP1-30M3.2-001"
/product="thioesterase superfamily member 2"
/note="match: ESTs: AA614605 AA928745 AI017897 AI052396
BF407527 BF475709 BF594161
match: cDNAs: AF155649.1 AF220186.1 AF274952.1 AK000508.1
AK002261.1 AK008624.1 BC000894.1"

gene complement(join(12316. .12571,15963. .16147,26337. .26628))
/gene="THEM2"
/locus_tag="RP1-30M3.2-002"

mRNA complement(join(12316. .12571,15963. .16147,26337. .26628))
/gene="THEM2"
/locus_tag="RP1-30M3.2-002"
/product="thioesterase superfamily member 2"
/note="match: ESTs: BG219890 BG256069"

polyA_site 12318

polyA_signal 12336. .12341

CDS complement(join(12415. .12571,15963. .16147,46686. .46766))
/gene="THEM2"
/locus_tag="RP1-30M3.2-001"
/standard_name="OTTHUMP00000016090"
/note="match: proteins: AAA28024 AAH18165 AAL48890 O95549
P34419 P87304 P93828 P95914 Q18187 Q9ASQ1 Q9NPJ3 Q9VZZ5
Q9VZZ6"
/codon_start=1
/product="thioesterase superfamily member 2"
/protein_id="CAD92509.1"

Query Match 14.9%; Score 196; DB 8; Length 102200;
Best Local Similarity 80.1%; Pred. No. 1.5e-30;
Matches 245; Conservative 0; Mismatches 55; Indels 6; Gaps 1;

```
Qy      957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016
            | |||| || ||||||||||||||||  ||||| || | |||| |||||| ||||| ||||
Db    62959 AGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCA 63018

Qy     1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076
            ||||||||||||||||||||||||||||||  ||    ||    || ||  | ||  |||
Db    63019 AACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTG 63078

Qy     1077 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACC 1130
            |||||| ||| | ||||||||||||||||||||||      ||||||||||| |||
Db    63079 CTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACA 63138

Qy     1131 TTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGA 1190
            ||||||| | |||||| |||||| |||| || ||||||||||||||||||||||||||| |
Db    63139 TTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTA 63198

Qy     1191 GTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGC 1250
            |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||  ||  ||
Db    63199 GTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGT 63258

Qy     1251 AGCTTT 1256
            | |||
Db    63259 GGGTTT 63264
```

RESULT 35
CR925830
LOCUS CR925830 167847 bp DNA linear PRI 08-JAN-2005
DEFINITION Human DNA sequence from clone RP11-195J19 on chromosome 6, complete sequence.
ACCESSION CR925830
VERSION CR925830.3 GI:56798018
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 167847)
  AUTHORS   Palmer,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 24, 2004 this sequence version replaced gi:56542150.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-195J19 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES              Location/Qualifiers
     source           1. .167847
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="6"
                      /clone="RP11-195J19"
                      /clone_lib="RPCI-11.1"
ORIGIN

  Query Match            14.9%;  Score 196;  DB 8;  Length 167847;
  Best Local Similarity   80.1%;  Pred. No. 1.6e-30;
  Matches  245;  Conservative    0;  Mismatches   55;  Indels    6;  Gaps    1;

Qy        957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016
              | |||| || |||||||||||||||  ||||| || | |||| |||||| ||||| ||||
Db      43202 AGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCA 43261

Qy       1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076
              |||||||||||||||||||||||||||||||  ||    ||    || ||  | ||  |||
Db      43262 AACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTG 43321

Qy       1077 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACC 1130
              |||||| ||| | ||||||||||||||||||||||||      |||||||||||| |||
Db      43322 CTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACA 43381

Qy       1131 TTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGA 1190
              ||||||| | |||||| |||||| |||| || ||||||||||||||||||||||||||| |
Db      43382 TTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTA 43441

Qy       1191 GTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGC 1250
              |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||  ||  ||

Db      43442 GTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGT 43501

Qy       1251 AGCTTT 1256
                 | |||
Db      43502 GGGTTT 43507


RESULT 36
CR942205
LOCUS       CR942205              176819 bp    DNA     linear   PRI 29-APR-2005
DEFINITION  Human DNA sequence from clone RP11-948M1 on chromosome 6, complete
            sequence.
ACCESSION   CR942205
VERSION     CR942205.4  GI:62989869
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 176819)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:62551221.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-948M1 is from the library RPCI-11.4 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES              Location/Qualifiers
     source           1. .176819
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="6"
                      /clone="RP11-948M1"
                      /clone_lib="RPCI-11.4"
ORIGIN

  Query Match            14.9%;  Score 196;  DB 8;  Length 176819;
  Best Local Similarity   80.1%;  Pred. No. 1.6e-30;
  Matches  245;  Conservative    0;  Mismatches   55;  Indels    6;  Gaps    1;
```

```
Qy        957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016
              | |||| || |||||||||||||||  ||||| || | |||| |||||| ||||| ||||
Db      37185 AGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCA 37244

Qy       1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076
              |||||||||||||||||||||||||||||||||||||  ||    ||     || ||  | ||  |||
Db      37245 AACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTG 37304

Qy       1077 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACC 1130
              |||||| ||| | |||||||||||||||||||||||||      |||||||||||| |||
Db      37305 CTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACA 37364

Qy       1131 TTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGA 1190
              ||||||| | |||||| |||||| |||| || |||||||||||||||||||||||||| |
Db      37365 TTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTA 37424

Qy       1191 GTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGC 1250
              |||||||||||| ||  |||||| |||  || || |||||| ||| ||||  ||  ||
Db      37425 GTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGT 37484

Qy       1251 AGCTTT 1256
                | |||
Db      37485 GGGTTT 37490
```

RESULT 37
BV203236

```
LOCUS       BV203236                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm211480 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV203236
VERSION     BV203236.1  GI:48172559
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN
```

```
  Query Match            11.1%;  Score 146.2;  DB 10;  Length 201;
  Best Local Similarity  82.6%;  Pred. No. 2.8e-20;
  Matches  166;  Conservative    1;  Mismatches   34;  Indels    0;  Gaps    0;

Qy        693 CTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAA 752
              ||||| | |||||||||  |||||||||||| ||||| || ||||||||||||||||||
Db          1 CTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAA 60

Qy        753 ATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTT 812
              |||||||||||||||||||||  |||| ||  |||||| : ||   ||||||||  |||
Db         61 GTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAYGTGTCAGGAAATGAGCTTT 120

Qy        813 GCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAAT 872
              |||||||||||||||||||||||||||||||  ||| ||||||||||||  |  || |
Db        121 GCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGT 180

Qy        873 TAAAAACTGTTCTTGGAAAAA 893
              ||||||  ||| |    ||||
Db        181 TAAAAATGGTTTTAAAGAAAA 201
```

RESULT 38
AX339803/c

```
LOCUS       AX339803                179 bp    DNA     linear   PAT 10-JAN-2002
DEFINITION  Sequence 50 from Patent WO0196388.
ACCESSION   AX339803
VERSION     AX339803.1  GI:18135784
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Harlocker,S.L. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196388-A 50 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .179
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match            10.3%;  Score 135.6;  DB 6;  Length 179;
  Best Local Similarity  84.4%;  Pred. No. 4.7e-18;
  Matches  151;  Conservative    1;  Mismatches   27;  Indels    0;  Gaps    0;

Qy        576 TGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAA 635
              | :| || |||||||||| |||||||||||||||||| |||||||| | |  ||||||||
Db        179 TNHACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAA 120

Qy        636 AGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTA 695
              |||||||| ||   | |||    |||||||||||| ||||||||||||||||||||||||
Db        119 AGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTA 60

Qy        696 TACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              || | |||||||||  |||||||||||| ||||| || |||||||||||||||||| |
Db         59 TAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 1
```

RESULT 39
AX351374

```
LOCUS       AX351374                176 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 121 from Patent WO0196390.
ACCESSION   AX351374
VERSION     AX351374.1  GI:18616721
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196390-A 121 20-DEC-2001;
            CORIXA CORPORATION (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .176
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             10.2%;  Score 134.4;  DB 6;  Length 176;
  Best Local Similarity   85.2%;  Pred. No. 8.5e-18;
  Matches  150;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy        579 ATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGA 638
              | || |||||||||| |||||||||||||||||||| |||||||| | |  ||||||||||
Db          1 ACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGA 60

Qy        639 AGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATAC 698
              ||||| ||   | |||     |||||||||||| ||||||||||||||||||||||||
Db         61 AGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAA 120

Qy        699 TATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              | ||||||||||  ||||||||||||| ||||| || ||||||||||||||||||| |
Db        121 TGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 176


RESULT 40
AR240556/c
LOCUS       AR240556                 268 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 312 from patent US 6468758.
ACCESSION   AR240556
VERSION     AR240556.1  GI:27285636
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 268)
  AUTHORS   Benson,D.R., Lodes,M.J., Mitcham,J.L. and King,G.E.
  TITLE     Compositions and methods for ovarian cancer therapy and diagnosis
  JOURNAL   Patent: US 6468758-A 312 22-OCT-2002;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .268
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             7.7%;  Score 101.4;  DB 6;  Length 268;
  Best Local Similarity   81.8%;  Pred. No. 8.2e-11;
  Matches  117;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1114 AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGA 1173
              |||||||||||| ||| ||||||| | |||||| |||||| |||| || ||||||||||
Db        262 AGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGA 203

Qy       1174 CTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTG 1233
              |||||||||||||| |||||||||||| ||  |||||| |||  || || ||||||
Db        202 CTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTA 143

Qy       1234 AAAAGCTTCCCACTTGCAGCTTT 1256
              ||| ||||  ||  ||  | |||
Db        142 AAATGCTTTTCAAGTGTGGGTTT 120


RESULT 41
HS133N16S/c
LOCUS       HS133N16S                411 bp    DNA     linear   STS 02-SEP-1999
DEFINITION  H.sapiens STS from genomic clone 133N16, sequence tagged site.
ACCESSION   AL110544
VERSION     AL110544.1  GI:5830260
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   Mungall,A.J., Huckle,E., Taylor,R. and Hunt,S.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-1999) E-mail contact: humquery@sanger.ac.uk
COMMENT     Marker stdJ133N16SP6 (Primer A : TGTTCCCTCTTTGCTGCAG; Primer B :
            AACAGTTCAAGGTTGTGGGC; amplimer size : 127 bp) is from sequence
            generated from the SP6 end of PAC 133N16. 133N16 is part of the
            bacterial clone contigs constructed by the Chromosome 6 Mapping
            Group.(http://www.sanger.ac.uk/HGP/Chr6/) 133N16 is from the
            library  constructed at the Roswell Park Cancer Institute by the
            group of Pieter de Jong.
            For further details see http://bacpac.med.buffalo.edu/.
FEATURES             Location/Qualifiers
     source          1. .411
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="133N16"
ORIGIN

  Query Match             7.3%;  Score 95.2;  DB 10;  Length 411;
  Best Local Similarity   78.0%;  Pred. No. 1.7e-09;
  Matches  128;  Conservative    0;  Mismatches   33;  Indels    3;  Gaps    1;

Qy        419 AACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTA 475
              || || ||  ||| |||| | |||| ||    ||  ||| |||||||  |   |||| | |
Db        411 AATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAA 352

Qy        476 GAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAAT 535
              ||| || ||||||   | ||| ||||||||||||||||||||||||||||  |  ||||
Db        351 GAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAAT 292

Qy        536 CTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTA 579
              ||| | |||||||||||||||||||||||| | ||||||||||
Db        291 CTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTA 248


RESULT 42
AC095717/c
LOCUS       AC095717              238494 bp    DNA     linear   HTG 13-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-9K13, WORKING DRAFT SEQUENCE.
ACCESSION   AC095717
VERSION     AC095717.6  GI:24940991
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 238494)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
```

```
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 238494)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 238494)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 13, 2002 this sequence version replaced gi:22773148.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GDDG
                Center clone name: CH230-9K13
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 221551 bases at least Q40
                Consensus quality: 223436 bases at least Q30
                Consensus quality: 224655 bases at least Q20
                Estimated insert size: 229610; sum-of-contigs estimation
                Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   238494: contig of 238494 bp in length.
FEATURES            Location/Qualifiers
     source         1. .238494
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-9K13"
     misc_feature   8872. .10696
                    /note="wgs_contig"
ORIGIN

  Query Match             7.2%;  Score 95;  DB 14;  Length 238494;
  Best Local Similarity   79.0%;  Pred. No. 3.5e-09;
  Matches  113;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy        958 AGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAA 1017
              | |||||||||||||||||||||||||| || ||| |||||||||||||||||| | |||||
Db     200730 ACTTATCAAATGTGGTGGTTTACCTGAAAATGTTCTTGATGCCTGGGAATTTGTGGGCAA 200671

Qy       1018 ACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGC 1077
              |||| |||||||||||||||||||    || ||||| |||||| || |||||| | |   |
Db     200670 ACCTGAACATTGCCAGTATACATGATTTATGAAAGTAAATAATAATCTCAGGTTACTGCC 200611

Qy       1078 TGCTTATAAGCATCGTTTTGATC 1100
              ||    ||    | ||| || ||
Db     200610 TGGGCCTATTTCTGGTTCTGCTC 200588


RESULT 43
BX957279
LOCUS       BX957279              198752 bp    DNA     linear   VRT 29-APR-2005
DEFINITION  Zebrafish DNA sequence from clone DKEY-218N20 in linkage group 19,
            complete sequence.
ACCESSION   BX957279
VERSION     BX957279.13  GI:62989971
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 198752)
  AUTHORS   Clark,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 30, 2005 this sequence version replaced gi:61673710.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
```

```
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-218N20
            is from a Zebrafish BAC library
            VECTOR: pIndigoBAC-5.
FEATURES             Location/Qualifiers
     source          1. .198752
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-218N20"
                     /clone_lib="DanioKey"
ORIGIN

  Query Match             6.3%;  Score 82.6;  DB 5;  Length 198752;
  Best Local Similarity   59.4%;  Pred. No. 1.4e-06;
  Matches  161;  Conservative    0;  Mismatches  104;  Indels    6;  Gaps    1;

Qy        957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016
              | ||   |||   ||| ||||| ||||     || | |||||  ||||||    | ||||
Db     148170 AGGTGGCCAATGTTGGAGGTTTGCCTGCAGGTGTGTGTGATGTGTGGGAACAGCTGGGCA 148229

Qy       1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076
              |||   |||| |||  ||| || ||||| || ||||| || | |||     |   | |||
Db     148230 AACAAGAACACTGCAGGTACACCTGGGACACCAAAGCTAACAGCAATAAGACTGTGCCTT 148289

Qy       1077 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGC------AGAAGAGGGGCACC 1130
                |     ||| || || || |||||  | ||  |      |    |     ||
Db     148290 ACGTCAGCAGATGTCGCTTCGACCGCATATTCCTGAGGTCTGCTAAAACCGCTCCGCCGG 148349

Qy       1131 TTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGA 1190
              | | |||| |     | | | |  ||||  |||| |||||||||| ||| | | | | |
Db     148350 TCACTCCTGACCACATGGCCTTGATTGGAATGGAGAAACTGGACTGCGGTCGTTATACCA 148409

Qy       1191 GTGATCACTGGGGGCTCTTGTGCACCTTGAA 1221
              ||||||||||||  | |  ||||| || ||
Db     148410 GTGATCACTGGGGTATTTACTGCACTTTCAA 148440


RESULT 44
BX511258
LOCUS       BX511258              200076 bp    DNA     linear   VRT 02-APR-2004
DEFINITION  Zebrafish DNA sequence from clone CH211-81E5 in linkage group 19,
            complete sequence.
ACCESSION   BX511258
VERSION     BX511258.6  GI:42820869
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 200076)
  AUTHORS   Heath,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 25, 2004 this sequence version replaced gi:42538803.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-81E5
            is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .200076
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-81E5"
                     /clone_lib="CHORI-211"
ORIGIN

  Query Match             6.3%;  Score 82.6;  DB 5;  Length 200076;
  Best Local Similarity   59.4%;  Pred. No. 1.4e-06;
  Matches  161;  Conservative    0;  Mismatches  104;  Indels    6;  Gaps    1;

Qy        957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016
              | ||   |||   ||| ||||| ||||     || | |||||  ||||||    | ||||
Db      90296 AGGTGGCCAATGTTGGAGGTTTGCCTGCAGGTGTGTGTGATGTGTGGGAACAGCTGGGCA 90355

Qy       1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076
              |||   |||| |||  ||| || ||||| || ||||| || | |||     |   | |||
Db      90356 AACAAGAACACTGCAGGTACACCTGGGACACCAAAGCTAACAGCAATAAGACTGTGCCTT 90415

Qy       1077 CTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACC------ 1130
                |     ||| || || || |||||  | ||  |  |    |     |  |
Db      90416 ACGTCAGCAGATGTCGCTTCGACCGCATATTCCTGAGGTCTGCTAAAACCCGTCCAGCGG 90475

Qy       1131 TTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGA 1190
              | | |||| |     | | | |  ||||  |||| |||||||||| ||| | | | | |
Db      90476 TCACTCCTGACCACATGGCCTTGATTGGAATGGAGAAACTGGACTGCGGTCGTTATACCA 90535

Qy       1191 GTGATCACTGGGGGCTCTTGTGCACCTTGAA 1221
              ||||||||||||  | |  ||||| || ||
Db      90536 GTGATCACTGGGGTATTTACTGCACTTTCAA 90566


RESULT 45
```

```
AC156673/c
LOCUS       AC156673               231985 bp    DNA     linear   HTG 01-JUL-2005
DEFINITION  Bos taurus clone CH240-59M13, *** SEQUENCING IN PROGRESS ***, 21
            unordered pieces.
ACCESSION   AC156673
VERSION     AC156673.2  GI:68266455
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 231985)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 231985)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 231985)
  AUTHORS   Cow Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jun 28, 2005 this sequence version replaced gi:58531393.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: FDJG
                Center clone name: CH240-59M13
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 224220 bases at least Q40
                Consensus quality: 226361 bases at least Q30
                Consensus quality: 227836 bases at least Q20
                Estimated insert size: 228962; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 21 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3298: contig of 3298 bp in length
            *     3299     3348: gap of 50 bp
            *     3349     5753: contig of 2405 bp in length
            *     5754     6081: gap of 328 bp
            *     6082     7706: contig of 1625 bp in length
            *     7707     7806: gap of unknown length
            *     7807    24881: contig of 17075 bp in length
            *    24882    24931: gap of 50 bp
            *    24932    53019: contig of 28088 bp in length
            *    53020    53069: gap of 50 bp
            *    53070    56864: contig of 3795 bp in length
            *    56865    56914: gap of 50 bp
            *    56915    64212: contig of 7298 bp in length
            *    64213    64400: gap of 188 bp
            *    64401    74558: contig of 10158 bp in length
            *    74559    74608: gap of 50 bp
            *    74609    98289: contig of 23681 bp in length
            *    98290    98389: gap of unknown length
            *    98390   100834: contig of 2445 bp in length
            *   100835   100884: gap of 50 bp
            *   100885   132970: contig of 32086 bp in length
            *   132971   133113: gap of 143 bp
            *   133114   134462: contig of 1349 bp in length
            *   134463   134562: gap of unknown length
            *   134563   141930: contig of 7368 bp in length
            *   141931   141980: gap of 50 bp
            *   141981   186048: contig of 44068 bp in length
            *   186049   186098: gap of 50 bp
            *   186099   203328: contig of 17230 bp in length
            *   203329   203378: gap of 50 bp
            *   203379   205841: contig of 2463 bp in length
```

```
            *   205842   205941: gap of unknown length
            *   205942   207514: contig of 1573 bp in length
            *   207515   207564: gap of 50 bp
            *   207565   212129: contig of 4565 bp in length
            *   212130   212979: gap of 850 bp
            *   212980   214003: contig of 1024 bp in length
            *   214004   214103: gap of unknown length
            *   214104   215330: contig of 1227 bp in length
            *   215331   215430: gap of unknown length
            *   215431   231985: contig of 16555 bp in length.
FEATURES             Location/Qualifiers
     source          1. .231985
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9913"
                     /clone="CH240-59M13"
     gap             3299. .3348
                     /estimated_length=50
     gap             5754. .6081
                     /estimated_length=328
     gap             7707. .7806
                     /estimated_length=unknown
     gap             24882. .24931
                     /estimated_length=50
     gap             53020. .53069
                     /estimated_length=50
     gap             56865. .56914
                     /estimated_length=50
     gap             64213. .64400
                     /estimated_length=188
     gap             74559. .74608
                     /estimated_length=50
     gap             98290. .98389
                     /estimated_length=unknown
     gap             100835. .100884
                     /estimated_length=50
     gap             132971. .133113
                     /estimated_length=143
     gap             134463. .134562
                     /estimated_length=unknown
     gap             141931. .141980
                     /estimated_length=50
     gap             186049. .186098
                     /estimated_length=50
     gap             203329. .203378
                     /estimated_length=50
     gap             205842. .205941
                     /estimated_length=unknown
     gap             207515. .207564
                     /estimated_length=50
     gap             212130. .212979
                     /estimated_length=850
     gap             214004. .214103
                     /estimated_length=unknown
     gap             215331. .215430

  Query Match             5.8%;  Score 76.6;  DB 14;  Length 231985;
  Best Local Similarity   65.5%;  Pred. No. 2.7e-05;
  Matches  112;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy       1092 GTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACC 1151
             | | ||  | ||   |   | |||||      |      |||||||   |||||| || |
Db       1530 GGTGTGTCCTAACTGTGAACTGAGCACTGCTAGTAGCGGTTATTCCCTGAAGTTTGGATC 1471

Qy       1152 TTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGT 1211
             |  | ||| ||||||||||||||||| |||||||||| ||||||||||||  ||  |||
Db       1470 TGCTAGGGCTGGAAAAACTGGACTGTAGTAGATTTCCCAGTGATCACTGGGATCTTCTGT 1411

Qy       1212 GCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGT 1262
             | | |||  ||||| || ||||||    || | | |   || |||  || |
Db       1410 GAAACTTAGATGTAATACTGTGAACCCTTTTCAAATATGAGTTTTCAACAT 1360
```

```
RESULT 46
BD077628
LOCUS       BD077628                 258 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  5'EST of secreted protein expressed in muscles and other mesodermal
            tissues.
ACCESSION   BD077628
VERSION     BD077628.1  GI:22623231
KEYWORDS    JP 2001512016-A/214.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 258)
  AUTHORS   Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
  TITLE     5'EST of secreted protein expressed in muscles and other mesodermal
            tissues
  JOURNAL   Patent: JP 2001512016-A 214 21-AUG-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001512016-A/214
            PD   21-AUG-2001
            PF   31-JUL-1998 JP 2000505295
            PR   01-AUG-1997 US    08/905134
            PI   JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO  PI
            LACROIX
            PC   C12N15/09,C12N15/09,C07K14/47,C12M1/00,C12N15/00,C12N15/00 CC
             blastn
            CC   identity 97
            CC   region 1. .180
            CC   id T31666
            CC   est
            CC   blastn
            CC   identity 92
            CC   region 88. .141
            CC   id R58665
            CC   est
            CC   blastn
            CC   identity 90
            CC   region 39. .93
            CC   id R58665
            CC   est
            CC   blastn
            CC   identity 98
            CC   region 63. .137
            CC   id R14990
            CC   est
            CC   blastn
            CC   identity 94
            CC   region 1. .50
            CC   id R14990
            CC   est
            CC   blastn
            CC   identity 99
            CC   region 1. .120
            CC   id T26956
            CC   est
            CC   Von Heijne matrix
            CC   score 6
            CC   seq LCXEFXSVASCDA/AV
            FH   Key             Location/Qualifiers
            FT   misc_feature    75. .254
            FT   misc_feature    73. .126
            FT   misc_feature    23. .77
            FT   misc_feature    157. .231
            FT   misc_feature    95. .144
            FT   misc_feature    135. .254
            FT   sig_peptide     31. .150.
FEATURES             Location/Qualifiers
     source          1. .258
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
```

```
ORIGIN

  Query Match             5.7%;  Score 75.2;  DB 6;  Length 258;
  Best Local Similarity   58.7%;  Pred. No. 2.8e-05;
  Matches  122;  Conservative    4;  Mismatches   82;  Indels    0;  Gaps    0;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    | |   |  || |   | | ||  |||     ||||| | |||||
Db         47 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 106

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| ||| :|:|  |||:| | ||| |   | |||||  || |  |||  |   ||
Db        107 GACTTCTGTGTSTRGAGTTTRCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 166

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   ||:|||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        167 TCCTGGCCGAKAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 226

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCC 378
              | | ||    |  | |   |||| ||
Db        227 TGGAGGAGAGCGCCTTGGAACGCCGACC 254


RESULT 47
AX939201
LOCUS       AX939201                 258 bp    DNA     linear   PAT 07-JAN-2004
DEFINITION  Sequence 220 from Patent EP1367124.
ACCESSION   AX939201
VERSION     AX939201.1  GI:40733556
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   dumas milne Edwards,J.B., Duclert,A. and Lacroix,B.
  TITLE     5' ests for secreted proteins expressed in muscle and other
            mesodermal tissues
  JOURNAL   Patent: EP 1367124-A 220 03-DEC-2003;
            GENSET (FR)
FEATURES             Location/Qualifiers
     source          1. .258
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /tissue_type="KIDNEY"
                     /dev_stage="FETAL"
     sig_peptide     31. .150
ORIGIN

  Query Match             5.7%;  Score 75.2;  DB 6;  Length 258;
  Best Local Similarity   58.7%;  Pred. No. 2.8e-05;
  Matches  122;  Conservative    4;  Mismatches   82;  Indels    0;  Gaps    0;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    | |   |  || |   | | ||  |||     ||||| | |||||
Db         47 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 106

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| ||| :|:|  |||:| | ||| |   | |||||  || |  |||  |   ||
Db        107 GACTTCTGTGTSTRGAGTTTRCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 166

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   ||:|||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        167 TCCTGGCCGAKAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 226

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCC 378
              | | ||    |  | |   |||| ||
Db        227 TGGAGGAGAGCGCCTTGGAACGCCGACC 254


RESULT 48
CQ069444
LOCUS       CQ069444                 444 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 5244 from Patent WO0157278.
ACCESSION   CQ069444
VERSION     CQ069444.1  GI:41039313
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 5244 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL031775.1
                     EXPRESSED IN HELA, SIGNAL = 1.9"
ORIGIN

  Query Match             5.7%;  Score 74.6;  DB 6;  Length 444;
  Best Local Similarity   71.3%;  Pred. No. 3.9e-05;
  Matches  112;  Conservative    0;  Mismatches   44;  Indels    1;  Gaps    1;

Qy        577 GTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAA 636
              ||| || ||||||||| ||||||||||||||||||| |||||||| | |  |||||||||
Db        253 GTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAA 312

Qy        637 GAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTAT 696
              ||||||| ||    | |||    |||||||||||||  |    | ||||  ||    | |
Db        313 GAAGAGATCAAGTAATTATGAGATTATTACAGGTAACAATTTTTGTTATTCAACTTGTCT 372

Qy        697 ACT-ATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 732
                |  ||||| ||  ||  |||||||   | || |||
Db        373 TTTCTTTGAAAAACAGATCAGTGAAAGAGATAAATCA 409


RESULT 49
CQ078617
LOCUS       CQ078617                 691 bp    DNA     linear   PAT 20-JAN-2004
DEFINITION  Sequence 14417 from Patent WO0157278.
ACCESSION   CQ078617
VERSION     CQ078617.1  GI:41048486
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Penn,S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
  TITLE     Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human hela cells or other human
            cervical epithelialcells
  JOURNAL   Patent: WO 0157278-A 14417 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="MAP TO AL031775.1
                     EXPRESSED IN HELA, SIGNAL = 1.9
                     SWISSPROT HIT: P18004, EVALUE 7.00e-01
                     NT HIT: AF201687.1, EVALUE 3.00e-45
                     EST_HUMAN HIT: AV717253.1, EVALUE 5.00e-45"
```

```
ORIGIN

  Query Match             5.7%;  Score 74.6;  DB 6;  Length 691;
  Best Local Similarity   71.3%;  Pred. No. 4.1e-05;
  Matches  112;  Conservative    0;  Mismatches   44;  Indels    1;  Gaps    1;

Qy        577 GTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAA 636
              ||| || ||||||||| |||||||||||||||||||| |||||||| | |  |||||||||
Db        115 GTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAA 174

Qy        637 GAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTAT 696
              ||||||| ||   | |||    ||||||||||||||  |     | ||||  ||   | ||
Db        175 GAAGAGATCAAGTAATTATGAGATTATTACAGGTAACAATTTTTGTTATTCAACTTGTCT 234

Qy        697 ACT-ATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 732
                |  ||||| ||  ||  |||||||   | || |||
Db        235 TTTCTTTGAAAAACAGATCAGTGAAAGAGATAAATCA 271


RESULT 50
AX525593
LOCUS       AX525593                705 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 115 from Patent WO02066682.
ACCESSION   AX525593
VERSION     AX525593.1  GI:25170471
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 115 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .705
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             5.1%;  Score 67.4;  DB 6;  Length 705;
  Best Local Similarity   60.8%;  Pred. No. 0.0014;
  Matches  110;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
              ||  |||| || ||||| | |   |      |    |||      |    |  ||| | |
Db         76 GGACAGCAACGGGAAGACCAGCCATGAGTCAAGTCAGCTGGATGAACCAAGCGTGGAAAC 135

Qy        137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
              |      ||  || | |  | | ||      |||| | |   | |    | || |  ||||
Db        136 ACACAGCCTGGAGCAGTCCAAGGAGTATAAGCAGAGGGCCAGCCACGAGAGCACTGAGCA 195

Qy        197 G 197
              |
Db        196 G 196


RESULT 51
AX525862
LOCUS       AX525862                631 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 384 from Patent WO02066682.
ACCESSION   AX525862
VERSION     AX525862.1  GI:25170740
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 384 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .631
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.9%;  Score 64.8;  DB 6;  Length 631;
  Best Local Similarity   78.0%;  Pred. No. 0.0048;
  Matches   78;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 77

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTG 116
              ||| | || |   | |   | |  |  |||  |||| ||
Db         78 TGGATGAAGAGAACATATGAGCTTGTGAGAGAAGACCGTG 117


RESULT 52
AX525757
LOCUS       AX525757                690 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 279 from Patent WO02066682.
ACCESSION   AX525757
VERSION     AX525757.1  GI:25170635
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 279 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .690
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.9%;  Score 63.8;  DB 6;  Length 690;
  Best Local Similarity   61.8%;  Pred. No. 0.0078;
  Matches  118;  Conservative    0;  Mismatches   72;  Indels    1;  Gaps    1;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGAT-GGCGTCTGGCAG 135
              | |    ||| | | |  |      ||  |||| | ||  |  |  || | |
Db         77 CGAACCTCGCGGGGCGGGAAGCCGCGAGATGGACACCCCTCCGCTCTCAGACTCGGACTC 136

Qy        136 CAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGC 195
              | | || |||| |     | || |||  |  |     || ||  ||||   | | |||
Db        137 CGGGTCGGATGAGTGCCTGGCCTCAGATCAAGAGTTGCAGGATGCGTTTTCCCGCGGACT 196

Qy        196 AGCACAGGCGG 206
                | || | |
Db        197 CCTAAAGCCAG 207
```

```
RESULT 53
AX525766
LOCUS       AX525766                696 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 288 from Patent WO02066682.
ACCESSION   AX525766
VERSION     AX525766.1  GI:25170644
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 288 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .696
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.8%;  Score 63.2;  DB 6;  Length 696;
  Best Local Similarity   84.5%;  Pred. No. 0.01;
  Matches   71;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              | |||||| || |    |  || ||
Db         76 GAGAAGCTGCCTTCTCATTGGCTG 99


RESULT 54
AX525570
LOCUS       AX525570                704 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 92 from Patent WO02066682.
ACCESSION   AX525570
VERSION     AX525570.1  GI:25170448
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 92 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .704
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.8%;  Score 63.2;  DB 6;  Length 704;
  Best Local Similarity   84.5%;  Pred. No. 0.01;
  Matches   71;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAT 77

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              ||| |   ||| || |||   |||
Db         78 GGGTATGTGCGGGATGAGACAGTG 101


RESULT 55
AX525719
LOCUS       AX525719                722 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 241 from Patent WO02066682.
ACCESSION   AX525719
VERSION     AX525719.1  GI:25170597
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 241 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .722
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.8%;  Score 62.8;  DB 6;  Length 722;
  Best Local Similarity   69.7%;  Pred. No. 0.013;
  Matches   85;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGG 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
                 || |  |    |||    |  |    ||| |  | || ||   |||| |||  |
Db         77 CCGCCGCCGTCTCTGCAGCCCACGGGTCCCTGGGCCGTTCCGCCGCCCGCGTGCGGCCCC 136

Qy        137 AG 138
               |
Db        137 GG 138


RESULT 56
AX525877
LOCUS       AX525877                630 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 399 from Patent WO02066682.
ACCESSION   AX525877
VERSION     AX525877.1  GI:25170755
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 399 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.7%;  Score 62.2;  DB 6;  Length 630;
  Best Local Similarity   76.8%;  Pred. No. 0.017;
  Matches   76;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGTG 79
```

```
Qy          77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCT 115
               |  | ||  |  |   |||||||   |||||   ||
Db          80 ATGTTGGAGAAGGCTGCCGGGGTGTTGCTAGGGGCGGCT 118


RESULT 57
AR216294
LOCUS       AR216294                1549 bp    DNA     linear   PAT 25-SEP-2002
DEFINITION  Sequence 5 from patent US 6410713.
ACCESSION   AR216294
VERSION     AR216294.1  GI:23314759
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1549)
  AUTHORS   Guerriero,V. and Raynes,D.A.
  TITLE     DNA encoding proteins that inhibit Hsp70 function
  JOURNAL   Patent: US 6410713-A 5 25-JUN-2002;
FEATURES             Location/Qualifiers
     source          1. .1549
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             4.7%;  Score 62.2;  DB 6;  Length 1549;
  Best Local Similarity   95.5%;  Pred. No. 0.018;
  Matches   64;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGCGG 78
               || ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db           2 TTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCCGG 61

Qy          79 GAAGCAG 85
               |||||||
Db          62 CAAGCAG 68


RESULT 58
AR680067
LOCUS       AR680067                1549 bp    DNA     linear   PAT 13-JUN-2005
DEFINITION  Sequence 5 from patent US 6903202.
ACCESSION   AR680067
VERSION     AR680067.1  GI:67622738
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1549)
  AUTHORS   Guerriero,V. Jr. and Raynes,D.A.
  TITLE     DNA encoding proteins that inhibit Hsp70 function
  JOURNAL   Patent: US 6903202-A 5 07-JUN-2005;
            Desert Genetics, Inc.; Tucson, AZ
FEATURES             Location/Qualifiers
     source          1. .1549
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             4.7%;  Score 62.2;  DB 6;  Length 1549;
  Best Local Similarity   95.5%;  Pred. No. 0.018;
  Matches   64;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          19 TTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCGG 78
               || ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db           2 TTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCCGG 61

Qy          79 GAAGCAG 85
               |||||||
Db          62 CAAGCAG 68


RESULT 59
AX525524
LOCUS       AX525524                701 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 46 from Patent WO02066682.
ACCESSION   AX525524
VERSION     AX525524.1  GI:25170402
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 46 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .701
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.7%;  Score 61.6;  DB 6;  Length 701;
  Best Local Similarity   56.4%;  Pred. No. 0.023;
  Matches  115;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy          17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
               | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76

Qy          77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
               || |   |      || |   |  |   |   |   ||| |   | | |   |
Db          77 GGAACCATGGCCAGCCCGCTGCGCTCCTTGATGCTACTGCTGGCCGTCCTGGCCGTGGCC 136

Qy         137 AGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCA 196
               |  |||   |  |   |||  |   ||        | | ||  ||  || | | || ||
Db         137 TGGGCCGGAACCTCCAGGCCACCCCCGCGATTGTTGGGAGCTCCGCAGGAGGCAGATGCC 196

Qy         197 GCACAGGCGGAGGAGGACCGGGTG 220
               |||  ||  | | | || | |
Db         197 AGCGAGGAGGGCGTGCAGCGAGCG 220


RESULT 60
AX525842
LOCUS       AX525842                718 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 364 from Patent WO02066682.
ACCESSION   AX525842
VERSION     AX525842.1  GI:25170720
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 364 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .718
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.7%;  Score 61.6;  DB 6;  Length 718;
  Best Local Similarity   88.2%;  Pred. No. 0.023;
  Matches   67;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;
```

```
Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || |||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCC 76

Qy         77 GGGAAGCAGCGTGAAG 92
                |||  |||   | || |
Db         77 TGGAGGCAAGGGGATG 92


RESULT 61
AX525494/c
LOCUS       AX525494                682 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 16 from Patent WO02066682.
ACCESSION   AX525494
VERSION     AX525494.1  GI:25170372
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 16 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .682
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.7%;  Score 61.4;  DB 6;  Length 682;
  Best Local Similarity   81.6%;  Pred. No. 0.025;
  Matches   71;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          7 TAATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCG 66
              | ||  |  | | || ||||||||||||||||||||||||||||||||||||||||||||
Db        672 TGATTACGCAATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCG 613

Qy         67 GCACGAGGGCGGGAAGCAGCGTGAAGA 93
              |||||||||| | | |   |   || |||
Db        612 GCACGAGGGTGAGGAAGAAGATGGAGA 586


RESULT 62
AX525612
LOCUS       AX525612                701 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 134 from Patent WO02066682.
ACCESSION   AX525612
VERSION     AX525612.1  GI:25170490
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 134 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .701
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.7%;  Score 61.2;  DB 6;  Length 701;
  Best Local Similarity   84.1%;  Pred. No. 0.028;
  Matches   69;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGG 98
              || |  ||  | |   ||| ||
Db         77 GGCAGCCATGGCGCCCAGCCGG 98


RESULT 63
AX224570
LOCUS       AX224570                956 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 52 from Patent WO0161006.
ACCESSION   AX224570
VERSION     AX224570.1  GI:15554762
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Pausch,M.H.
  TITLE     Two pore potassium channels, nucleotide sequences encoding them,
            and methods of using same
  JOURNAL   Patent: WO 0161006-A 52 23-AUG-2001;
            Basf Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .956
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN

  Query Match             4.7%;  Score 61.2;  DB 6;  Length 956;
  Best Local Similarity   95.5%;  Pred. No. 0.029;
  Matches   63;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          9 ATGATTCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 68
              ||||| ||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGATACGATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGC 60

Qy         69 ACGAGG 74
              ||||||
Db         61 ACGAGG 66


RESULT 64
AX525594/c
LOCUS       AX525594                581 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 116 from Patent WO02066682.
ACCESSION   AX525594
VERSION     AX525594.1  GI:25170472
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 116 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.6%;  Score 60.6;  DB 6;  Length 581;
```

```
  Best Local Similarity   79.1%;  Pred. No. 0.036;
  Matches   72;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCT 505

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAG 107
              |   || |||     |  | | | || ||||
Db        504 GAAGAGAAGCACCCTGCACAGTTCTTCTGAG 474


RESULT 65
AX525588
LOCUS       AX525588                692 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 110 from Patent WO02066682.
ACCESSION   AX525588
VERSION     AX525588.1  GI:25170466
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 110 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .692
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.6%;  Score 60.4;  DB 6;  Length 692;
  Best Local Similarity   91.4%;  Pred. No. 0.041;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 77

Qy         77 GGGAAGCAGC 86
              ||  ||||
Db         78 TAGAGCCAGC 87


RESULT 66
AX525704
LOCUS       AX525704                706 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 226 from Patent WO02066682.
ACCESSION   AX525704
VERSION     AX525704.1  GI:25170582
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 226 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .706
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.6%;  Score 60.4;  DB 6;  Length 706;
  Best Local Similarity   91.4%;  Pred. No. 0.041;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGG 76

Qy         77 GGGAAGCAGC 86
              |||| |||
Db         77 AAGAAGAAGC 86


RESULT 67
AX525660
LOCUS       AX525660                657 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 182 from Patent WO02066682.
ACCESSION   AX525660
VERSION     AX525660.1  GI:25170538
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 182 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.6%;  Score 60.2;  DB 6;  Length 657;
  Best Local Similarity   89.0%;  Pred. No. 0.045;
  Matches   65;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 77

Qy         77 GGGAAGCAGCGTG 89
              | | ||  | | |
Db         78 GCGCAGGGGAGCG 90


RESULT 68
AX525724
LOCUS       AX525724                669 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 246 from Patent WO02066682.
ACCESSION   AX525724
VERSION     AX525724.1  GI:25170602
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 246 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .669
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN
```

```
  Query Match             4.6%;  Score 60;  DB 6;  Length 669;
  Best Local Similarity   82.1%;  Pred. No. 0.049;
  Matches   69;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 77

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              |    | |||  || || |  | |
Db         78 GATGTGGAGCTCGATGACCTAGAG 101


RESULT 69
AX525746
LOCUS       AX525746                710 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 268 from Patent WO02066682.
ACCESSION   AX525746
VERSION     AX525746.1  GI:25170624
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 268 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .710
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.6%;  Score 60;  DB 6;  Length 710;
  Best Local Similarity   82.1%;  Pred. No. 0.05;
  Matches   69;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCG 88

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTG 100
              ||||  | ||  ||   |  || |
Db         89 GGGAGCCGGCTGGAGACGGTGGGG 112


RESULT 70
AX525583/c
LOCUS       AX525583                594 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 105 from Patent WO02066682.
ACCESSION   AX525583
VERSION     AX525583.1  GI:25170461
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 105 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .594
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.6%;  Score 59.8;  DB 6;  Length 594;
  Best Local Similarity   84.8%;  Pred. No. 0.054;
  Matches   67;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 512

Qy         77 GGGAAGCAGCGTGAAGAGC 95
              |  |  |||  ||| |
Db        511 ACCACCCCTCGTCTAGACC 493


RESULT 71
AX525652
LOCUS       AX525652                762 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 174 from Patent WO02066682.
ACCESSION   AX525652
VERSION     AX525652.1  GI:25170530
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 174 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .762
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.6%;  Score 59.8;  DB 6;  Length 762;
  Best Local Similarity   80.5%;  Pred. No. 0.055;
  Matches   70;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 79

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTT 103
              |||  || | |  || |   |||
Db         80 ATTCAGCCTCGCGGACGGCTGTCATTT 106


RESULT 72
AX525581
LOCUS       AX525581                720 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 103 from Patent WO02066682.
ACCESSION   AX525581
VERSION     AX525581.1  GI:25170459
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 103 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .720
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
```

```
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.5%;  Score 59.6;  DB 6;  Length 720;
  Best Local Similarity   62.2%;  Pred. No. 0.06;
  Matches   92;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | |   ||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         17 ATTNANTACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAC 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGC 136
              ||| |   | | | ||| |     ||| || | ||   |   |   |   |   |
Db         77 AGGGACAGATGGGCTGTGCGCGCTGCGCGAGCTGAGCGTGGACCTTCGAGCAGAGCGCTC 136

Qy        137 AGTTCCGATGCGGCGGAGCCCGCAGGGC 164
              ||| |  || | |  || |   || | |
Db        137 AGTGCTCATCCCGGAGACCTACCAAGCC 164


RESULT 73
AX525860
LOCUS       AX525860                761 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 382 from Patent WO02066682.
ACCESSION   AX525860
VERSION     AX525860.1  GI:25170738
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 382 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .761
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.5%;  Score 59.2;  DB 6;  Length 761;
  Best Local Similarity   79.5%;  Pred. No. 0.074;
  Matches   70;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGT 75

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTT 104
              | ||   |  ||| |    ||   || |
Db         76 GTGACTGATGGTGGAAGCAGGCCTTTCT 103


RESULT 74
AX525728
LOCUS       AX525728                686 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 250 from Patent WO02066682.
ACCESSION   AX525728
VERSION     AX525728.1  GI:25170606
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 250 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .686
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.5%;  Score 59;  DB 6;  Length 686;
  Best Local Similarity   78.0%;  Pred. No. 0.08;
  Matches   71;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         17 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCC 76

Qy         77 GGGAAGCAGCGTGAAGAGCGGGTGTTTTGAG 107
              | |  ||  | |    ||  |||| ||
Db         77 TGCAGTGAGAAAGTGGCAAAGGGCTTTTCAG 107


RESULT 75
AX525495/c
LOCUS       AX525495                691 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 17 from Patent WO02066682.
ACCESSION   AX525495
VERSION     AX525495.1  GI:25170373
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 17 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match             4.5%;  Score 59;  DB 6;  Length 691;
  Best Local Similarity   92.5%;  Pred. No. 0.081;
  Matches   62;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         17 AATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGC 76
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        675 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGGA 616

Qy         77 GGGAAGC 83
              | ||| |
Db        615 GAGAACC 609


Search completed: March 11, 2006, 12:48:46
Job time : 6926 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 03:44:14 ; Search time 26.0462 Seconds
(without alignments)
797.263 Million cell updates/sec

Title: US-10-757-745-2_COPY_145_153
Perfect score: 45
Sequence: 1 PDVIFLQEV 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7673375 seqs, 1153648444 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170339_4330/app_query.fasta_1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02h
-USER=US10757745_@CGN_1_1_1676_@runat_09032006_170339_4330 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	38	84.4	1509	6	US-09-925-065A-33485	Sequence 33485, A
c	2	38	84.4	1611	8	US-10-750-185-54202	Sequence 54202, A
c	3	38	84.4	1611	8	US-10-750-623-54202	Sequence 54202, A
c	4	37	82.2	637	6	US-09-925-065A-497337	Sequence 497337,
c	5	37	82.2	1179	9	US-11-096-568A-30375	Sequence 30375, A
	6	37	82.2	208765	7	US-10-330-773-407	Sequence 407, App
	7	36	80.0	201	8	US-10-995-561-4650	Sequence 4650, Ap
	8	36	80.0	201	8	US-10-995-561-4654	Sequence 4654, Ap
	9	36	80.0	201	8	US-10-995-561-4660	Sequence 4660, Ap
	10	36	80.0	201	8	US-10-995-561-4664	Sequence 4664, Ap
	11	36	80.0	201	8	US-10-995-561-27410	Sequence 27410, A
	12	36	80.0	201	8	US-10-995-561-27508	Sequence 27508, A
	13	36	80.0	606	6	US-09-925-065A-704289	Sequence 704289,
	14	36	80.0	1152	6	US-09-925-065A-77580	Sequence 77580, A
	15	36	80.0	1515	7	US-10-932-182A-82471	Sequence 82471, A
	16	36	80.0	1515	7	US-10-932-182A-82471	Sequence 82471, A
c	17	36	80.0	2764	12	US-11-136-527-3829	Sequence 3829, Ap
	18	36	80.0	3100	8	US-10-775-169-254	Sequence 254, App
	19	36	80.0	37500	8	US-10-522-037-1	Sequence 1, Appli
c	20	35	77.8	201	12	US-11-124-367A-14480	Sequence 14480, A
	21	35	77.8	604	6	US-09-925-065A-790230	Sequence 790230,
c	22	35	77.8	617	6	US-09-925-065A-953853	Sequence 953853,
c	23	35	77.8	930	12	US-11-136-527-3369	Sequence 3369, Ap
c	24	35	77.8	930	12	US-11-136-527-7465	Sequence 7465, Ap
	25	35	77.8	1542	7	US-10-932-182A-3163	Sequence 3163, Ap
	26	35	77.8	1542	7	US-10-932-182A-3163	Sequence 3163, Ap
c	27	35	77.8	2637	8	US-10-750-185-64239	Sequence 64239, A
c	28	35	77.8	2637	8	US-10-750-623-64239	Sequence 64239, A
c	29	35	77.8	156297	12	US-11-121-086-65	Sequence 65, Appl
	30	35	77.8	175023	12	US-11-121-086-18	Sequence 18, Appl
c	31	34	75.6	200	9	US-11-261-135-3	Sequence 3, Appli
c	32	34	75.6	415	6	US-09-925-065A-574404	Sequence 574404,
	33	34	75.6	417	6	US-09-925-065A-574441	Sequence 574441,
c	34	34	75.6	419	6	US-09-925-065A-226520	Sequence 226520,
c	35	34	75.6	419	6	US-09-925-065A-226521	Sequence 226521,
c	36	34	75.6	474	6	US-09-925-065A-562814	Sequence 562814,
	37	34	75.6	495	6	US-09-925-065A-344600	Sequence 344600,
	38	34	75.6	495	6	US-09-925-065A-344601	Sequence 344601,
c	39	34	75.6	496	6	US-09-925-065A-229262	Sequence 229262,
c	40	34	75.6	496	6	US-09-925-065A-229263	Sequence 229263,
	41	34	75.6	506	6	US-09-925-065A-20548	Sequence 20548, A
	42	34	75.6	510	6	US-09-925-065A-21557	Sequence 21557, A
	43	34	75.6	510	6	US-09-925-065A-21558	Sequence 21558, A
c	44	34	75.6	513	6	US-09-925-065A-437773	Sequence 437773,
c	45	34	75.6	541	6	US-09-925-065A-48038	Sequence 48038, A
c	46	34	75.6	541	6	US-09-925-065A-48039	Sequence 48039, A
c	47	34	75.6	541	6	US-09-925-065A-48040	Sequence 48040, A
c	48	34	75.6	541	6	US-09-925-065A-48041	Sequence 48041, A
c	49	34	75.6	541	6	US-09-925-065A-48042	Sequence 48042, A
	50	34	75.6	560	6	US-09-925-065A-649710	Sequence 649710,
c	51	34	75.6	562	6	US-09-925-065A-627003	Sequence 627003,
c	52	34	75.6	562	6	US-09-925-065A-627004	Sequence 627004,
c	53	34	75.6	562	6	US-09-925-065A-627005	Sequence 627005,
c	54	34	75.6	566	6	US-09-925-065A-908784	Sequence 908784,
c	55	34	75.6	569	6	US-09-925-065A-442932	Sequence 442932,
c	56	34	75.6	569	6	US-09-925-065A-442933	Sequence 442933,
c	57	34	75.6	570	6	US-09-925-065A-262283	Sequence 262283,
	58	34	75.6	582	6	US-09-925-065A-325943	Sequence 325943,
c	59	34	75.6	583	6	US-09-925-065A-686631	Sequence 686631,
c	60	34	75.6	589	6	US-09-925-065A-393309	Sequence 393309,
c	61	34	75.6	593	6	US-09-925-065A-655382	Sequence 655382,
c	62	34	75.6	600	8	US-10-750-185-20586	Sequence 20586, A
c	63	34	75.6	600	8	US-10-750-623-20586	Sequence 20586, A
	64	34	75.6	602	6	US-09-925-065A-619704	Sequence 619704,
	65	34	75.6	606	6	US-09-925-065A-619706	Sequence 619706,
c	66	34	75.6	607	6	US-09-925-065A-461094	Sequence 461094,
c	67	34	75.6	607	6	US-09-925-065A-461095	Sequence 461095,
c	68	34	75.6	607	6	US-09-925-065A-461096	Sequence 461096,
c	69	34	75.6	609	6	US-09-925-065A-262281	Sequence 262281,
c	70	34	75.6	609	6	US-09-925-065A-262282	Sequence 262282,
c	71	34	75.6	609	6	US-09-925-065A-262284	Sequence 262284,
c	72	34	75.6	609	6	US-09-925-065A-262285	Sequence 262285,
	73	34	75.6	610	6	US-09-925-065A-770547	Sequence 770547,
	74	34	75.6	610	6	US-09-925-065A-770548	Sequence 770548,
	75	34	75.6	610	6	US-09-925-065A-770549	Sequence 770549,

ALIGNMENTS

```
RESULT 1
US-09-925-065A-33485/c
; Sequence 33485, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33485
;   LENGTH: 1509
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-33485

Alignment Scores:
Pred. No.:                295             Length:         1509
Score:                    38.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   3
Best Local Similarity:    66.7%           Mismatches:     0
Query Match:              84.4%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-33485 (1-1509)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||:::|||:::|||
Db        662 CCGGATGTGGTTTTCATTCAGAAAGTA 636

RESULT 2
US-10-750-185-54202/c
; Sequence 54202, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 54202
;   LENGTH: 1611
;   TYPE: DNA
;   ORGANISM: Bovine  19866881264862
US-10-750-185-54202

Alignment Scores:
Pred. No.:                321             Length:         1611
Score:                    38.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              84.4%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-750-185-54202 (1-1611)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||   ||||||:::||||||
Db        338 CCTGATGTTTTTTTCCTTGAAGAAGTT 312

RESULT 3
US-10-750-623-54202/c
; Sequence 54202, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 54202
;   LENGTH: 1611
;   TYPE: DNA
;   ORGANISM: Bovine  19866881264862
US-10-750-623-54202

Alignment Scores:
Pred. No.:                321             Length:         1611
Score:                    38.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              84.4%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-750-623-54202 (1-1611)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||   ||||||:::||||||
Db        338 CCTGATGTTTTTTTCCTTGAAGAAGTT 312

RESULT 4
US-09-925-065A-497337/c
; Sequence 497337, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
```

```
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 497337
;   LENGTH: 637
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-497337

Alignment Scores:
Pred. No.:                171          Length:         637
Score:                    37.00        Matches:        6
Percent Similarity:       100.0%       Conservative:   2
Best Local Similarity:    75.0%        Mismatches:     0
Query Match:              82.2%        Indels:         0
DB:                       6            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-497337 (1-637)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||:::::|||||||||
Db        466 CCAGATGTGGTATATTTGCAGGAA 443

RESULT 5
US-11-096-568A-30375/c
; Sequence 30375, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30375
;   LENGTH: 1179
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1179)
;   OTHER INFORMATION: Ceres Seq. ID no. 4957757
US-11-096-568A-30375

Alignment Scores:
Pred. No.:                369          Length:         1179
Score:                    37.00        Matches:        6
Percent Similarity:       100.0%       Conservative:   2
Best Local Similarity:    75.0%        Mismatches:     0
Query Match:              82.2%        Indels:         0
DB:                       9            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-096-568A-30375 (1-1179)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||:::|||||||||:::
Db        472 CCGGATGTAGTTTTCTTGCAAGAC 449

RESULT 6
US-10-330-773-407
; Sequence 407, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 407
;   LENGTH: 208765
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(208765)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-407

Alignment Scores:
Pred. No.:                2.41e+05     Length:         208765
Score:                    37.00        Matches:        6
Percent Similarity:       100.0%       Conservative:   2
Best Local Similarity:    75.0%        Mismatches:     0
Query Match:              82.2%        Indels:         0
DB:                       7            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-330-773-407 (1-208765)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||||||:::
Db     113412 CCAGACATAATCTTTCTGCAGCAG 113435

RESULT 7
US-10-995-561-4650
; Sequence 4650, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4650
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-4650

Alignment Scores:
Pred. No.:                68.4         Length:         201
Score:                    36.00        Matches:        6
Percent Similarity:       87.5%        Conservative:   1
Best Local Similarity:    75.0%        Mismatches:     1
Query Match:              80.0%        Indels:         0
DB:                       8            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-995-561-4650 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24

RESULT 8
US-10-995-561-4654
; Sequence 4654, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4654
;   LENGTH: 201
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-10-995-561-4654

Alignment Scores:
Pred. No.:                68.4          Length:         201
Score:                    36.00         Matches:        6
Percent Similarity:       87.5%         Conservative:   1
Best Local Similarity:    75.0%         Mismatches:     1
Query Match:              80.0%         Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-995-561-4654 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129

RESULT 9
US-10-995-561-4660
; Sequence 4660, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4660
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-4660

Alignment Scores:
Pred. No.:                68.4          Length:         201
Score:                    36.00         Matches:        6
Percent Similarity:       87.5%         Conservative:   1
Best Local Similarity:    75.0%         Mismatches:     1
Query Match:              80.0%         Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-995-561-4660 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24

RESULT 10
US-10-995-561-4664
; Sequence 4664, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4664
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-4664

Alignment Scores:
Pred. No.:                68.4          Length:         201
Score:                    36.00         Matches:        6
Percent Similarity:       87.5%         Conservative:   1
Best Local Similarity:    75.0%         Mismatches:     1
Query Match:              80.0%         Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-995-561-4664 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129

RESULT 11
US-10-995-561-27410
; Sequence 27410, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27410
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-27410

Alignment Scores:
Pred. No.:                68.4          Length:         201
Score:                    36.00         Matches:        6
Percent Similarity:       87.5%         Conservative:   1
Best Local Similarity:    75.0%         Mismatches:     1
Query Match:              80.0%         Indels:         0
DB:                       8             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-995-561-27410 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24

RESULT 12
US-10-995-561-27508
; Sequence 27508, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27508
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-27508

Alignment Scores:
Pred. No.:                68.4          Length:         201
Score:                    36.00         Matches:        6
Percent Similarity:       87.5%         Conservative:   1
Best Local Similarity:    75.0%         Mismatches:     1
Query Match:              80.0%         Indels:         0
DB:                       8             Gaps:           0
```

```
US-10-757-745-2_COPY_145_153 (1-9) x US-10-995-561-27508 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129

RESULT 13
US-09-925-065A-704289
; Sequence 704289, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 704289
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-704289

Alignment Scores:
Pred. No.:                273             Length:         606
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    75.0%           Mismatches:     0
Query Match:              80.0%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-704289 (1-606)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||:::|||:::
Db         27 CCTGATGTAATCTTCATCCAGCAG 50

RESULT 14
US-09-925-065A-77580
; Sequence 77580, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 77580
;   LENGTH: 1152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1152)
;   OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-77580

Alignment Scores:
Pred. No.:                610             Length:         1152
Score:                    36.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
Best Local Similarity:    66.7%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-77580 (1-1152)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||   :::
Db        865 CCTGATGTAATGTTTCTACAGCATATA 891

RESULT 15
US-10-932-182A-82471
; Sequence 82471, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
;  CURRENT FILING DATE:  2004-09-02
;  NUMBER OF SEQ ID NOS: 197023
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82471
;   LENGTH: 1515
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-10-932-182A-82471

Alignment Scores:
Pred. No.:                860             Length:         1515
Score:                    36.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-932-182A-82471 (1-1515)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   ||||||||||
Db        454 CCTGATGTAGTATCGTTGCAGGAGGTC 480

RESULT 16
US-10-932-182A-82471
; Sequence 82471, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAO, YOSHIHIRO
;  APPLICANT:  NAKAMURA, NORIHISA
;  APPLICANT:  KODAMA, YUKIKO
;  APPLICANT:  FUJIMURA, TOMOKO
;  APPLICANT:  ASHIKARI, TOSHIHIKO
;  TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;  FILE REFERENCE: 030685-043
;  CURRENT APPLICATION NUMBER: US/10/932,182A
```

; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82471
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-82471

Alignment Scores:

Pred. No.:	860	Length:	1515
Score:	36.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-932-182A-82471 (1-1515)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||:::   ||||||||||||
Db        454 CCTGATGTAGTATCGTTGCAGGAGGTC 480
```

RESULT 17
US-11-136-527-3829/c
; Sequence 3829, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3829
; LENGTH: 2764
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3829

Alignment Scores:

Pred. No.:	1.83e+03	Length:	2764
Score:	36.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	87.5%	Mismatches:	0
Query Match:	80.0%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-136-527-3829 (1-2764)

```
Qy          2 AspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||||||||
Db        325 GATGTTCTTTTTCTGCAGGAAGTC 302
```

RESULT 18
US-10-775-169-254
; Sequence 254, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dorner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 254
; LENGTH: 3100
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-254

Alignment Scores:

Pred. No.:	2.11e+03	Length:	3100
Score:	36.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-775-169-254 (1-3100)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        732 CCAGACGCTGTTTTCCTGCAAGAA 755
```

RESULT 19
US-10-522-037-1
; Sequence 1, Application US/10522037
; Publication No. US20050282166A1
; GENERAL INFORMATION:
; APPLICANT: LIBRAGEN
; TITLE OF INVENTION: Method for the expression of unknown environmental DNA into adapt.
; TITLE OF INVENTION: cells.
; FILE REFERENCE: B0149WO
; CURRENT APPLICATION NUMBER: US/10/522,037
; CURRENT FILING DATE: 2005-01-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 37500
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence of clone FS3-124.
US-10-522-037-1

Alignment Scores:

Pred. No.:	4.81e+04	Length:	37500
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	80.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-522-037-1 (1-37500)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||:::||||||||||:::
Db    29025 CCGGACGTCTTGTTCCTGCAGCAG 29048
```

RESULT 20
US-11-124-367A-14480/c
; Sequence 14480, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,846
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554

```
;  PRIOR FILING DATE: 2004-08-09
;  NUMBER OF SEQ ID NOS: 34460
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14480
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-124-367A-14480

Alignment Scores:
Pred. No.:                116          Length:         201
Score:                    35.00        Matches:        6
Percent Similarity:       77.8%        Conservative:   1
Best Local Similarity:    66.7%        Mismatches:     2
Query Match:              77.8%        Indels:         0
DB:                       12           Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-124-367A-14480 (1-201)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||   |||||||||   :::
Db        188 CCTGATGTATTTTTCCTCCAAAATATT 162

RESULT 21
US-09-925-065A-790230
; Sequence 790230, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 790230
;   LENGTH: 604
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-790230

Alignment Scores:
Pred. No.:                462          Length:         604
Score:                    35.00        Matches:        5
Percent Similarity:       100.0%       Conservative:   3
Best Local Similarity:    62.5%        Mismatches:     0
Query Match:              77.8%        Indels:         0
DB:                       6            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-790230 (1-604)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||::::::|||||||||:::
Db        241 CCAGATATTTTATTTCTTCAGGAT 264

RESULT 22
US-09-925-065A-953853/c
; Sequence 953853, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 953853
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-953853

Alignment Scores:
Pred. No.:                475          Length:         617
Score:                    35.00        Matches:        6
Percent Similarity:       100.0%       Conservative:   1
Best Local Similarity:    85.7%        Mismatches:     0
Query Match:              77.8%        Indels:         0
DB:                       6            Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-953853 (1-617)

Qy          1 ProAspValIlePheLeuGln 7
              ||||||:::|||||||||||
Db        477 CCAGATATTATTTTTTTGCAA 457

RESULT 23
US-11-136-527-3369/c
; Sequence 3369, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3369
;   LENGTH: 930
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-3369

Alignment Scores:
Pred. No.:                794          Length:         930
Score:                    35.00        Matches:        6
Percent Similarity:       77.8%        Conservative:   1
Best Local Similarity:    66.7%        Mismatches:     2
Query Match:              77.8%        Indels:         0
DB:                       12           Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-136-527-3369 (1-930)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||   |||   ||||||:::
Db        407 CCGGATGTATTTTTCACCCAAGAAATT 381

RESULT 24
US-11-136-527-7465/c
```

; Sequence 7465, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7465
; LENGTH: 930
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-7465

Alignment Scores:

Pred. No.:	794	Length:	930
Score:	35.00	Matches:	6
Percent Similarity:	77.8%	Conservative:	1
Best Local Similarity:	66.7%	Mismatches:	2
Query Match:	77.8%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-136-527-7465 (1-930)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||   |||   ||||||:::
Db        407 CCGGATGTATTTTTCACCCAAGAAATT 381
```

RESULT 25
US-10-932-182A-3163
; Sequence 3163, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3163
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-3163

Alignment Scores:

Pred. No.:	1.5e+03	Length:	1542
Score:	35.00	Matches:	6
Percent Similarity:	88.9%	Conservative:	2
Best Local Similarity:	66.7%	Mismatches:	1
Query Match:	77.8%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-932-182A-3163 (1-1542)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::   ||||||||||||
Db        481 CCCGATATTGTATCATTACAAGAAGTC 507
```

RESULT 26
US-10-932-182A-3163
; Sequence 3163, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3163
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-3163

Alignment Scores:

Pred. No.:	1.5e+03	Length:	1542
Score:	35.00	Matches:	6
Percent Similarity:	88.9%	Conservative:	2
Best Local Similarity:	66.7%	Mismatches:	1
Query Match:	77.8%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-932-182A-3163 (1-1542)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::   ||||||||||||
Db        481 CCCGATATTGTATCATTACAAGAAGTC 507
```

RESULT 27
US-10-750-185-64239/c
; Sequence 64239, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 64239
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Bovine 19866881160446
US-10-750-185-64239

Alignment Scores:

Pred. No.:	2.93e+03	Length:	2637
Score:	35.00	Matches:	4
Percent Similarity:	100.0%	Conservative:	5
Best Local Similarity:	44.4%	Mismatches:	0
Query Match:	77.8%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-750-185-64239 (1-2637)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::||||||:::::::::::::
Db       1843 CCCGACATTATTTTTATGGAACAGATC 1817
```

RESULT 28

US-10-750-623-64239/c
; Sequence 64239, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 64239
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Bovine 19866881160446
US-10-750-623-64239

Alignment Scores:

Pred. No.:	2.93e+03	Length:	2637
Score:	35.00	Matches:	4
Percent Similarity:	100.0%	Conservative:	5
Best Local Similarity:	44.4%	Mismatches:	0
Query Match:	77.8%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-750-623-64239 (1-2637)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::||||||:::::::::::
Db       1843 CCCGACATTATTTTTATGGAACAGATC 1817
```

RESULT 29
US-11-121-086-65/c
; Sequence 65, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 65
; LENGTH: 156297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-65

Alignment Scores:

Pred. No.:	4.82e+05	Length:	156297
Score:	35.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	77.8%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-121-086-65 (1-156297)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||:::::|||||||||
Db      42551 CCAGATGTGATGTATTTACAGGAG 42528
```

RESULT 30
US-11-121-086-18
; Sequence 18, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18
; LENGTH: 175023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-18

Alignment Scores:

Pred. No.:	5.54e+05	Length:	175023
Score:	35.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	77.8%	Indels:	0
DB:	12	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-121-086-18 (1-175023)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||||||||   |||||||||:::
Db     161675 CCCGATGTCATCAAGTTGCAAGAGATT 161701
```

RESULT 31
US-11-261-135-3/c
; Sequence 3, Application US/11261135
; Publication No. US20060041948A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Keith D.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING CHEMOKINE
; TITLE OF INVENTION: RECEPTOR 9A GENE DISRUPTIONS
; FILE REFERENCE: R-365
; CURRENT APPLICATION NUMBER: US/11/261,135
; CURRENT FILING DATE: 2005-10-27
; PRIOR APPLICATION NUMBER: US/09/903,377
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/217,255
; PRIOR FILING DATE: 2000-07-10
; PRIOR APPLICATION NUMBER: US 60/221,483
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: US 60/262,113
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Targeting vector
US-11-261-135-3

Alignment Scores:

Pred. No.:	197	Length:	200
Score:	34.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	75.6%	Indels:	0
DB:	9	Gaps:	0

```
US-10-757-745-2_COPY_145_153 (1-9) x US-11-261-135-3 (1-200)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::|||||||||   |||
Db        123 CCTGACATTATTTTCTTACAGAA 100

RESULT 32
US-09-925-065A-574404/c
; Sequence 574404, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 574404
;   LENGTH: 415
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-574404

Alignment Scores:
Pred. No.:              492           Length:         415
Score:                  34.00         Matches:        5
Percent Similarity:     100.0%        Conservative:   4
Best Local Similarity:  55.6%         Mismatches:     0
Query Match:            75.6%         Indels:         0
DB:                     6             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-574404 (1-415)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        134 CCCAATGTGATATTTATTCAACAAATA 108

RESULT 33
US-09-925-065A-574441
; Sequence 574441, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 574441
;   LENGTH: 417
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-574441

Alignment Scores:
Pred. No.:              495           Length:         417
Score:                  34.00         Matches:        5
Percent Similarity:     100.0%        Conservative:   4
Best Local Similarity:  55.6%         Mismatches:     0
Query Match:            75.6%         Indels:         0
DB:                     6             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-574441 (1-417)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        284 CCCAATGTGATATTTATTCAACAAATA 310

RESULT 34
US-09-925-065A-226520/c
; Sequence 226520, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 226520
;   LENGTH: 419
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-226520

Alignment Scores:
Pred. No.:              498           Length:         419
Score:                  34.00         Matches:        5
Percent Similarity:     100.0%        Conservative:   4
Best Local Similarity:  55.6%         Mismatches:     0
Query Match:            75.6%         Indels:         0
DB:                     6             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-226520 (1-419)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        134 CCCAATGTGATATTTATTCAACAAATA 108

RESULT 35
US-09-925-065A-226521/c
; Sequence 226521, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
```

```
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 226521
;   LENGTH: 419
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-226521

Alignment Scores:
Pred. No.:                498             Length:         419
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-226521 (1-419)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        134 CCCAATGTGATATTTATTCAACAAATA 108

RESULT 36
US-09-925-065A-562814/c
; Sequence 562814, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 562814
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-562814

Alignment Scores:
Pred. No.:                581             Length:         474
Score:                    34.00           Matches:        4
Percent Similarity:       100.0%          Conservative:   5
Best Local Similarity:    44.4%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-562814 (1-474)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::::::::|||||:::|||:::
Db        239 CCAGAAATTCTTTTCTTAGAAGAAATT 213

RESULT 37
US-09-925-065A-344600
; Sequence 344600, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 344600
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-344600

Alignment Scores:
Pred. No.:                613             Length:         495
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-344600 (1-495)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        161 CCCAATGTGATATTTATTCAACAAATA 187

RESULT 38
US-09-925-065A-344601
; Sequence 344601, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 344601
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-344601
```

```
Alignment Scores:
Pred. No.:                613             Length:         495
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-344601 (1-495)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||:::::::
Db        161 CCCAATGTGATATTTATTCAACAAATA 187

RESULT 39
US-09-925-065A-229262/c
; Sequence 229262, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 229262
;   LENGTH: 496
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-229262

Alignment Scores:
Pred. No.:                615             Length:         496
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-229262 (1-496)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||:::::::
Db        125 CCCAATGTGATATTTATTCAACAAATA 99

RESULT 40
US-09-925-065A-229263/c
; Sequence 229263, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 229263
;   LENGTH: 496
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-229263

Alignment Scores:
Pred. No.:                615             Length:         496
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-229263 (1-496)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||:::::::
Db        125 CCCAATGTGATATTTATTCAACAAATA 99

RESULT 41
US-09-925-065A-20548
; Sequence 20548, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20548
;   LENGTH: 506
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-20548

Alignment Scores:
Pred. No.:                630             Length:         506
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-20548 (1-506)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||:::::::
Db        443 CCCAATGTGATATTTATTCAACAAATA 469

RESULT 42
US-09-925-065A-21557
```

```
; Sequence 21557, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21557
;   LENGTH: 510
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-21557

Alignment Scores:
Pred. No.:                637             Length:         510
Score:                    34.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-21557 (1-510)

Qy          2 AspValIlePheLeuGlnGlu 8
              ||||||||||||||||||||
Db        201 GATGTTATTTTCTTGCAAGAG 221

RESULT 43
US-09-925-065A-21558
; Sequence 21558, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21558
;   LENGTH: 510
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-21558

Alignment Scores:
Pred. No.:                637             Length:         510
Score:                    34.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-21558 (1-510)

Qy          2 AspValIlePheLeuGlnGlu 8
              ||||||||||||||||||||
Db        201 GATGTTATTTTCTTGCAAGAG 221

RESULT 44
US-09-925-065A-437773/c
; Sequence 437773, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 437773
;   LENGTH: 513
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-437773

Alignment Scores:
Pred. No.:                641             Length:         513
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-437773 (1-513)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        353 CCCAATGTGATATTTATTCAACAAATA 327

RESULT 45
US-09-925-065A-48038/c
; Sequence 48038, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
```

; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48038
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-48038

Alignment Scores:

Pred. No.:	685	Length:	541
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-48038 (1-541)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        378 CCCAATGTGATATTTATTCAACAAATA 352
```

RESULT 46
US-09-925-065A-48039/c
; Sequence 48039, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48039
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-48039

Alignment Scores:

Pred. No.:	685	Length:	541
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-48039 (1-541)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        378 CCCAATGTGATATTTATTCAACAAATA 352
```

RESULT 47
US-09-925-065A-48040/c
; Sequence 48040, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48040
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-48040

Alignment Scores:

Pred. No.:	685	Length:	541
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-48040 (1-541)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        378 CCCAATGTGATATTTATTCAACAAATA 352
```

RESULT 48
US-09-925-065A-48041/c
; Sequence 48041, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48041
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-48041

Alignment Scores:

Pred. No.:	685	Length:	541
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-48041 (1-541)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        378 CCCAATGTGATATTTATTCAACAAATA 352
```

RESULT 49
US-09-925-065A-48042/c
; Sequence 48042, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48042
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-48042

Alignment Scores:

Pred. No.:	685	Length:	541
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-48042 (1-541)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        378 CCCAATGTGATATTTATTCAACAAATA 352
```

RESULT 50
US-09-925-065A-649710
; Sequence 649710, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 649710
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-649710

Alignment Scores:

Pred. No.:	716	Length:	560
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-649710 (1-560)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::|||||||||::::::
Db        446 CCAGATCTCCTTTTCTTACAGGATTTA 472
```

RESULT 51
US-09-925-065A-627003/c
; Sequence 627003, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 627003
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-627003

Alignment Scores:

Pred. No.:	719	Length:	562
Score:	34.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	85.7%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-627003 (1-562)

```
Qy          1 ProAspValIlePheLeuGln 7
              |||||||||:::|||||||||
Db        459 CCAGATGTCCTGTTCCTACAA 439
```

RESULT 52
US-09-925-065A-627004/c
; Sequence 627004, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A

```
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 627004
;   LENGTH: 562
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-627004

Alignment Scores:
Pred. No.:                719             Length:         562
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-627004 (1-562)

Qy          1 ProAspValIlePheLeuGln 7
              ||||||||||:::|||||||||
Db        459 CCAGATGTCCTGTTCCTACAA 439

RESULT 53
US-09-925-065A-627005/c
; Sequence 627005, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 627005
;   LENGTH: 562
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-627005

Alignment Scores:
Pred. No.:                719             Length:         562
Score:                    34.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    85.7%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-627005 (1-562)

Qy          1 ProAspValIlePheLeuGln 7
              ||||||||||:::|||||||||
Db        459 CCAGATGTCCTGTTCCTACAA 439

RESULT 54
US-09-925-065A-908784/c
; Sequence 908784, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 908784
;   LENGTH: 566
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-908784

Alignment Scores:
Pred. No.:                725             Length:         566
Score:                    34.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
Best Local Similarity:    66.7%           Mismatches:     1
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-908784 (1-566)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||   ::::::||||||||||||||||
Db        559 CCCAGAATTTTATTTTTGCAGGAAGTA 533

RESULT 55
US-09-925-065A-442932/c
; Sequence 442932, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 442932
;   LENGTH: 569
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-09-925-065A-442932

Alignment Scores:

Pred. No.:	730	Length:	569
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-442932 (1-569)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        231 CCCAATGTGATATTTATTCAACAAATA 205
```

RESULT 56
US-09-925-065A-442933/c
; Sequence 442933, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 442933
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-442933

Alignment Scores:

Pred. No.:	730	Length:	569
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-442933 (1-569)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        231 CCCAATGTGATATTTATTCAACAAATA 205
```

RESULT 57
US-09-925-065A-262283/c
; Sequence 262283, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 262283
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-262283

Alignment Scores:

Pred. No.:	732	Length:	570
Score:	34.00	Matches:	5
Percent Similarity:	100.0%	Conservative:	4
Best Local Similarity:	55.6%	Mismatches:	0
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-262283 (1-570)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        232 CCCAATGTGATATTTATTCAACAAATA 206
```

RESULT 58
US-09-925-065A-325943
; Sequence 325943, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 325943
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-325943

Alignment Scores:

Pred. No.:	751	Length:	582
Score:	34.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	75.6%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-325943 (1-582)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||   |||:::||||||
Db        298 CCTGATGTCATCACATTGGAGGAGGTA 324
```

RESULT 59

```
US-09-925-065A-686631/c
; Sequence 686631, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 686631
;   LENGTH: 583
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-686631

Alignment Scores:
Pred. No.:                753             Length:         583
Score:                    34.00           Matches:        7
Percent Similarity:       77.8%           Conservative:   0
Best Local Similarity:    77.8%           Mismatches:     2
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-686631 (1-583)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||   ||||||||||||   |||
Db        426 CCTGATTTCATCTTCCTCCAGCCTGTG 400

RESULT 60
US-09-925-065A-393309/c
; Sequence 393309, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 393309
;   LENGTH: 589
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-393309

Alignment Scores:
Pred. No.:                763             Length:         589
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-393309 (1-589)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||::::::||||||:::|||::: 
Db         48 CCAGATCTTGTGTTTCTCAGAGAGCTC 22

RESULT 61
US-09-925-065A-655382/c
; Sequence 655382, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 655382
;   LENGTH: 593
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-655382

Alignment Scores:
Pred. No.:                769             Length:         593
Score:                    34.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   0
Best Local Similarity:    100.0%          Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-655382 (1-593)

Qy          2 AspValIlePheLeuGlnGlu 8
              |||||||||||||||||||||
Db         68 GATGTTATTTTCTTGCAAGAG 48

RESULT 62
US-10-750-185-20586/c
; Sequence 20586, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
```

```
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 20586
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Bovine  MMBT16526
US-10-750-185-20586

Alignment Scores:
Pred. No.:                780             Length:         600
Score:                    34.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              75.6%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-750-185-20586 (1-600)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        242 CCTGACTCAGTATTTTTACAGGAG 219

RESULT 63
US-10-750-623-20586/c
; Sequence 20586, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 20586
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Bovine  MMBT16526
US-10-750-623-20586

Alignment Scores:
Pred. No.:                780             Length:         600
Score:                    34.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              75.6%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-750-623-20586 (1-600)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        242 CCTGACTCAGTATTTTTACAGGAG 219

RESULT 64
US-09-925-065A-619704
; Sequence 619704, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 619704
;   LENGTH: 602
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-619704

Alignment Scores:
Pred. No.:                784             Length:         602
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-619704 (1-602)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||:::::
Db        445 CCCAATGTGATATTTATTCAACAAATA 471

RESULT 65
US-09-925-065A-619706
; Sequence 619706, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 619706
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-619706

Alignment Scores:
Pred. No.:                790             Length:         606
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-619706 (1-606)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||:::::
```

```
Db        497 CCCAATGTGATATTTATTCAACAAATA 523

RESULT 66
US-09-925-065A-461094/c
; Sequence 461094, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 461094
;   LENGTH: 607
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-461094

Alignment Scores:
Pred. No.:                792             Length:         607
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-461094 (1-607)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        125 CCCAATGTGATATTTATTCAACAAATA 99

RESULT 67
US-09-925-065A-461095/c
; Sequence 461095, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 461095
;   LENGTH: 607
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-461095

Alignment Scores:
Pred. No.:                792             Length:         607
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-461095 (1-607)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        125 CCCAATGTGATATTTATTCAACAAATA 99

RESULT 68
US-09-925-065A-461096/c
; Sequence 461096, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 461096
;   LENGTH: 607
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-461096

Alignment Scores:
Pred. No.:                792             Length:         607
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-461096 (1-607)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        125 CCCAATGTGATATTTATTCAACAAATA 99

RESULT 69
US-09-925-065A-262281/c
; Sequence 262281, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
```

```
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 262281
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(609)
;   OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-262281

Alignment Scores:
Pred. No.:                795             Length:         609
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-262281 (1-609)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        165 CCCAATGTGATATTTATTCAACAAATA 139

RESULT 70
US-09-925-065A-262282/c
; Sequence 262282, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 262282
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(609)
;   OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-262282

Alignment Scores:
Pred. No.:                795             Length:         609
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-262282 (1-609)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        165 CCCAATGTGATATTTATTCAACAAATA 139

RESULT 71
US-09-925-065A-262284/c
; Sequence 262284, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 262284
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(609)
;   OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-262284

Alignment Scores:
Pred. No.:                795             Length:         609
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-262284 (1-609)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||::::::
Db        165 CCCAATGTGATATTTATTCAACAAATA 139

RESULT 72
US-09-925-065A-262285/c
; Sequence 262285, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
```

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 262285
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(609)
;   OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-262285

Alignment Scores:
Pred. No.:                795             Length:         609
Score:                    34.00           Matches:        5
Percent Similarity:       100.0%          Conservative:   4
Best Local Similarity:    55.6%           Mismatches:     0
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-262285 (1-609)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::|||||||||:::|||:::::::
Db        165 CCCAATGTGATATTTATTCAACAAATA 139

RESULT 73
US-09-925-065A-770547
; Sequence 770547, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 770547
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-770547

Alignment Scores:
Pred. No.:                797             Length:         610
Score:                    34.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
Best Local Similarity:    66.7%           Mismatches:     1
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-770547 (1-610)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::::::|||   ||||||||||||
Db        577 CCTGAAATAATTTTATTACAAGAAGTT 603

RESULT 74
US-09-925-065A-770548
; Sequence 770548, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 770548
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-770548

Alignment Scores:
Pred. No.:                797             Length:         610
Score:                    34.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
Best Local Similarity:    66.7%           Mismatches:     1
Query Match:              75.6%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-770548 (1-610)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::::::|||   ||||||||||||
Db        577 CCTGAAATAATTTTATTACAAGAAGTT 603

RESULT 75
US-09-925-065A-770549
; Sequence 770549, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 770549
;   LENGTH: 610
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-770549

Alignment Scores:
Pred. No.:                797             Length:         610
Score:                    34.00           Matches:        6
Percent Similarity:       88.9%           Conservative:   2
```

Best Local Similarity: 66.7% Mismatches: 1
Query Match: 75.6% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-770549 (1-610)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||:::::|||   ||||||||||||
Db        577 CCTGAAATAATTTTATTACAAGAAGTT 603
```

Search completed: March 11, 2006, 12:42:29
Job time : 54.0462 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 11, 2006, 00:16:23 ; Search time 46.8465 Seconds
(without alignments)
1588.688 Million cell updates/sec

Title: US-10-757-745-2_COPY_145_153
Perfect score: 45
Sequence: 1 PDVIFLQEV 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10757745/runat_09032006_170337_4279/app_query.fasta_1
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=75 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss08
-USER=US10757745_@CGN_1_1_2715_@runat_09032006_170337_4279 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	45	100.0	176	3	US-09-878-722-121	Sequence 121, App
c	2	45	100.0	176	3	US-09-878-178-50	Sequence 50, Appl
	3	45	100.0	176	3	US-09-904-456-121	Sequence 121, App
c	4	45	100.0	176	5	US-10-046-935-50	Sequence 50, Appl
c	5	45	100.0	176	5	US-10-146-502-50	Sequence 50, Appl
c	6	45	100.0	176	5	US-10-060-036-1367	Sequence 1367, Ap
	7	45	100.0	401	5	US-10-066-543-1363	Sequence 1363, Ap
	8	45	100.0	444	3	US-09-864-761-16762	Sequence 16762, A
	9	45	100.0	553	5	US-10-066-543-983	Sequence 983, App
	10	45	100.0	625	5	US-10-066-543-1031	Sequence 1031, Ap
	11	45	100.0	633	5	US-10-066-543-1084	Sequence 1084, Ap
	12	45	100.0	644	5	US-10-066-543-1070	Sequence 1070, Ap
	13	45	100.0	674	5	US-10-066-543-1067	Sequence 1067, Ap
	14	45	100.0	691	3	US-09-864-761-33229	Sequence 33229, A
c	15	45	100.0	1079	3	US-09-745-288-63	Sequence 63, Appl
c	16	45	100.0	1079	7	US-10-453-919-63	Sequence 63, Appl
	17	45	100.0	1088	9	US-10-450-763-22058	Sequence 22058, A
	18	45	100.0	1296	3	US-09-925-299-170	Sequence 170, App
	19	45	100.0	1296	3	US-09-925-299-170	Sequence 170, App
	20	45	100.0	1920	9	US-10-757-745-1	Sequence 1, Appli
	21	45	100.0	1940	9	US-10-783-271-28	Sequence 28, Appl
	22	45	100.0	1948	5	US-10-037-270-889	Sequence 889, App
	23	45	100.0	1948	6	US-10-117-722-889	Sequence 889, App
	24	45	100.0	1948	9	US-10-122-851-889	Sequence 889, App
	25	45	100.0	3152	3	US-09-981-353-88	Sequence 88, Appl
	26	44	97.8	1312	9	US-10-757-745-3	Sequence 3, Appli
	27	42	93.3	581	7	US-10-437-963-29447	Sequence 29447, A
	28	40	88.9	625	5	US-10-066-543-548	Sequence 548, App
	29	40	88.9	1224	7	US-10-437-963-99622	Sequence 99622, A
c	30	39	86.7	600	9	US-10-972-079-46608	Sequence 46608, A
	31	38	84.4	912	3	US-09-738-626-746	Sequence 746, App
c	32	38	84.4	1509	4	US-09-925-065A-33485	Sequence 33485, A
	33	38	84.4	1838	7	US-10-425-114-25133	Sequence 25133, A
	34	38	84.4	2209	5	US-10-027-632-103398	Sequence 103398,
	35	38	84.4	2209	6	US-10-027-632-103398	Sequence 103398,
	36	38	84.4	2468	8	US-10-425-115-149831	Sequence 149831,
	37	38	84.4	5337	7	US-10-437-963-81672	Sequence 81672, A
	38	38	84.4	3309400	3	US-09-738-626-1	Sequence 1, Appli
c	39	37	82.2	516	3	US-09-770-961-110	Sequence 110, App
c	40	37	82.2	637	4	US-09-925-065A-497337	Sequence 497337,
	41	37	82.2	753	7	US-10-424-599-123282	Sequence 123282,
c	42	37	82.2	768	6	US-10-342-224-27	Sequence 27, Appl
	43	37	82.2	774	6	US-10-369-493-45451	Sequence 45451, A
	44	37	82.2	777	6	US-10-369-493-31729	Sequence 31729, A
	45	37	82.2	852	6	US-10-369-493-47104	Sequence 47104, A
	46	37	82.2	1350	10	US-11-097-143-9713	Sequence 9713, Ap
	47	37	82.2	1628	7	US-10-424-599-43531	Sequence 43531, A
	48	36	80.0	201	7	US-10-741-601-1929	Sequence 1929, Ap
	49	36	80.0	201	7	US-10-741-601-1933	Sequence 1933, Ap
	50	36	80.0	201	7	US-10-741-601-1938	Sequence 1938, Ap
	51	36	80.0	201	7	US-10-741-601-1942	Sequence 1942, Ap
	52	36	80.0	201	7	US-10-741-601-11767	Sequence 11767, A
	53	36	80.0	201	7	US-10-741-601-11806	Sequence 11806, A
	54	36	80.0	201	8	US-10-741-600-7337	Sequence 7337, Ap
	55	36	80.0	201	8	US-10-741-600-7341	Sequence 7341, Ap
	56	36	80.0	201	8	US-10-741-600-7346	Sequence 7346, Ap
	57	36	80.0	201	8	US-10-741-600-7350	Sequence 7350, Ap
	58	36	80.0	201	8	US-10-741-600-29526	Sequence 29526, A
	59	36	80.0	201	8	US-10-741-600-29565	Sequence 29565, A
	60	36	80.0	245	3	US-09-922-293-3628	Sequence 3628, Ap
	61	36	80.0	254	3	US-09-922-293-3627	Sequence 3627, Ap
	62	36	80.0	492	5	US-10-027-632-60370	Sequence 60370, A
	63	36	80.0	492	5	US-10-027-632-107582	Sequence 107582,
	64	36	80.0	492	6	US-10-027-632-60370	Sequence 60370, A
	65	36	80.0	492	6	US-10-027-632-107582	Sequence 107582,
c	66	36	80.0	559	8	US-10-425-115-27387	Sequence 27387, A
c	67	36	80.0	583	7	US-10-437-963-26483	Sequence 26483, A
	68	36	80.0	606	4	US-09-925-065A-704289	Sequence 704289,
	69	36	80.0	714	7	US-10-424-599-156	Sequence 156, App
	70	36	80.0	717	7	US-10-282-122A-35952	Sequence 35952, A
	71	36	80.0	1051	3	US-09-882-945A-303	Sequence 303, App
	72	36	80.0	1051	8	US-10-807-114-303	Sequence 303, App
	73	36	80.0	1068	5	US-10-290-058A-2	Sequence 2, Appli
	74	36	80.0	1068	5	US-10-251-385-129	Sequence 129, App
	75	36	80.0	1068	5	US-10-251-385-231	Sequence 231, App

ALIGNMENTS

RESULT 1

US-09-878-722-121
; Sequence 121, Application US/09878722
; Publication No. US20020040127A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Hepler, William T.
; APPLICANT: Clapper, Jonathan
; APPLICANT: Wang, Aijun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.524
; CURRENT APPLICATION NUMBER: US/09/878,722
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 245
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-722-121

Alignment Scores:

Pred. No.:	1.05	Length:	176
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-878-722-121 (1-176)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          6 CCAGATGTGATATTTCTACAGGAAGTT 32
```

RESULT 2
US-09-878-178-50/c
; Sequence 50, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-178-50

Alignment Scores:

Pred. No.:	1.05	Length:	176
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-878-178-50 (1-176)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145
```

RESULT 3
US-09-904-456-121
; Sequence 121, Application US/09904456
; Publication No. US20030017167A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.524C1
; CURRENT APPLICATION NUMBER: US/09/904,456
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 247
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-456-121

Alignment Scores:

Pred. No.:	1.05	Length:	176
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-904-456-121 (1-176)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db          6 CCAGATGTGATATTTCTACAGGAAGTT 32
```

RESULT 4
US-10-046-935-50/c
; Sequence 50, Application US/10046935
; Publication No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-50

Alignment Scores:

Pred. No.:	1.05	Length:	176
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-046-935-50 (1-176)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145
```

RESULT 5
US-10-146-502-50/c
; Sequence 50, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-502-50

Alignment Scores:

Pred. No.:	1.05	Length:	176
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-146-502-50 (1-176)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145
```

RESULT 6
US-10-060-036-1367/c
; Sequence 1367, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1367
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-1367

Alignment Scores:

Pred. No.:	1.05	Length:	176
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-060-036-1367 (1-176)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        171 CCAGATGTGATATTTCTACAGGAAGTT 145
```

RESULT 7
US-10-066-543-1363
; Sequence 1363, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1363
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1363

Alignment Scores:

Pred. No.:	2.72	Length:	401
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-066-543-1363 (1-401)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         55 CCAGATGTGATATTTCTACAGGAAGTT 81
```

RESULT 8
US-09-864-761-16762
; Sequence 16762, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16762
;   LENGTH: 444
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
US-09-864-761-16762

Alignment Scores:
Pred. No.:                3.06          Length:          444
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-864-761-16762 (1-444)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        260 CCAGATGTGATATTTCTACAGGAAGTT 286

RESULT 9
US-10-066-543-983
; Sequence 983, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 983
;   LENGTH: 553
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 6, 35
;   OTHER INFORMATION: n = A,T,C or G

US-10-066-543-983

Alignment Scores:
Pred. No.:                3.95          Length:          553
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-066-543-983 (1-553)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         78 CCAGATGTGATATTTCTACAGGAAGTT 104

RESULT 10
US-10-066-543-1031
; Sequence 1031, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1031
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 10
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1031

Alignment Scores:
Pred. No.:                4.55          Length:          625
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-066-543-1031 (1-625)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         78 CCAGATGTGATATTTCTACAGGAAGTT 104

RESULT 11
US-10-066-543-1084
; Sequence 1084, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
```

```
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1084
;   LENGTH: 633
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1084

Alignment Scores:
Pred. No.:                4.62          Length:          633
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-066-543-1084 (1-633)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         70 CCAGATGTGATATTTCTACAGGAAGTT 96

RESULT 12
US-10-066-543-1070
; Sequence 1070, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1070
;   LENGTH: 644
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-066-543-1070

Alignment Scores:
Pred. No.:                4.71          Length:          644
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-066-543-1070 (1-644)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         68 CCAGATGTGATATTTCTACAGGAAGTT 94

RESULT 13
US-10-066-543-1067
; Sequence 1067, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1067
;   LENGTH: 674
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11, 377, 640, 654
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1067

Alignment Scores:
Pred. No.:                4.96          Length:          674
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       5             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-066-543-1067 (1-674)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         54 CCAGATGTGATATTTCTACAGGAAGTT 80

RESULT 14
US-09-864-761-33229
; Sequence 33229, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
```

```
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33229
;   LENGTH: 691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031775.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
;   OTHER INFORMATION: SWISSPROT HIT: P18004, EVALUE 7.00e-01
;   OTHER INFORMATION: NT HIT: AF201687.1, EVALUE 3.00e-45
;   OTHER INFORMATION: EST_HUMAN HIT: AV717253.1, EVALUE 5.00e-45
US-09-864-761-33229

Alignment Scores:
Pred. No.:                5.11          Length:          691
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-864-761-33229 (1-691)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        122 CCAGATGTGATATTTCTACAGGAAGTT 148

RESULT 15
US-09-745-288-63/c
; Sequence 63, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  TITLE OF INVENTION:  COMPOUNDS FOR IMMUNOTHERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE:  210121.446D1
;  CURRENT APPLICATION NUMBER: US/09/745,288
;  CURRENT FILING DATE:  2000-12-19
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-745-288-63

Alignment Scores:
Pred. No.:                8.55          Length:          1079
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       3             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-745-288-63 (1-1079)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024

RESULT 16
US-10-453-919-63/c
; Sequence 63, Application US/10453919
; Publication No. US20040033230A1
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.446C7
;  CURRENT APPLICATION NUMBER: US/10/453,919
;  CURRENT FILING DATE:  2003-06-03
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 63
;   LENGTH: 1079
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-453-919-63

Alignment Scores:
Pred. No.:                8.55          Length:          1079
Score:                    45.00         Matches:         9
Percent Similarity:       100.0%        Conservative:    0
Best Local Similarity:    100.0%        Mismatches:      0
Query Match:              100.0%        Indels:          0
DB:                       7             Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-453-919-63 (1-1079)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1050 CCAGATGTGATATTTCTACAGGAAGTT 1024

RESULT 17
US-10-450-763-22058
; Sequence 22058, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;  FILE REFERENCE: 790CIP3/US
;  CURRENT APPLICATION NUMBER: US/10/450,763
;  CURRENT FILING DATE:  2003-06-11
;  PRIOR APPLICATION NUMBER: PCT/US01/08631
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: 09/649,167
```

```
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 60736
;  SOFTWARE: Custom
; SEQ ID NO 22058
;   LENGTH: 1088
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIMILAR
;   LOCATION: (1)..(687)
;   OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein
;   OTHER INFORMATION: similar to C. elegans Y63D3A.4),accession number AL031775,Smith-
;   OTHER INFORMATION: Waterman Score=1199.
US-10-450-763-22058

Alignment Scores:
Pred. No.:                8.63          Length:         1088
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       9             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-450-763-22058 (1-1088)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db         37 CCAGATGTGATATTTCTACAGGAAGTT 63

RESULT 18
US-09-925-299-170
; Sequence 170, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:                10.6          Length:         1296
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-299-170 (1-1296)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        456 CCAGATGTGATATTTCTACAGGAAGTT 482

RESULT 19
US-09-925-299-170
; Sequence 170, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 1296
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1261)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1276)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.:                10.6          Length:         1296
Score:                    45.00         Matches:        9
Percent Similarity:       100.0%        Conservative:   0
Best Local Similarity:    100.0%        Mismatches:     0
Query Match:              100.0%        Indels:         0
DB:                       3             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-299-170 (1-1296)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        456 CCAGATGTGATATTTCTACAGGAAGTT 482

RESULT 20
US-10-757-745-1
; Sequence 1, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1920
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1627)..(1627)
;   OTHER INFORMATION: N stands for any nucleotide.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (20)..(1108)
;   OTHER INFORMATION:
;   FEATURE:
```

; NAME/KEY: misc_feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

Alignment Scores:

Pred. No.:	16.6	Length:	1920
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-757-745-1 (1-1920)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        452 CCAGATGTGATATTTCTACAGGAAGTT 478
```

RESULT 21
US-10-783-271-28
; Sequence 28, Application US/10783271
; Publication No. US20050186577A1
; GENERAL INFORMATION:
; APPLICANT: Veridex, LLC
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
; FILE REFERENCE: VDX-5003 USNP
; CURRENT APPLICATION NUMBER: US/10/783,271
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: human
US-10-783-271-28

Alignment Scores:

Pred. No.:	16.8	Length:	1940
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-783-271-28 (1-1940)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        458 CCAGATGTGATATTTCTACAGGAAGTT 484
```

RESULT 22
US-10-037-270-889
; Sequence 889, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-037-270-889

Alignment Scores:

Pred. No.:	16.9	Length:	1948
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-037-270-889 (1-1948)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        471 CCAGATGTGATATTTCTACAGGAAGTT 497
```

RESULT 23
US-10-117-722-889
; Sequence 889, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-117-722-889

Alignment Scores:

Pred. No.:	16.9	Length:	1948
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0

```
DB:                    6              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-117-722-889 (1-1948)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        471 CCAGATGTGATATTTCTACAGGAAGTT 497

RESULT 24
US-10-122-851-889
; Sequence 889, Application US/10122851
; Publication No. US20050239060A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: Novel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2BDV3
;  CURRENT APPLICATION NUMBER: US/10/122,851
;  CURRENT FILING DATE:  2002-04-12
;  PRIOR APPLICATION NUMBER: 09/620,312
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1104
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
;   LENGTH: 1948
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (39)..(1127)
US-10-122-851-889

Alignment Scores:
Pred. No.:              16.9          Length:        1948
Score:                  45.00         Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            100.0%        Indels:        0
DB:                     9             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-122-851-889 (1-1948)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db        471 CCAGATGTGATATTTCTACAGGAAGTT 497

RESULT 25
US-09-981-353-88
; Sequence 88, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE: PA-0038 US
;  CURRENT APPLICATION NUMBER: US/09/981,353
;  CURRENT FILING DATE:  2001-10-11
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 88
;   LENGTH: 3152
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020160382A1 232992.1
;   NAME/KEY: unsure
;   LOCATION: 1171
;   OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

Alignment Scores:
Pred. No.:              29.5          Length:        3152
Score:                  45.00         Matches:       9
Percent Similarity:     100.0%        Conservative:  0
Best Local Similarity:  100.0%        Mismatches:    0
Query Match:            100.0%        Indels:        0
DB:                     3             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-981-353-88 (1-3152)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||||||||||||||
Db       1684 CCAGATGTGATATTTCTACAGGAAGTT 1710

RESULT 26
US-10-757-745-3
; Sequence 3, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
;  APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
;  TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
;  FILE REFERENCE: 2676-4555US
;  CURRENT APPLICATION NUMBER: US/10/757,745
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: US/09/697,863A
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: PCT/EP99/03025
;  PRIOR FILING DATE: 1999-04-28
;  PRIOR APPLICATION NUMBER: EPO 98201392.2
;  PRIOR FILING DATE: 1998-04-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1312
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (122)..(1234)
;   OTHER INFORMATION:
US-10-757-745-3

Alignment Scores:
Pred. No.:              17.9          Length:        1312
Score:                  44.00         Matches:       8
Percent Similarity:     100.0%        Conservative:  1
Best Local Similarity:  88.9%         Mismatches:    0
Query Match:            97.8%         Indels:        0
DB:                     9             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-757-745-3 (1-1312)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||||||||||||||
Db        584 CCAGATGTGGTATTTCTACAGGAAGTT 610

RESULT 27
US-10-437-963-29447
; Sequence 29447, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
```

```
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 29447
;   LENGTH: 581
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_33949C.1
US-10-437-963-29447

Alignment Scores:
Pred. No.:                19.6          Length:         581
Score:                    42.00         Matches:        8
Percent Similarity:       100.0%        Conservative:   1
Best Local Similarity:    88.9%         Mismatches:     0
Query Match:              93.3%         Indels:         0
DB:                       7             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-437-963-29447 (1-581)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||||||||||||
Db         62 CCTGACCTGATCTTCCTCCAGGAGGTC 88

RESULT 28
US-10-066-543-548
; Sequence 548, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Pyle, Ruth A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  Durham, Margarita
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.563
;  CURRENT APPLICATION NUMBER: US/10/066,543
;  CURRENT FILING DATE:  2002-01-31
;  NUMBER OF SEQ ID NOS: 3417
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 19, 71
;   OTHER INFORMATION: n = A,T,C or G
US-10-066-543-548

Alignment Scores:
Pred. No.:                59.5          Length:         625
Score:                    40.00         Matches:        8
Percent Similarity:       88.9%         Conservative:   0
Best Local Similarity:    88.9%         Mismatches:     1
Query Match:              88.9%         Indels:         0
DB:                       5             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-066-543-548 (1-625)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||   |||||||||||||||
Db         62 CCAGATGTGNTATTTCTACAGGAAGTT 88

RESULT 29
US-10-437-963-99622
; Sequence 99622, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 99622
;   LENGTH: 1224
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1224)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_97416C.1
US-10-437-963-99622

Alignment Scores:
Pred. No.:                129           Length:         1224
Score:                    40.00         Matches:        7
Percent Similarity:       88.9%         Conservative:   1
Best Local Similarity:    77.8%         Mismatches:     1
Query Match:              88.9%         Indels:         0
DB:                       7             Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-437-963-99622 (1-1224)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||||||   ||||||:::
Db        204 CCGGACGTGATCTTCTTCCAGGAAATC 230

RESULT 30
US-10-972-079-46608/c
; Sequence 46608, Application US/10972079
; Publication No. US20050153317A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  KERR, Richard
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  HOLM, Tom
;  TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
;  TITLE OF INVENTION:  LIVESTOCK
;  FILE REFERENCE: MMI1110-2
;  CURRENT APPLICATION NUMBER: US/10/972,079
;  CURRENT FILING DATE:  2004-10-22
;  PRIOR APPLICATION NUMBER: US 60/514,333
;  PRIOR FILING DATE: 2003-10-24
;  NUMBER OF SEQ ID NOS: 96631
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 46608
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Chicken  19866894287722_1
```

US-10-972-079-46608

Alignment Scores:

Pred. No.:	95	Length:	600
Score:	39.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	86.7%	Indels:	0
DB:	9	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-972-079-46608 (1-600)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::::||||||||||||
Db        160 CCAGATATTGTTTTCCTTCAAGAA 137
```

RESULT 31
US-09-738-626-746
; Sequence 746, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 746
; LENGTH: 912
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-746

Alignment Scores:

Pred. No.:	258	Length:	912
Score:	38.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	84.4%	Indels:	0
DB:	3	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-738-626-746 (1-912)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   ||||||||||||
Db        106 CCGGACGTTGTCCTCCTCCAAGAAGTC 132
```

RESULT 32
US-09-925-065A-33485/c
; Sequence 33485, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33485
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-33485

Alignment Scores:

Pred. No.:	461	Length:	1509
Score:	38.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	3
Best Local Similarity:	66.7%	Mismatches:	0
Query Match:	84.4%	Indels:	0
DB:	4	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-33485 (1-1509)

```
Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||:::|||:::|||
Db        662 CCGGATGTGGTTTTCATTCAGAAAGTA 636
```

RESULT 33
US-10-425-114-25133
; Sequence 25133, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25133
; LENGTH: 1838
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-257-D4_FLI
US-10-425-114-25133

Alignment Scores:

Pred. No.:	579	Length:	1838
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	87.5%	Mismatches:	0
Query Match:	84.4%	Indels:	0
DB:	7	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-425-114-25133 (1-1838)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||||||||:::
Db       1034 CCAGATGTTATTTTTCTACAGCAA 1057
```

RESULT 34

```
US-10-027-632-103398
; Sequence 103398, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103398
;   LENGTH: 2209
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-103398

Alignment Scores:
Pred. No.:                716             Length:         2209
Score:                    38.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   3
Best Local Similarity:    66.7%           Mismatches:     0
Query Match:              84.4%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-027-632-103398 (1-2209)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||:::|||:::|||
Db        226 CCGGATGTGGTTTTCATTCAGAAAGTA 252

RESULT 35
US-10-027-632-103398
; Sequence 103398, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103398
;   LENGTH: 2209
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-103398

Alignment Scores:
Pred. No.:                716             Length:         2209
Score:                    38.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   3
Best Local Similarity:    66.7%           Mismatches:     0
Query Match:              84.4%           Indels:         0
DB:                       6               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-027-632-103398 (1-2209)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::|||:::|||:::|||
Db        226 CCGGATGTGGTTTTCATTCAGAAAGTA 252

RESULT 36
US-10-425-115-149831
; Sequence 149831, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 149831
;   LENGTH: 2468
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(2468)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_68171C.1
US-10-425-115-149831

Alignment Scores:
Pred. No.:                814             Length:         2468
Score:                    38.00           Matches:        7
Percent Similarity:       100.0%          Conservative:   1
Best Local Similarity:    87.5%           Mismatches:     0
Query Match:              84.4%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-425-115-149831 (1-2468)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||||||||||||:::
Db       1437 CCAGATGTTATTTTTCTACAGCAA 1460

RESULT 37
US-10-437-963-81672
; Sequence 81672, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
```

```
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 81672
;   LENGTH: 5337
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_81175C.1
US-10-437-963-81672

Alignment Scores:
Pred. No.:                1.99e+03        Length:          5337
Score:                    38.00           Matches:         8
Percent Similarity:       88.9%           Conservative:    0
Best Local Similarity:    88.9%           Mismatches:      1
Query Match:              84.4%           Indels:          0
DB:                       7               Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-437-963-81672 (1-5337)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||   ||||||||||||||||||||
Db        112 CCGCACGTCATCTTCTTGCAGGAGGTC 138

RESULT 38
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
;   LENGTH: 3309400
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Alignment Scores:
Pred. No.:                3.08e+06        Length:          3309400
Score:                    38.00           Matches:         7
Percent Similarity:       88.9%           Conservative:    1
Best Local Similarity:    77.8%           Mismatches:      1
Query Match:              84.4%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-738-626-1 (1-3309400)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
                |||||||||:::   ||||||||||||
Db     686540 CCGGACGTTGTCCTCCTCCAAGAAGTC 686566

RESULT 39
US-09-770-961-110/c
; Sequence 110, Application US/09770961
; Publication No. US20030115639A1
; GENERAL INFORMATION:
;  APPLICANT: Gorlach, Jorn
;  APPLICANT:  An, Yong-Qiang
;  APPLICANT:  Hamilton, Carol M.
;  APPLICANT:  Price, Jennifer L.
;  APPLICANT:  Raines, Tracy M.
;  APPLICANT:  Yu, Yang
;  APPLICANT:  Rameaka, Joshua G.
;  APPLICANT:  Page, Amy
;  APPLICANT:  Matthew, Abraham V.
;  APPLICANT:  Ledford, Brooke L.
;  APPLICANT:  Woessner, Jeffrey P.
;  APPLICANT:  Haas, William David
;  APPLICANT:  Garcia, Carlos A.
;  APPLICANT:  Kricker, Maja
;  APPLICANT:  Slader, Ted
;  APPLICANT:  Davis, Keith R.
;  APPLICANT:  Allen, Keith
;  APPLICANT:  Hoffman, Neil
;  APPLICANT:  Hurban, Patrick
;  TITLE OF INVENTION: Expressed Sequences of Arabidopsis
;  TITLE OF INVENTION:  thaliana
;  FILE REFERENCE: 2026 (PARA-015PRV)
;  CURRENT APPLICATION NUMBER: US/09/770,961
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER: 60/178,466
;  PRIOR FILING DATE: 2000-01-27
;  NUMBER OF SEQ ID NOS: 999
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 110
;   LENGTH: 516
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-770-961-110

Alignment Scores:
Pred. No.:                223             Length:          516
Score:                    37.00           Matches:         7
Percent Similarity:       88.9%           Conservative:    1
Best Local Similarity:    77.8%           Mismatches:      1
Query Match:              82.2%           Indels:          0
DB:                       3               Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-770-961-110 (1-516)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||   ||||||
Db        478 CCTGACATCATCTTCCTTCTTGAGGTT 452

RESULT 40
US-09-925-065A-497337/c
; Sequence 497337, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
```

```
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 497337
;   LENGTH: 637
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-497337

Alignment Scores:
Pred. No.:              285          Length:          637
Score:                  37.00        Matches:         6
Percent Similarity:     100.0%       Conservative:    2
Best Local Similarity:  75.0%        Mismatches:      0
Query Match:            82.2%        Indels:          0
DB:                     4            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-497337 (1-637)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||::::::|||||||||
Db        466 CCAGATGTGGTATATTTGCAGGAA 443

RESULT 41
US-10-424-599-123282
; Sequence 123282, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 123282
;   LENGTH: 753
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_82330C.1
US-10-424-599-123282

Alignment Scores:
Pred. No.:              346          Length:          753
Score:                  37.00        Matches:         7
Percent Similarity:     87.5%        Conservative:    0
Best Local Similarity:  87.5%        Mismatches:      1
Query Match:            82.2%        Indels:          0
DB:                     7            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-424-599-123282 (1-753)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||   ||||||||||||
Db         47 CCTGATGTTTTTTTTTTACAAGAA 70

RESULT 42
US-10-342-224-27/c
; Sequence 27, Application US/10342224
; Publication No. US20030162294A1
; GENERAL INFORMATION:
;  APPLICANT: Nathalie Verbruggen
;  TITLE OF INVENTION: Genes Involved in Tolerance to Environmental Stress
;  FILE REFERENCE: CNN-012US
;  CURRENT APPLICATION NUMBER: US/10/342,224
;  CURRENT FILING DATE:  2003-01-13
;  PRIOR APPLICATION NUMBER: US/09/762,154
;  PRIOR FILING DATE: 2002-02-02
;  PRIOR APPLICATION NUMBER: EP 98202634.6
;  PRIOR FILING DATE: 1998-08-04
;  NUMBER OF SEQ ID NOS: 123
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
;   LENGTH: 768
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)...(752)
US-10-342-224-27

Alignment Scores:
Pred. No.:              354          Length:          768
Score:                  37.00        Matches:         7
Percent Similarity:     88.9%        Conservative:    1
Best Local Similarity:  77.8%        Mismatches:      1
Query Match:            82.2%        Indels:          0
DB:                     6            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-342-224-27 (1-768)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||||||||   |||||||
Db        467 CCTGACATCATCTTCCTTCTTGAGGTT 441

RESULT 43
US-10-369-493-45451
; Sequence 45451, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 45451
;   LENGTH: 774
;   TYPE: DNA
;   ORGANISM: Rhodobacter capsulatus
US-10-369-493-45451

Alignment Scores:
Pred. No.:              357          Length:          774
Score:                  37.00        Matches:         6
Percent Similarity:     88.9%        Conservative:    2
Best Local Similarity:  66.7%        Mismatches:      1
Query Match:            82.2%        Indels:          0
DB:                     6            Gaps:            0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-369-493-45451 (1-774)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   |||||||||:::
Db         79 CCCGATGTCGTGCTGCTGCAGGAGATC 105

RESULT 44
US-10-369-493-31729
; Sequence 31729, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
```

```
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 31729
;   LENGTH: 777
;   TYPE: DNA
;   ORGANISM: Rhodobacter sphaeroides
US-10-369-493-31729

Alignment Scores:
Pred. No.:              358          Length:        777
Score:                  37.00        Matches:       6
Percent Similarity:     88.9%        Conservative:  2
Best Local Similarity:  66.7%        Mismatches:    1
Query Match:            82.2%        Indels:        0
DB:                     6            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-369-493-31729 (1-777)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   |||||||||:::
Db         85 CCTGACGTGGTGCTGCTGCAGGAGATC 111

RESULT 45
US-10-369-493-47104
; Sequence 47104, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 47104
;   LENGTH: 852
;   TYPE: DNA
;   ORGANISM: Deinococcus radiodurans
US-10-369-493-47104

Alignment Scores:
Pred. No.:              399          Length:        852
Score:                  37.00        Matches:       7
Percent Similarity:     88.9%        Conservative:  1
Best Local Similarity:  77.8%        Mismatches:    1
Query Match:            82.2%        Indels:        0
DB:                     6            Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-369-493-47104 (1-852)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||:::   ||||||||||||
Db        184 CCCGACGTGCTGCTTCTGCAAGAAGTC 210

RESULT 46
US-11-097-143-9713
; Sequence 9713, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9713
;   LENGTH: 1350
;   TYPE: DNA
;   ORGANISM: DROSOPHILA
US-11-097-143-9713

Alignment Scores:
Pred. No.:              678          Length:        1350
Score:                  37.00        Matches:       8
Percent Similarity:     88.9%        Conservative:  0
Best Local Similarity:  88.9%        Mismatches:    1
Query Match:            82.2%        Indels:        0
DB:                     10           Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-11-097-143-9713 (1-1350)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||||   |||||||||||
Db        487 CCCGATGTGATTTGTCTGCAAGAGGTA 513

RESULT 47
US-10-424-599-43531
; Sequence 43531, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 43531
;   LENGTH: 1628
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_139307C.1
US-10-424-599-43531

Alignment Scores:
Pred. No.:              842          Length:        1628
```

```
Score:                  37.00          Matches:       7
Percent Similarity:     88.9%          Conservative:  1
Best Local Similarity:  77.8%          Mismatches:    1
Query Match:            82.2%          Indels:        0
DB:                     7              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-424-599-43531 (1-1628)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||||||||   ||||||||||||:::
Db        462 CCCGACGTTCCCTTCCTCCAAGAAATC 488

RESULT 48
US-10-741-601-1929
; Sequence 1929, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1929
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-1929

Alignment Scores:
Pred. No.:              126            Length:        201
Score:                  36.00          Matches:       6
Percent Similarity:     87.5%          Conservative:  1
Best Local Similarity:  75.0%          Mismatches:    1
Query Match:            80.0%          Indels:        0
DB:                     7              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-601-1929 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24

RESULT 49
US-10-741-601-1933
; Sequence 1933, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1933
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-1933

Alignment Scores:
Pred. No.:              126            Length:        201
Score:                  36.00          Matches:       6
Percent Similarity:     87.5%          Conservative:  1
Best Local Similarity:  75.0%          Mismatches:    1
Query Match:            80.0%          Indels:        0
DB:                     7              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-601-1933 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129

RESULT 50
US-10-741-601-1938
; Sequence 1938, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1938
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-1938

Alignment Scores:
Pred. No.:              126            Length:        201
Score:                  36.00          Matches:       6
Percent Similarity:     87.5%          Conservative:  1
Best Local Similarity:  75.0%          Mismatches:    1
Query Match:            80.0%          Indels:        0
DB:                     7              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-601-1938 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24

RESULT 51
US-10-741-601-1942
; Sequence 1942, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1942
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-1942

Alignment Scores:
Pred. No.:              126            Length:        201
Score:                  36.00          Matches:       6
Percent Similarity:     87.5%          Conservative:  1
Best Local Similarity:  75.0%          Mismatches:    1
Query Match:            80.0%          Indels:        0
DB:                     7              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-601-1942 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129

RESULT 52
US-10-741-601-11767
```

```
; Sequence 11767, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11767
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-11767

Alignment Scores:
Pred. No.:                126             Length:         201
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-601-11767 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24

RESULT 53
US-10-741-601-11806
; Sequence 11806, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11806
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-11806

Alignment Scores:
Pred. No.:                126             Length:         201
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-601-11806 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129

RESULT 54
US-10-741-600-7337
; Sequence 7337, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7337
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-7337

Alignment Scores:
Pred. No.:                126             Length:         201
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-600-7337 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24

RESULT 55
US-10-741-600-7341
; Sequence 7341, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7341
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-600-7341

Alignment Scores:
Pred. No.:                126             Length:         201
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-600-7341 (1-201)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129

RESULT 56
US-10-741-600-7346
; Sequence 7346, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7346
;   LENGTH: 201
;   TYPE: DNA
```

; ORGANISM: Homo sapiens
US-10-741-600-7346

Alignment Scores:

Pred. No.:	126	Length:	201
Score:	36.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-600-7346 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24
```

RESULT 57
US-10-741-600-7350
; Sequence 7350, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7350
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-7350

Alignment Scores:

Pred. No.:	126	Length:	201
Score:	36.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-600-7350 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129
```

RESULT 58
US-10-741-600-29526
; Sequence 29526, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29526
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-29526

Alignment Scores:

Pred. No.:	126	Length:	201
Score:	36.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-600-29526 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db          1 CCAGACGCTGTTTTCCTGCAAGAA 24
```

RESULT 59
US-10-741-600-29565
; Sequence 29565, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29565
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-29565

Alignment Scores:

Pred. No.:	126	Length:	201
Score:	36.00	Matches:	6
Percent Similarity:	87.5%	Conservative:	1
Best Local Similarity:	75.0%	Mismatches:	1
Query Match:	80.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-741-600-29565 (1-201)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        106 CCAGACGCTGTTTTCCTGCAAGAA 129
```

RESULT 60
US-09-922-293-3628
; Sequence 3628, Application US/09922293
; Publication No. US20040123339A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Heck, Gregory R.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Transcription in Plants
; FILE REFERENCE: 16517.254
; CURRENT APPLICATION NUMBER: US/09/922,293
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/067,000
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: US 60/069,472
; PRIOR FILING DATE: 1997-12-09
; PRIOR APPLICATION NUMBER: US 60/071,479
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/074,201
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 60/074,282
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 60/074,280
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 60/074,281
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 60/074,566
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: US 60/074,567

; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: US 60/074,565
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: US 60/075,462
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/074,789
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/075,459
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/075,461
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/075,464
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/075,460
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/075,463
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/077,231
; PRIOR FILING DATE: 1998-03-09
; PRIOR APPLICATION NUMBER: US 60/077,229
; PRIOR FILING DATE: 1998-03-09
; PRIOR APPLICATION NUMBER: US 60/077,230
; PRIOR FILING DATE: 1998-03-09
; PRIOR APPLICATION NUMBER: US 60/078,368
; PRIOR FILING DATE: 1998-03-18
; PRIOR APPLICATION NUMBER: US 60/080,844
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 60/083,067
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: US 60/083,386
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: US 60/083,387
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: US 60/083,388
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: US 60/083,389
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: US 60/085,224
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,223
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,222
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,533
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: US 60/086,186
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,187
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,185
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,184
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,183
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,188
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/089,524
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: US 60/089,810
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,814
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,808
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,812
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,807
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,806
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,813
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,811
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/089,793
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/091,405
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/091,247
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/099,667
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 60/099,668
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 60/099,670
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 60/099,697
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 60/100,674
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 60/100,673
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 60/100,672
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 60/100,963
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 60/101,131
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 60/101,132
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 60/101,130
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 60/101,508
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/101,344
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/101,347
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/101,343
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/101,707
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: US 60/104,126
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/104,128
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/104,127
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/104,124
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/109,018
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: US 60/108,996
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: US 09/199,129
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: US 09/210,297
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 60/111,981
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 60/113,224
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/229,413
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 3853
; SEQ ID NO 3628
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Glycine max
US-09-922-293-3628

Alignment Scores:

Pred. No.:	158	Length:	245
Score:	36.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1

```
Best Local Similarity:  77.8%          Mismatches:    1
Query Match:            80.0%          Indels:        0
DB:                     3              Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-922-293-3628 (1-245)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||   ||||||||||||
Db         42 CCAGACATTATTTGTCTTCAGGAGGTT 68

RESULT 61
US-09-922-293-3627
; Sequence 3627, Application US/09922293
; Publication No. US20040123339A1
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W.
;  APPLICANT:  Heck, Gregory R.
;  APPLICANT:  Liu, Jingdong
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Transcription in Plants
;  FILE REFERENCE: 16517.254
;  CURRENT APPLICATION NUMBER: US/09/922,293
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: US 60/067,000
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: US 60/069,472
;  PRIOR FILING DATE: 1997-12-09
;  PRIOR APPLICATION NUMBER: US 60/071,479
;  PRIOR FILING DATE: 1998-01-13
;  PRIOR APPLICATION NUMBER: US 60/074,201
;  PRIOR FILING DATE: 1998-02-10
;  PRIOR APPLICATION NUMBER: US 60/074,282
;  PRIOR FILING DATE: 1998-02-10
;  PRIOR APPLICATION NUMBER: US 60/074,280
;  PRIOR FILING DATE: 1998-02-10
;  PRIOR APPLICATION NUMBER: US 60/074,281
;  PRIOR FILING DATE: 1998-02-10
;  PRIOR APPLICATION NUMBER: US 60/074,566
;  PRIOR FILING DATE: 1998-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,567
;  PRIOR FILING DATE: 1998-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,565
;  PRIOR FILING DATE: 1998-02-12
;  PRIOR APPLICATION NUMBER: US 60/075,462
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/074,789
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/075,459
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/075,461
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/075,464
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/075,460
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/075,463
;  PRIOR FILING DATE: 1998-02-19
;  PRIOR APPLICATION NUMBER: US 60/077,231
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 60/077,229
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 60/077,230
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 60/078,368
;  PRIOR FILING DATE: 1998-03-18
;  PRIOR APPLICATION NUMBER: US 60/080,844
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: US 60/083,067
;  PRIOR FILING DATE: 1998-04-27
;  PRIOR APPLICATION NUMBER: US 60/083,386
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: US 60/083,387
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: US 60/083,388
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: US 60/083,389
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: US 60/085,224
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: US 60/085,223
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: US 60/085,222
;  PRIOR FILING DATE: 1998-05-13
;  PRIOR APPLICATION NUMBER: US 60/085,533
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: US 60/086,186
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: US 60/086,187
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: US 60/086,185
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: US 60/086,184
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: US 60/086,183
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: US 60/086,188
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: US 60/089,524
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: US 60/089,810
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: US 60/089,814
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: US 60/089,808
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: US 60/089,812
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: US 60/089,807
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: US 60/089,806
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: US 60/089,813
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: US 60/089,811
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: US 60/089,793
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: US 60/091,405
;  PRIOR FILING DATE: 1998-06-30
;  PRIOR APPLICATION NUMBER: US 60/091,247
;  PRIOR FILING DATE: 1998-06-30
;  PRIOR APPLICATION NUMBER: US 60/099,667
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: US 60/099,668
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: US 60/099,670
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: US 60/099,697
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: US 60/100,674
;  PRIOR FILING DATE: 1998-09-16
;  PRIOR APPLICATION NUMBER: US 60/100,673
;  PRIOR FILING DATE: 1998-09-16
;  PRIOR APPLICATION NUMBER: US 60/100,672
;  PRIOR FILING DATE: 1998-09-16
;  PRIOR APPLICATION NUMBER: US 60/100,963
;  PRIOR FILING DATE: 1998-09-17
;  PRIOR APPLICATION NUMBER: US 60/101,131
;  PRIOR FILING DATE: 1998-09-21
;  PRIOR APPLICATION NUMBER: US 60/101,132
;  PRIOR FILING DATE: 1998-09-21
;  PRIOR APPLICATION NUMBER: US 60/101,130
;  PRIOR FILING DATE: 1998-09-21
;  PRIOR APPLICATION NUMBER: US 60/101,508
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: US 60/101,344
```

```
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: US 60/101,347
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: US 60/101,343
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: US 60/101,707
;  PRIOR FILING DATE: 1998-09-25
;  PRIOR APPLICATION NUMBER: US 60/104,126
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: US 60/104,128
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: US 60/104,127
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: US 60/104,124
;  PRIOR FILING DATE: 1998-10-13
;  PRIOR APPLICATION NUMBER: US 60/109,018
;  PRIOR FILING DATE: 1998-11-18
;  PRIOR APPLICATION NUMBER: US 60/108,996
;  PRIOR FILING DATE: 1998-11-18
;  PRIOR APPLICATION NUMBER: US 09/199,129
;  PRIOR FILING DATE: 1998-11-24
;  PRIOR APPLICATION NUMBER: US 09/210,297
;  PRIOR FILING DATE: 1998-12-08
;  PRIOR APPLICATION NUMBER: US 60/111,981
;  PRIOR FILING DATE: 1998-12-11
;  PRIOR APPLICATION NUMBER: US 60/113,224
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: US 09/229,413
;  PRIOR FILING DATE: 1999-01-12
;  NUMBER OF SEQ ID NOS: 3853
; SEQ ID NO 3627
;   LENGTH: 254
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-922-293-3627

Alignment Scores:
Pred. No.:                165             Length:         254
Score:                    36.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-922-293-3627 (1-254)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||   ||||||||||||
Db         33 CCAGACATTATTTGTCTTCAGGAGGTT 59

RESULT 62
US-10-027-632-60370
; Sequence 60370, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60370
;   LENGTH: 492
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-60370

Alignment Scores:
Pred. No.:                353             Length:         492
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    75.0%           Mismatches:     0
Query Match:              80.0%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-027-632-60370 (1-492)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||:::|||:::
Db        247 CCTGATGTAATCTTCATCCAGCAG 270

RESULT 63
US-10-027-632-107582
; Sequence 107582, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107582
;   LENGTH: 492
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-107582

Alignment Scores:
Pred. No.:                353             Length:         492
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    75.0%           Mismatches:     0
Query Match:              80.0%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-027-632-107582 (1-492)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||:::|||:::
Db        247 CCTGATGTAATCTTCATCCAGCAG 270

RESULT 64
US-10-027-632-60370
```

; Sequence 60370, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60370
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-60370

Alignment Scores:

Pred. No.:	353	Length:	492
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	80.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-027-632-60370 (1-492)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||:::|||:::
Db        247 CCTGATGTAATCTTCATCCAGCAG 270
```

RESULT 65
US-10-027-632-107582
; Sequence 107582, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107582
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-107582

Alignment Scores:

Pred. No.:	353	Length:	492
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	80.0%	Indels:	0
DB:	6	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-027-632-107582 (1-492)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||||||||||:::|||:::
Db        247 CCTGATGTAATCTTCATCCAGCAG 270
```

RESULT 66
US-10-425-115-27387/c
; Sequence 27387, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 27387
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(559)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_124993C.1
US-10-425-115-27387

Alignment Scores:

Pred. No.:	410	Length:	559
Score:	36.00	Matches:	6
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	75.0%	Mismatches:	0
Query Match:	80.0%	Indels:	0
DB:	8	Gaps:	0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-425-115-27387 (1-559)

```
Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||:::||||||:::||||||
Db        528 CCTGATTTGATCTTCATACAAGAA 505
```

RESULT 67
US-10-437-963-26483/c
; Sequence 26483, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

```
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 26483
;   LENGTH: 583
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_31269C.1
US-10-437-963-26483

Alignment Scores:
Pred. No.:                430             Length:         583
Score:                    36.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-437-963-26483 (1-583)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              |||   |||||||||||||||||:::
Db        209 CCTCGCGTGATTTTTCTACAGGAGATC 183

RESULT 68
US-09-925-065A-704289
; Sequence 704289, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 704289
;   LENGTH: 606
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-704289

Alignment Scores:
Pred. No.:                450             Length:         606
Score:                    36.00           Matches:        6
Percent Similarity:       100.0%          Conservative:   2
Best Local Similarity:    75.0%           Mismatches:     0
Query Match:              80.0%           Indels:         0
DB:                       4               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-925-065A-704289 (1-606)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              |||||||||||||||:::|||:::
Db         27 CCTGATGTAATCTTCATCCAGCAG 50

RESULT 69
US-10-424-599-156
; Sequence 156, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 156
;   LENGTH: 714
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_100140C.1
US-10-424-599-156

Alignment Scores:
Pred. No.:                544             Length:         714
Score:                    36.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-424-599-156 (1-714)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||:::|||   ||||||||||||
Db        426 CCTGATATTATCTGCTTACAAGAGGTG 452

RESULT 70
US-10-282-122A-35952
; Sequence 35952, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
```

```
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35952
;    LENGTH: 717
;    TYPE: DNA
;    ORGANISM: Streptococcus mutans
US-10-282-122A-35952

Alignment Scores:
Pred. No.:                546             Length:         717
Score:                    36.00           Matches:        7
Percent Similarity:       88.9%           Conservative:   1
Best Local Similarity:    77.8%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       7               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-282-122A-35952 (1-717)

Qy          1 ProAspValIlePheLeuGlnGluVal 9
              ||||||   ||||||||||:::|||||||
Db        385 CCAGATAAAATATTTCTAGAAGAAGTC 411

RESULT 71
US-09-882-945A-303
; Sequence 303, Application US/09882945A
; Publication No. US20030143535A1
; GENERAL INFORMATION:
;  APPLICANT: Lyamichev, Victor
;  APPLICANT:  Allawi, Hatim
;  APPLICANT:  Dong, Fang
;  APPLICANT:  Neri, Bruce
;  APPLICANT:  Vener, Tatiana
;  TITLE OF INVENTION: Nucleic Acid Accessible Hybridization Sites
;  FILE REFERENCE: FORS-04586
;  CURRENT APPLICATION NUMBER: US/09/882,945A
;  CURRENT FILING DATE:  2001-06-15
;  NUMBER OF SEQ ID NOS: 334
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 303
;    LENGTH: 1051
;    TYPE: RNA
;    ORGANISM: Rattus rattus
US-09-882-945A-303

Alignment Scores:
Pred. No.:                850             Length:         1051
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       3               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-09-882-945A-303 (1-1051)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||
Db        467 CCGGACGCUGUUUUCCUGCAAGAA 490

RESULT 72
US-10-807-114-303
; Sequence 303, Application US/10807114
; Publication No. US20040235024A1
; GENERAL INFORMATION:
;  APPLICANT: Lyamichev, Victor
;  APPLICANT:  Allawi, Hatim
;  APPLICANT:  Dong, Fang
;  APPLICANT:  Neri, Bruce
;  APPLICANT:  Vener, Tatiana
;  TITLE OF INVENTION: Nucleic Acid Accessible Hybridization Sites
;  FILE REFERENCE: FORS-04586
;  CURRENT APPLICATION NUMBER: US/10/807,114
;  CURRENT FILING DATE:  2004-03-23
;  PRIOR APPLICATION NUMBER: US/09/882,945
;  PRIOR FILING DATE: 2001-06-15
;  NUMBER OF SEQ ID NOS: 334
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 303
;    LENGTH: 1051
;    TYPE: RNA
;    ORGANISM: Rattus rattus
US-10-807-114-303

Alignment Scores:
Pred. No.:                850             Length:         1051
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       8               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-807-114-303 (1-1051)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||
Db        467 CCGGACGCUGUUUUCCUGCAAGAA 490

RESULT 73
US-10-290-058A-2
; Sequence 2, Application US/10290058A
; Publication No. US20030104455A1
; GENERAL INFORMATION:
;  APPLICANT: Silos-Santiago, Inmaculada
;  TITLE OF INVENTION: Methods and Compositions for Treating
;  TITLE OF INVENTION:  Urological Disorders Using 313, 333, 5464, 18817 or 33524
;  FILE REFERENCE: MPI01-289P1RM
;  CURRENT APPLICATION NUMBER: US/10/290,058A
;  CURRENT FILING DATE:  2002-11-07
;  PRIOR APPLICATION NUMBER: 60/344,552
;  PRIOR FILING DATE: 2001-11-07
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;    LENGTH: 1068
;    TYPE: DNA
;    ORGANISM: Homo Sapien
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (1)...(1068)
US-10-290-058A-2

Alignment Scores:
Pred. No.:                866             Length:         1068
Score:                    36.00           Matches:        6
Percent Similarity:       87.5%           Conservative:   1
Best Local Similarity:    75.0%           Mismatches:     1
Query Match:              80.0%           Indels:         0
DB:                       5               Gaps:           0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-290-058A-2 (1-1068)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||
Db        645 CCAGACGCTGTTTTCCTGCAAGAA 668

RESULT 74
US-10-251-385-129
; Sequence 129, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
;  APPLICANT: Behan, Dominic P.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Liaw, Chen W.
```

```
;  TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
;  TITLE OF INVENTION:  Protein-Coupled
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0040
;  CURRENT APPLICATION NUMBER: US/10/251,385
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US/09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 294
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 129
;   LENGTH: 1068
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-251-385-129

Alignment Scores:
Pred. No.:              866           Length:        1068
Score:                  36.00         Matches:       6
Percent Similarity:     87.5%         Conservative:  1
Best Local Similarity:  75.0%         Mismatches:    1
Query Match:            80.0%         Indels:        0
DB:                     5             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-251-385-129 (1-1068)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        645 CCAGACGCTGTTTTCCTGCAAGAA 668

RESULT 75
US-10-251-385-231
; Sequence 231, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
;  APPLICANT: Behan, Dominic P.
;  APPLICANT:  Chalmers, Derek T.
;  APPLICANT:  Liaw, Chen W.
;  TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
;  TITLE OF INVENTION:  Protein-Coupled
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN-0040
;  CURRENT APPLICATION NUMBER: US/10/251,385
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US/09/170,496
;  PRIOR FILING DATE: 1998-10-13
;  NUMBER OF SEQ ID NOS: 294
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 231
;   LENGTH: 1068
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-251-385-231

Alignment Scores:
Pred. No.:              866           Length:        1068
Score:                  36.00         Matches:       6
Percent Similarity:     87.5%         Conservative:  1
Best Local Similarity:  75.0%         Mismatches:    1
Query Match:            80.0%         Indels:        0
DB:                     5             Gaps:          0

US-10-757-745-2_COPY_145_153 (1-9) x US-10-251-385-231 (1-1068)

Qy          1 ProAspValIlePheLeuGlnGlu 8
              ||||||   :::|||||||||||||
Db        645 CCAGACGCTGTTTTCCTGCAAGAA 668


Search completed: March 11, 2006, 05:27:18
Job time : 73.8465 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2006, 12:10:44 ; Search time 5827 Seconds
(without alignments)
10534.521 Million cell updates/sec

Title: US-10-757-745-3
Perfect score: 1312
Sequence: 1 agctattaatgattcgaatt.....tctcaacctttcaggacatc 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 75 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1156.6	88.2	1289	4	AK009089	AK009089 Mus muscu
	2	712.6	54.3	942	2	BG298330	BG298330 602397112
	3	712.4	54.3	815	5	BU936308	BU936308 AGENCOURT
	4	710.8	54.2	767	8	CX226137	CX226137 MBM01355
	5	669	51.0	841	2	BF784792	BF784792 602111139
	6	666.6	50.8	1168	4	CR601303	CR601303 full-leng
	7	666.6	50.8	1894	4	CR597293	CR597293 full-leng
	8	666.6	50.8	1909	4	CR595644	CR595644 full-leng
c	9	658.4	50.2	714	2	BG070243	BG070243 H3085C02-
	10	657.4	50.1	1089	11	DQ049205	DQ049205 Homo sapi
	11	655	49.9	1089	11	DQ049206	DQ049206 Pan trogl
	12	649.4	49.5	963	3	BI656151	BI656151 603283518
	13	647	49.3	647	1	AA536734	AA536734 vj88b09.r
	14	642.4	49.0	644	2	BG083167	BG083167 H3085C02-
	15	625	47.6	1620	4	CR592636	CR592636 full-leng
	16	617	47.0	617	7	CN694288	CN694288 E0348E09-
	17	616.6	47.0	1743	4	CR602029	CR602029 full-leng
	18	612	46.6	703	5	BY719852	BY719852 BY719852
	19	608.6	46.4	712	2	BB609556	BB609556 BB609556
	20	606.8	46.2	648	1	AW321603	AW321603 uo36b06.y
	21	600	45.7	600	3	BI990829	BI990829 4081-88 M
	22	593.6	45.2	965	7	CN803230	CN803230 ILLUMIGEN
	23	591.8	45.1	614	2	BG695251	BG695251 NISC_iv14
	24	587.4	44.8	621	2	BF182465	BF182465 601804232
	25	586.8	44.7	734	7	CV119018	CV119018 AGENCOURT
	26	580	44.2	580	3	BM503555	BM503555 ih33e07.y
	27	579	44.1	1081	5	BX444691	BX444691 BX444691
	28	576	43.9	904	5	BX422491	BX422491 BX422491
c	29	573	43.7	952	5	BX433489	BX433489 BX433489
	30	572.6	43.6	936	5	BY708937	BY708937 BY708937
	31	572.4	43.6	1018	5	BX337141	BX337141 BX337141
	32	563.8	43.0	1087	7	CN642527	CN642527 ILLUMIGEN
	33	560.8	42.7	911	2	BF101875	BF101875 601753033
	34	560.8	42.7	972	3	BM468826	BM468826 AGENCOURT
	35	559	42.6	1046	3	BM926092	BM926092 AGENCOURT
	36	553.2	42.2	1067	5	BX337905	BX337905 BX337905
	37	552.6	42.1	981	5	BX338160	BX338160 BX338160
	38	551.4	42.0	942	5	BX358707	BX358707 BX358707
	39	551	42.0	645	5	BY737520	BY737520 BY737520
	40	545	41.5	637	2	BF471995	BF471995 UI-M-BH3-
	41	543.8	41.4	600	2	BG800090	BG800090 2112-54 M
	42	542.8	41.4	883	5	BU179107	BU179107 AGENCOURT
	43	542.8	41.4	1090	7	CN641671	CN641671 ILLUMIGEN
	44	538.6	41.1	922	7	CN801806	CN801806 ILLUMIGEN
	45	524.2	40.0	639	6	CB545297	CB545297 AMGNNUC:N
	46	520	39.6	531	1	AA824213	AA824213 vy20h08.r
	47	517	39.4	517	9	CC248786	CC248786 XK529 Bay
	48	516.2	39.3	961	2	BI150788	BI150788 602914879
	49	512.2	39.0	682	10	CZ905189	CZ905189 NAISTrap_
	50	511.8	39.0	544	1	AA208842	AA208842 mw73d02.r
	51	505.8	38.6	846	7	CR765451	CR765451 DKFZp469E
	52	505	38.5	948	1	AL555333	AL555333 AL555333
	53	503	38.3	749	7	CR767800	CR767800 DKFZp469C
	54	495.2	37.7	810	8	CX166335	CX166335 HESC2_39_
	55	492.6	37.5	870	2	BG740396	BG740396 602634171
	56	489.8	37.3	1192	3	BM553049	BM553049 AGENCOURT
	57	485.6	37.0	870	5	BU159911	BU159911 AGENCOURT
	58	479.4	36.5	493	1	AA163045	AA163045 mr28g09.r
c	59	474.4	36.2	587	1	AU015298	AU015298 AU015298
	60	474	36.1	615	6	CB579756	CB579756 AMGNNUC:N
	61	472	36.0	932	2	BG740339	BG740339 602635289
	62	469.4	35.8	471	1	AA681654	AA681654 vr43d12.s
	63	468	35.7	468	1	AA636415	AA636415 vq86d01.s
	64	466	35.5	680	7	CK952220	CK952220 4091710 B
	65	465.2	35.5	1035	3	BM554324	BM554324 AGENCOURT
	66	459.6	35.0	1103	3	BM555041	BM555041 AGENCOURT
	67	459.4	35.0	883	5	BU169945	BU169945 AGENCOURT
	68	458	34.9	547	1	AA476073	AA476073 vh26g06.r
	69	456	34.8	627	6	CB577890	CB577890 AMGNNUC:N
	70	455.8	34.7	546	6	CB600618	CB600618 AGENCOURT
	71	454	34.6	729	7	CN298924	CN298924 170004244
	72	452.8	34.5	767	2	BI258848	BI258848 602969663
	73	452.2	34.5	525	2	BE290901	BE290901 601084205
	74	451.2	34.4	1046	6	CA976386	CA976386 AGENCOURT
	75	449.8	34.3	691	3	BI648868	BI648868 603275826

ALIGNMENTS

RESULT 1
AK009089
LOCUS AK009089 1289 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310002D07 product:Traf and Tnf receptor associated protein, full insert sequence.
ACCESSION AK009089
VERSION AK009089.1 GI:12843662
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1

```
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1289)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.jp/) for further
            details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved
            with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
            Host: SOLR.
FEATURES              Location/Qualifiers
     source           1. .1289
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="C57BL/6J"
                      /db_xref="FANTOM_DB:2310002D07"
                      /db_xref="taxon:10090"
                      /clone="2310002D07"
                      /sex="male"
                      /tissue_type="tongue"
                      /clone_lib="RIKEN full-length enriched mouse cDNA library"
                      /dev_stage="adult"
     misc_feature     55. .1160
                      /note="Traf and Tnf receptor associated protein
                      (MGD|MGI:1860486, GB|NM_019551, evidence: BLASTN, 99%,
                      match=1184)
                      putative"
     polyA_signal     1260. .1265
                      /note="putative"
     polyA_site       1289
                      /note="putative"
ORIGIN

  Query Match             88.2%;  Score 1156.6;  DB 4;  Length 1289;
  Best Local Similarity   99.0%;  Pred. No. 4.1e-304;
  Matches 1227;  Conservative    0;  Mismatches    5;  Indels    8;  Gaps    6;

Qy         73 GGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGG 132
               |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db          7 GCGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGA-CCTGCGGCGATGGCGTCTGG 65

Qy        133 CAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGA 192
              |||||||||||||||||| | |    ||| ||||||||||||||||||||||||||||||
Db         66 CAGCAGTTCCGATGCGGCGGACCGC---AGGCNGGCAGGGCGGGCGGCGTCGGCGCCCGA 122

Qy        193 AGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGC 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGC 182

Qy        253 GTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCA 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCA 242

Qy        313 GACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCG 372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCG 302

Qy        373 CCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAA 432
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 -CAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAA 361

Qy        433 TGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTAT 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTAT 421

Qy        493 TTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCG 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCG 481

Qy        553 AGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTAT 612
              |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 AGGGGTGTG-TCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTAT 540

Qy        613 CCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAA 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 -CCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAA 599

Qy        673 TGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 732
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        600 TGAAGAAGGATATTTCACAGCTATACTATTGAAGAAA-GAAGAGTGAAATTTAAAAGTCA 658

Qy        733 GGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAG 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 GGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAG 718

Qy        793 TTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTC 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        719 TTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTC 778

Qy        853 TGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTC 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 TGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTC 838

Qy        913 AACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGG 972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 AACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGG 898

Qy        973 TGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 TGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCA 958

Qy       1033 GTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCG 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCG 1018

Qy       1093 TTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCT 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 TTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCT 1078

Qy       1153 TGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTG 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 TGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTG 1138

Qy       1213 CACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCAC 1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 CACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCAC 1198

Qy       1273 TAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              ||||||||||||||||||||||||||||||||||||||||
Db       1199 TAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1238
```

RESULT 2
BG298330
LOCUS BG298330 942 bp mRNA linear EST 21-FEB-2001
DEFINITION 602397112F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4511666 5', mRNA sequence.
ACCESSION BG298330
VERSION BG298330.1 GI:13062875
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 942)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10395 row: h column: 03
High quality sequence stop: 715.
FEATURES Location/Qualifiers
source 1. .942
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4511666"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
ORIGIN

Query Match 54.3%; Score 712.6; DB 2; Length 942;
Best Local Similarity 95.2%; Pred. No. 6.4e-183;
Matches 778; Conservative 0; Mismatches 34; Indels 5; Gaps 4;

```
Qy         86 CGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGAT 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGAT 60

Qy        146 GCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCG 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCG 120

Qy        206 GAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGA 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGA 180

Qy        266 TGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCC 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCC 240

Qy        326 CTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACG 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACG 300

Qy        386 TCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATT 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATT 360

Qy        446 TTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACC 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACC 420

Qy        506 TGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCC 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCC 480

Qy        566 TGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGT 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGT 540

Qy        626 GCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATAT 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATAT 600

Qy        686 TTCACAGCTATACTATT-GAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCC 744
              |||||||||||||||| |||||||||||||||||| ||| ||||||||||||||||||||
Db        601 TTCACAGCTATACTATTGGAAGAAAGGAAGAGTGAACTTTCAAAGTCAGGAGATTATTCC 660

Qy        745 TTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAAT-GTGAGTTTGGGTGGAA 803
              |||||||||||||||||||||||||||||| |||||||||||||| ||||||| |||||||
Db        661 TTTTCCAAATACCAAAATGATGAGAAAACTGCTATGCGTAAATGGTGAGTTGGGGTGGAC 720

Qy        804 ATGAATTTTGCCTTATGACA-TCCCATTTGGAGAGCACCAGAGAACATT--CTGCGGAAC 860
              ||| |||||| ||||||| | ||| ||| || ||| |||||  | |||    | |||||
Db        721 ATGGATTTTGGCTTATGAAATTCCGATTGGGGGAGGACCAGAAGAAATTTCTGGGGGAAC 780

Qy        861 GAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCA 897
              | || |   | ||||||  ||    |||  | ||||
Db        781 GGATTACGACACTAAAAATGGTCTGGGGACCATTGCA 817
```

RESULT 3
BU936308
LOCUS BU936308 815 bp mRNA linear EST 18-OCT-2002
DEFINITION AGENCOURT_10522379 NIH_MGC_169 Mus musculus cDNA clone IMAGE:6704064 5', mRNA sequence.
ACCESSION BU936308

```
VERSION     BU936308.1  GI:24125127
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 815)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jonathan Kuo, NIMH
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2999  row: e  column: 24
            High quality sequence stop: 600.
FEATURES             Location/Qualifiers
     source          1. .815
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6704064"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_169"
                     /note="Organ: Testicles; Vector: pDNR-LIB; Site_1: SfiI
                     (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); cDNA made
                     by oligo-dT priming and directionally cloned. 5' and 3'
                     adaptors were used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 0.5 kb
                     size fraction. Library created in thelaboratory of M.
                     Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
ORIGIN
```

```
  Query Match             54.3%;  Score 712.4;  DB 5;  Length 815;
  Best Local Similarity   99.9%;  Pred. No. 7e-183;
  Matches  713;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        590 GTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCC 649
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCC 60

Qy        650 AGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAA 709
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAA 120

Qy        710 GGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGA 769
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGA 180

Qy        770 AACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCAT 829
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCAT 240

Qy        830 TTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGA 889
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGA 300

Qy        890 AAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGA 949
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGA 360

Qy        950 GATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTT 1009
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTT 420

Qy       1010 TTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGG 1069
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGG 480

Qy       1070 ATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCAC 1129
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCAC 540

Qy       1130 CTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCG 1189
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCG 600

Qy       1190 AGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTG 1249
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTG 660

Qy       1250 CAGCTTTACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTT 1303
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGCTTTACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTT 714
```

```
RESULT 4
CX226137
LOCUS       CX226137                767 bp    mRNA    linear   EST 29-DEC-2004
DEFINITION  MBM01355 Mus Musculus hematopoietic BM-HPC5 cDNA library Mus
            musculus cDNA 5', mRNA sequence.
ACCESSION   CX226137
VERSION     CX226137.1  GI:56881429
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 767)
  AUTHORS   Williams,C., Wirta,V., Richter,K., Karlsson,C., Lundeberg,J. and
            Carlsson,L.
  TITLE     Expressed sequence tags of cDNA clones from a hematopoietic stem
            cell line expressing Lhx2
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Williams, C.
            Molecular Biotechnology
            Institution of Biotechnology
            AlbaNova University Center, KTH-Royal Institute of Technology, 106
            91 Stockholm, Sweden
            Tel: +46855378332.
            Fax: +46855378481
            Email: cecilia.williams@biotech.kth.se
            Seq primer: M13REV.
FEATURES             Location/Qualifiers
     source          1. .767
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6-cast"
                     /db_xref="taxon:10090"
                     /sex="Male and Female"
                     /tissue_type="Bone Marrow"
                     /cell_type="Hematopoietic progenitor/stem cells
                     immortalized by Lhx2"
                     /cell_line="hematopoietic stem cell-like cell line
                     BM-HPC5"
                     /dev_stage="adult"
                     /clone_lib="Mus Musculus hematopoietic BM-HPC5 cDNA
                     library"
                     /note="Organ: Bone Marrow; Vector: pCMVSport6.1;
                     Preamplified custom cDNA library by Invitrogen/ResGen"
ORIGIN
```

```
  Query Match             54.2%;  Score 710.8;  DB 8;  Length 767;
  Best Local Similarity   99.3%;  Pred. No. 1.9e-182;
  Matches  745;  Conservative    0;  Mismatches    2;  Indels    3;  Gaps    3;
```

```
Qy        131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 60

Qy        191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 120

Qy        251 GCGTTGGTGGGGGGATGC-GACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTG 309
              ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        121 GCGTTGGTGGGGGGATGCTGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTG 180

Qy        310 GCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCC 369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCC 240

Qy        370 GCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGC 429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGC 300

Qy        430 AAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCAC 489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCAC 360

Qy        490 TATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGC 549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGC 420

Qy        550 TCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGT 609
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        421 TCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT-CAGGAAGT 479

Qy        610 TATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGG 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGG 539

Qy        670 TAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAG 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 TAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAG 599

Qy        730 TCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGT 659

Qy        790 GAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACA 849
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        660 GAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAAAACA 719

Qy        850 TTC-TGCGGAACGAATAAGACAATTAAAAA 878
              ||| ||||| ||||||||||||||||||||
Db        720 TTCTTGCGGGACGAATAAGACAATTAAAAA 749
```

```
RESULT 5
BF784792
LOCUS       BF784792                841 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  602111139F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4239537
            5', mRNA sequence.
ACCESSION   BF784792
VERSION     BF784792.1  GI:12089828
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 841)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9852  row: m  column: 10
            High quality sequence stop: 719.
FEATURES             Location/Qualifiers
     source          1. .841
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4239537"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Kid14"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library. |"
ORIGIN
```

Query Match 51.0%; Score 669; DB 2; Length 841;
Best Local Similarity 97.7%; Pred. No. 5e-171;
Matches 753; Conservative 0; Mismatches 10; Indels 8; Gaps 7;

```
Qy         82 GCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTC 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTC 60

Qy        142 CGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACA 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACA 120

Qy        202 GGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGG 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGG 180

Qy        262 GGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAA 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAA 240

Qy        322 AGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCC 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCC 300

Qy        382 CACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAAC 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAAC 360

Qy        442 CATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCAT 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCAT 420

Qy        502 TACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTG 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTG 480

Qy        562 TTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATA 540

Qy        622 CTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAA-TGAAGAAG 680
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        541 CTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAACTGAAGAAG 600

Qy        681 GATATTTCACAGCT-ATACTATTGAA-GAAAGGAAGAGTGAAAT--TTAAAAGTCAGGAG 736
              ||||||||||||| |||||||||| |||||||||||||||||  |  | |||||||||
Db        601 GATATTTCACAGCTAATACTATTGAACGAAAGGAAGAGTGAAATTCTACACAGTCAGGAG 660

Qy        737 AT-TATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTT 795
              || |||||||||||||||||||||||||||||||||| |||||||||||||||||| ||
```

```
Db        661 ATATATTCCTTTTCCAAATACCAAAATGATGAGAAATCTGCTATGCGTAAATGTGAG-TT 719

Qy        796 GGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGA 846

Db        720 GGGTGGACATGAA-TTTGCCTTATGACATCCATTTTGGAGAGCCCGAGAAA 769


RESULT 6
CR601303
LOCUS       CR601303               1168 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DN005YN02 of Adult brain of Homo sapiens
            (human).
ACCESSION   CR601303
VERSION     CR601303.1  GI:50482110
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1168)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1168)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1168
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="Adult brain"
                     /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match             50.8%;  Score 666.6;  DB 4;  Length 1168;
  Best Local Similarity   76.9%;  Pred. No. 2.5e-170;
  Matches  842;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230

Db         42 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 101

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290

Db        102 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 161

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db        162 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 221

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410

Db        222 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467

Db        282 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 341

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527

Db        342 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 401

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587

Db        402 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 461

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647

Db        462 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 521

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707

Db        522 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 581

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767

Db        582 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 641

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827

Db        642 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 701

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887

Db        702 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 761

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        762 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 821

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007

Db        822 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 881

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067

Db        882 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 941

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121

Db        942 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1001

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181

Db       1002 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1061

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241

Db       1062 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1121

Qy       1242 CCCACTTGCAGCTTT 1256

Db       1122 TTCAAGTGTGGGTTT 1136


RESULT 7
CR597293
LOCUS       CR597293               1894 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DK007YK05 of HeLa cells Cot 25-normalized
            of Homo sapiens (human).
ACCESSION   CR597293
VERSION     CR597293.1  GI:50478100
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1894)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
```

```
  REMARK     Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
             Faraday Avenue
REFERENCE    2  (bases 1 to 1894)
  AUTHORS    Genoscope.
  TITLE      Direct Submission
  JOURNAL    Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT      1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
             end enriched, double-strand cDNA was digested with Not I and cloned
             into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
             was normalized. Library was constructed by Life Technologies, a
             division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DK007YK05"
                     /tissue_type="HeLa cells Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match           50.8%;  Score 666.6;  DB 4;  Length 1894;
  Best Local Similarity   76.9%;  Pred. No. 2.9e-170;
  Matches  842;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     ||||||  | |||||
Db         35 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 94

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db         95 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 154

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||   || |
Db        155 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 214

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||   |  | |   |||| |||   ||   ||  || || || |||||||||
Db        215 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 274

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        275 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 334

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        335 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 394

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||| | ||||||||||| || ||||
Db        395 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 454

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||||| |||||||| | |  ||||||||||||||||| ||
Db        455 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 514

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||| | |||||||
Db        515 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 574

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        575 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 634

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  ||||||   ||   ||||||||  ||||||||||||||||||
Db        635 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 694

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||  |  || |||||||  ||| |
Db        695 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 754

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        755 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 814

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  |||||  || | |||| ||||||  |
Db        815 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 874

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||||||||  ||   ||    || ||
Db        875 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 934

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||| ||| | |||||||||||||||||||||||      ||||||||
Db        935 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 994

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db        995 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1054

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1055 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1114

Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1115 TTCAAGTGTGGGTTT 1129


RESULT 8
CR595644
LOCUS        CR595644               1909 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION   full-length cDNA clone CS0DI052YN13 of Placenta Cot 25-normalized
             of Homo sapiens (human).
ACCESSION    CR595644
VERSION      CR595644.1  GI:50476451
KEYWORDS     HTC; CNSLT_cDNA.
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Hominidae; Homo.
REFERENCE    1  (bases 1 to 1909)
  AUTHORS    Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE      Full-length cDNA libraries and normalization
  JOURNAL    Unpublished
  REMARK     Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
             Faraday Avenue
REFERENCE    2  (bases 1 to 1909)
  AUTHORS    Genoscope.
  TITLE      Direct Submission
  JOURNAL    Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT      1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
             end enriched, double-strand cDNA was digested with Not I and cloned
             into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
             was normalized. Library was constructed by Life Technologies, a
             division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI052YN13"
                     /tissue_type="Placenta Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN
```

```
Query Match           50.8%;  Score 666.6;  DB 4;  Length 1909;
Best Local Similarity 76.9%;  Pred. No. 2.9e-170;
Matches  842;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  |||     ||||| | |||||
Db         41 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 100

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        101 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 160

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   ||||||||||||| ||| | |  || |||| | ||||||||||||  || |
Db        161 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 220

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |   ||||  |||   ||   ||   || || || |||||||||
Db        221 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 280

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||| |   |
Db        281 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 340

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||    | ||| |||||||||||||||||||||||||||||
Db        341 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 400

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | ||||||||||||||||||||||||| | |||||||||| || ||||
Db        401 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 460

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        461 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 520

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||||| | |||||||
Db        521 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 580

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        581 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 640

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  ||||||   ||   ||||||||  ||||||||||||||||||
Db        641 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 700

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||  ||| ||||||||||| |  || |||||||  ||| |
Db        701 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 760

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        761 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 820

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  ||||| || | |||| ||||||  |
Db        821 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 880

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| ||||||||||||||||||||||||||||||||||  ||   ||   || ||
Db        881 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 940

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||| | |||||||||||||||||||||||      ||||||||
Db        941 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 1000

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| ||||| || ||||||||||||||||||
Db       1001 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1060

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              ||||||| |||||||||||||| ||  |||||| |||  || || |||||| ||| ||||
Db       1061 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTT 1120

Qy       1242 CCCACTTGCAGCTTT 1256
                 ||  ||  | |||
Db       1121 TTCAAGTGTGGGTTT 1135


RESULT 9
BG070243/c
LOCUS       BG070243                714 bp    mRNA    linear   EST 18-DEC-2003
DEFINITION  H3085C02-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
            H3085C02 3', mRNA sequence.
ACCESSION   BG070243
VERSION     BG070243.2  GI:40069891
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 714)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
            Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,
            Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
  PUBMED    10922068
COMMENT     On Jan 26, 2001 this sequence version replaced gi:12552812.
            Other_ESTs: H3085C02-5
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3085  row: C  column: 02
            Seq primer: -21M13 Forward
            High quality sequence stop: 714
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .714
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:H3085C02-3"
                     /db_xref="taxon:10090"
                     /clone="H3085C02"
                     /sex="Clones arrayed from a variety of cDNA libraries"
                     /dev_stage="Clones arrayed from a variety of cDNA
                     libraries"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse 15K cDNA Clone Set"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 15,247 clones from 11
                     embryo  cDNA libraries (including preimplantation stage
                     embryos from  unfertilized egg to blastocyst, embryonic
                     part of E7.5 embryos,  extraembryonic part of E7.5
                     embryos, and E12.5 female mesonephros/gonad)  and one
                     newborn ovary cDNA library. Average insert size 1.5 kb.
                     All  source libraries are cloned unidirectionally with
                     Oligo(dT)-Not primers.  References include: (1)
                     Genome-wide expression profiling of mid-gestation
                     placenta and embryo using a 15,000 mouse developmental
                     cDNA microarray,  2000, Proc. Natl. Acad. Sci. U S A, 97:
                     9127-9132; (2) Large-scale cDNA  analysis reveals phased
                     gene expression patterns during preimplantation  mouse
                     develolpment, 2000, Development, 127: 1737-1749; (3)
                     Genome-wide  mapping of unselected transcripts from
                     extraembryonic tissue of 7.5-day  mouse embryos reveals
                     enrichment in the t-complex and under-representation  on
```

the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Query Match 50.2%; Score 658.4; DB 2; Length 714;
Best Local Similarity 99.8%; Pred. No. 3.7e-168;
Matches 659; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        653 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 652

Qy        713 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAAC 772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        651 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAGC 592

Qy        773 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 532

Qy        833 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 472

Qy        893 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 412

Qy        953 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 352

Qy       1013 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 292

Qy       1073 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTT 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         291 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTT 232

Qy       1133 ATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGT 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         231 ATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGT 172

Qy       1193 GATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAG 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         171 GATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAG 112

Qy       1253 CTTTACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         111 CTTTACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 52
```

RESULT 10
DQ049205
LOCUS DQ049205 1089 bp DNA linear GSS 02-JUN-2005
DEFINITION Homo sapiens TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION DQ049205
VERSION DQ049205.1 GI:66902404
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1089)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325

REFERENCE 2 (bases 1 to 1089)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.
FEATURES Location/Qualifiers
source 1. .1089
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
gene <1. .>1089
/gene="TTRAP"
/locus_tag="HC17203"

ORIGIN

Query Match 50.1%; Score 657.4; DB 11; Length 1089;
Best Local Similarity 77.2%; Pred. No. 8e-168;
Matches 828; Conservative 0; Mismatches 236; Indels 9; Gaps 2;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    | |  | || |  | | ||  |||     ||||| | |||||
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 76

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db         77 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | |||||||||||  || |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |   ||||  |||   ||   ||   || || || |||||||||
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |   |
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||| | |||||||||| || ||||
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              |||||  |||||||||||||||||| |||||||| | |  |||||||||||||||| ||
Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||||| | ||||||||
Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  ||||||   ||   ||||||||  |||||||||||||||||||
Db        617 GAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 676

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||||||||||||| |  || |||||||  ||| |
Db        677 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 736
```

```
Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        737 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 796

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  ||||| || | |||| |||||| |
Db        797 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 856

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| |||||||||||||||||||||||||||||||||||  ||   ||    || ||
Db        857 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 916

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||| | ||||||||||||||||||||||||      ||||||||
Db        917 GAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 976

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db        977 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1036

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGA 1234
              ||||||| |||||||||||||| ||  |||||| ||| || |||||| |
Db       1037 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAA 1089
```

RESULT 11
DQ049206
LOCUS DQ049206 1089 bp DNA linear GSS 02-JUN-2005
DEFINITION Pan troglodytes TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION DQ049206
VERSION DQ049206.1 GI:66902405
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 1089)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 1089)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.
FEATURES Location/Qualifiers
source 1. .1089
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1. .>1089
/gene="TTRAP"
/locus_tag="HC17203"
ORIGIN

Query Match 49.9%; Score 655; DB 11; Length 1089;
Best Local Similarity 76.6%; Pred. No. 3.6e-167;
Matches 822; Conservative 0; Mismatches 242; Indels 9; Gaps 2;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   |  || |   | | ||  ||| |  |||||| | |||||
Db         17 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCCGAGGTGAAAAAGCGGC 76
```

```
Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||          |||| | ||| |   | |||||  || |  |||  |  ||
Db         77 GACTTCTGTGTNNNNNGTTTGCCTCGGTCGCAAGCTGCGATGCCGCTGTGGCTCAGTGCT 136

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        137 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 196

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   |||| |||  ||   ||  || ||         |||||
Db        197 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCANNNNNNNTGTTG 256

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||    ||  ||| |||||| |   |
Db        257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 316

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||    | ||| ||||||||||||||||||||||||||||||
Db        317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 376

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||||| | ||||||||||  || ||||
Db        377 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 436

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||||| |||||||| | |  ||||||||||||||||| ||
Db        437 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 496

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||||| | |||||||
Db        497 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 556

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        557 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 616

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   ||||||||  |||||||||||||||||
Db        617 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 676

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              |||||||||||||||||  ||| ||| |||||||||  |  || |||||||  ||| |
Db        677 ATTTGGAGAGCACCAGAGGGCATGCTGTGGAACGAATGAATCAGTTAAAAATGGTTTTAA 736

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||||||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        737 AGAAAATGCAAGAGGCTCCAGAATCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 796

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||  ||||| || | |||| |||||| |
Db        797 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGT 856

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| |||||||||||||||||||||||||||||||||||  ||   ||    || ||
Db        857 TTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG 916

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAG 1121
              | ||  |||||||||| ||||| |||||||||||||||||||||||||      ||||||||
Db        917 GAATAACTGCTGCTTGTAAGCTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAG 976

Qy       1122 AGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTA 1181
              |||| ||| ||||||| | |||||| |||||| |||| || ||||||||||||||||||
Db        977 AGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1036

Qy       1182 GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGA 1234
              ||||||| |||||||||||||| ||  |||||| ||| |||||  |||||| |
Db       1037 GATTTCCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATGTAATATTGTAA 1089
```

RESULT 12
BI656151
LOCUS BI656151 963 bp mRNA linear EST 12-SEP-2001

```
DEFINITION  603283518F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5327572 5',
            mRNA sequence.
ACCESSION   BI656151
VERSION     BI656151.1  GI:15570387
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 963)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
            Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11831  row: d  column: 05
            High quality sequence stop: 693.
FEATURES             Location/Qualifiers
     source          1. .963
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="NMRI"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5327572"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam4"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Priscilla Furth,
                     NIH Reference for transgenic model: Li et al., Cell Growth
                     and  Differentiation 7, 3-11 (1996)."
ORIGIN
```

```
  Query Match             49.5%;  Score 649.4;  DB 3;  Length 963;
  Best Local Similarity   87.3%;  Pred. No. 1.2e-165;
  Matches  788;  Conservative    0;  Mismatches   91;  Indels   24;  Gaps    6;

Qy          71 GAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCT 130
               | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          60 GCGCGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCT 119

Qy         131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 190
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         120 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCC 179

Qy         191 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 250
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         180 GAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTT 239

Qy         251 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 310
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         240 GCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGG 299

Qy         311 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 370
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         300 CAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG 359

Qy         371 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 430
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         360 CGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCA 419
```

```
Qy         431 AATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACT 490
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         420 AATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACT 479

Qy         491 ATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCT 550
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         480 ATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCT 539

Qy         551 CGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTT 610
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         540 CGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTT 599

Qy         611 ATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTA-CAGG 669
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db         600 ATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACCAGG 659

Qy         670 TAATGAAGAAGGATATTTCACAGC-TATACTATTGAAGAAAGG----AAGAGTGAAATTT 724
               || |||||| |||||||||||||| |||||||||||||   || |      || | ||  ||
Db         660 TACTGAAGACGGATATTTCACAGCTTATACTATTGAACGAACGGACAGAGTGAGACTTTA 719

Qy         725 AAAAGTCAGGAG--ATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCG 782
               || |||||||||    | | |||||||||||| ||||||||||||||||||||||||||||
Db         720 AACAGTCAGGAGCATTATTCCCTTTTCCAAATGCCAAAATGATGAGAAACCTGCTATGCG 779

Qy         783 TAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACAT----------CCCATTTG 832
               | ||  || |  |  | ||     | ||     |   | | |          ||  || |
Db         780 TTAAATTGTGAGTTTGGGGTGGGAACTTGAATTTCTGGCCTTATTGAACATCCCATTTGG 839

Qy         833 GAGAGCACCAGAGAACATTCTGCGGAACGAA------TAAGACAATTAAAAACTGTTCTT 886
               ||||||  | ||| |   |    |      ||       || |||  |||||||||| |||
Db         840 GAGAGCCACCGAGGAACATTCCTGCGGGAAACGGAATTAGGACCCTTAAAAACTGGTCTG 899

Qy         887 GGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTA 946
               || |||||||   |||| ||| ||||  |   ||| || ||  | |||    ||| ||
Db         900 GGCAAAATGCCGAAGGCCCCAAATTCCCCACGGGTAATCTTGCCGGGGACTACCAACTTGG 959

Qy         947 AGA 949
               |||
Db         960 AGA 962
```

```
RESULT 13
AA536734
LOCUS       AA536734                 647 bp    mRNA    linear   EST 29-JUL-1997
DEFINITION  vj88b09.r1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
            IMAGE:944153 3', mRNA sequence.
ACCESSION   AA536734
VERSION     AA536734.1  GI:2282727
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:541009
            High quality sequence stop: 473.
```

```
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J x DBA/2J F1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:944153"
                     /tissue_type="embryo"
                     /dev_stage="2-cell"
                     /lab_host="DH10B"
                     /clone_lib="Knowles Solter mouse 2 cell"
                     /note="Organ: embryo; Vector: pBluescribe (modified);
                     Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
                     mRNA prepared from 13,500 2-cell stage  embryos.  Primer:
                     SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTTT-3'.   cDNAs
                     were cloned into the MluI/SalI sites of a modified
                     pBluescribe  vector using commercial linkers (NEB).
                     Average insert size:  1.2 kb."
ORIGIN

  Query Match           49.3%;  Score 647;  DB 1;  Length 647;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-165;
  Matches  647;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        666 CAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTA 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTA 60

Qy        726 AAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAA 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAA 120

Qy        786 ATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAG 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAG 180

Qy        846 AACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTC 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTC 240

Qy        906 CAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCA 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCA 300

Qy        966 AATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAAC 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAAC 360

Qy       1026 ATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATA 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATA 420

Qy       1086 AGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTT 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTT 480

Qy       1146 TAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGC 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGC 540

Qy       1206 TCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTG 1265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTG 600

Qy       1266 TTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 647


RESULT 14
BG083167
LOCUS       BG083167                644 bp    mRNA    linear   EST 18-DEC-2003
DEFINITION  H3085C02-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
            H3085C02 5', mRNA sequence.
ACCESSION   BG083167
VERSION     BG083167.2  GI:40069892
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 644)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
            Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,
            Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
  PUBMED    10922068
COMMENT     On Jan 26, 2001 this sequence version replaced gi:12565735.
            Other_ESTs: H3085C02-3
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3085  row: C  column: 02
            Seq primer: -21M13 Reverse
            High quality sequence stop: 644
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .644
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:H3085C02-5"
                     /db_xref="taxon:10090"
                     /clone="H3085C02"
                     /sex="Clones arrayed from a variety of cDNA libraries"
                     /dev_stage="Clones arrayed from a variety of cDNA
                     libraries"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse 15K cDNA Clone Set"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 15,247 clones from 11
                     embryo  cDNA libraries (including preimplantation stage
                     embryos from  unfertilized egg to blastocyst, embryonic
                     part of E7.5 embryos,  extraembryonic part of E7.5
                     embryos, and E12.5 female mesonephros/gonad)  and one
                     newborn ovary cDNA library. Average insert size 1.5 kb.
                     All  source libraries are cloned unidirectionally with
                     Oligo(dT)-Not primers.  References include: (1)
                     Genome-wide expression profiling of mid-gestation
                     placenta and embryo using a 15,000 mouse developmental
                     cDNA microarray,  2000, Proc. Natl. Acad. Sci. U S A, 97:
                     9127-9132; (2) Large-scale cDNA  analysis reveals phased
                     gene expression patterns during preimplantation  mouse
                     develolpment, 2000, Development, 127: 1737-1749; (3)
                     Genome-wide  mapping of unselected transcripts from
                     extraembryonic tissue of 7.5-day  mouse embryos reveals
                     enrichment in the t-complex and under-representation  on
                     the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
ORIGIN

  Query Match           49.0%;  Score 642.4;  DB 2;  Length 644;
  Best Local Similarity   99.8%;  Pred. No. 8.5e-164;
  Matches  643;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        356 GACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTA 415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTA 60

Qy        416 ACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTA 475
```

```
Db        61 ACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTA 120

Qy       476 GAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAAT 535
Db       121 GAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAAT 180

Qy       536 CTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTA 595
Db       181 CTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTA 240

Qy       596 TTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTAC 655
Db       241 TTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTAC 300

Qy       656 ACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGA 715
Db       301 ACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGA 360

Qy       716 GTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTG 775
Db       361 GTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAGCCTG 420

Qy       776 CTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAG 835
Db       421 CTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAG 480

Qy       836 AGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATG 895
Db       481 AGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATG 540

Qy       896 CAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAA 955
Db       541 CAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAA 600

Qy       956 GAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGC 999
Db       601 GAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGC 644
```

```
RESULT 15
CR592636
LOCUS       CR592636              1620 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DM007YE11 of Fetal liver of Homo sapiens
            (human).
ACCESSION   CR592636
VERSION     CR592636.1  GI:50473443
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1620)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1620)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1..1620
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DM007YE11"
                     /tissue_type="Fetal liver"
                     /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match             47.6%;  Score 625;  DB 4;  Length 1620;
  Best Local Similarity   78.7%;  Pred. No. 6.3e-159;
  Matches  774;  Conservative    0;  Mismatches  200;  Indels    9;  Gaps    2;

Qy       217 GGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCAC 276
Db         6 GGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGC 65

Qy       277 GATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTA 336
Db        66 AGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTA 125

Qy       337 CTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTC 396
Db       126 CTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCC 185

Qy       397 CGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAG 456
Db       186 CAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAG 245

Qy       457 TCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATAT 513
Db       246 CCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATAT 305

Qy       514 TGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGC 573
Db       306 TGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGC 365

Qy       574 TTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCT 633
Db       366 TTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCT 425

Qy       634 AAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGC 693
Db       426 AAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGC 485

Qy       694 TATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAA 753
Db       486 TATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAG 545

Qy       754 TACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTG 813
Db       546 TACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTG 605

Qy       814 CCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATT 873
Db       606 CCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTT 665

Qy       874 AAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGG 933
Db       666 AAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGG 725

Qy       934 AGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTT 993
Db       726 AGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTGT 785

Qy       994 TGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGC 1053
Db       786 GGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAAT 845

Qy      1054 AAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG 1113
Db       846 GAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAG 905

Qy      1114 ------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAA 1167
```

```
                  |||||||||||| ||| ||||||| | |||||| |||||| |||| || |||||
Db        906 AGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGGATTAGAAAA 965

Qy       1168 ACTGGACTGTGGTAGATTTCCGA 1190
                  ||||||||||||||||||||| |
Db        966 ACTGGACTGTGGTAGATTTCCTA 988
```

RESULT 16
CN694288
LOCUS CN694288 617 bp mRNA linear EST 18-MAY-2004
DEFINITION E0348E09-5 NIA Mouse E10.5 whole embryo cDNA library (Long) Mus musculus cDNA clone NIA:E0348E09 IMAGE:30863576 5', mRNA sequence.
ACCESSION CN694288
VERSION CN694288.1 GI:47463037
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 617)
AUTHORS Sharov,A.A., Piao,Y., Matoba,R., Dudekula,D.B., Qian,Y., VanBuren,V., Falco,G., Martin,P.R., Stagg,C.A., Bassey,U.C., Wang,Y., Carter,M.G., Hamatani,T., Aiba,K., Akutsu,H., Sharova,L., Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S., Nagaraja,R., Boheler,K.R., Taub,D., Hodes,R.J., Longo,D.L., Schlessinger,D., Keller,J., Klotz,E., Kelsoe,G., Umezawa,A., Vescovi,A.L., Rossant,J., Kunath,T., Hogan,B.L., Curci,A., D'Urso,M., Kelso,J., Hide,W. and Ko,M.S.
TITLE Transcriptome analysis of mouse stem cells and early embryos
JOURNAL PLoS Biol. 1 (3), 410-419 (2003)
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: E0348 row: E column: 09
Seq primer: M13 Reverse
High quality sequence stop: 617
POLYA=No.

FEATURES Location/Qualifiers
source 1. .617
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:E0348E09-5"
/db_xref="taxon:10090"
/clone="NIA:E0348E09 IMAGE:30863576"
/tissue_type="whole embryo including extraembryonic tissues at 10.5-days postcoitum"
/dev_stage="E10.5"
/lab_host="DH10B"
/clone_lib="NIA Mouse E10.5 whole embryo cDNA library (Long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were extracted from a pool of 8 embryos at 10.5-days postcoitum. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from 2ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.4Kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 47.0%; Score 617; DB 7; Length 617;
Best Local Similarity 100.0%; Pred. No. 7.2e-157;
Matches 617; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        493 TTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCG 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCG 60

Qy        553 AGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTAT 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTAT 120

Qy        613 CCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAA 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAA 180

Qy        673 TGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 240

Qy        733 GGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAG 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAG 300

Qy        793 TTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTC 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTC 360

Qy        853 TGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTC 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTC 420

Qy        913 AACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGG 972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGG 480

Qy        973 TGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCA 540

Qy       1033 GTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCG 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCG 600

Qy       1093 TTTTGATCGAATATTTT 1109
              |||||||||||||||||
Db        601 TTTTGATCGAATATTTT 617
```

RESULT 17
CR602029
LOCUS CR602029 1743 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DI042YL19 of Placenta Cot 25-normalized of Homo sapiens (human).
ACCESSION CR602029
VERSION CR602029.1 GI:50482836
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1743)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished

```
  REMARK     Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
             Paraday Avenue
REFERENCE    2  (bases 1 to 1743)
  AUTHORS    Genoscope.
  TITLE      Direct Submission
  JOURNAL    Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT      1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
             end enriched, double-strand cDNA was digested with Not I and cloned
             into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
             was normalized. Library was constructed by Life Technologies, a
             division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1743
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI042YL19"
                     /tissue_type="Placenta Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match             47.0%;  Score 616.6;  DB 4;  Length 1743;
  Best Local Similarity   79.7%;  Pred. No. 1.3e-156;
  Matches  756;  Conservative    0;  Mismatches  184;  Indels    9;  Gaps    2;

Qy        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376
              ||||  || |||| | ||||||||||||  || | | | ||    |  | |   ||||
Db         32 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 91

Qy        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436
              |||   ||   ||   || || || |||||||||| |||||||| || ||  ||| ||||
Db         92 CCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 151

Qy        437 ACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATT 493
              | |||| ||   ||  ||| ||||||| |   ||||  | |||| || ||||||   |
Db        152 TCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTC 211

Qy        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553
              ||| ||||||||||||||||||||||||||||  |  ||||||| | ||||||||||||
Db        212 TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 271

Qy        554 GGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATC 613
              |||||||||||| | |||||||||| || |||||||||| ||||||||||||||||||||
Db        272 GGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATT 331

Qy        614 CCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAAT 673
              |||||||| | |  |||||||||||||||| ||   | |||   ||||||||||| ||
Db        332 CCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCAT 391

Qy        674 GAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAG 733
              |||||||||||||||||||||| | |||||||||  |||||||||||| ||||| ||
Db        392 GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAA 451

Qy        734 GAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT 793
              |||||||||||||||||| |||||||||||||||||||  |||| ||  ||||||
Db        452 GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAC 511

Qy        794 TTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCT 853
              ||   ||||||||  |||||||||||||||||||||||||||||||||||||  ||| ||
Db        512 GTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCT 571

Qy        854 GCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCA 913
              |||||||||| |  || |||||||  ||| |    |||||||||||||||||||| |||
Db        572 GCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCA 631

Qy        914 ACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGT 973
              | || |||||||||||||||||||||| |||| |||| ||| |||| || ||||||||
Db        632 GCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGT 691

Qy        974 GGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAG 1033
              |||||||  |||| || | |||| |||||| ||||| ||||||||||||||||||||
Db        692 GGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAG 751

Qy       1034 TATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGT 1093
              ||||||||||||| ||   ||    || ||  | ||  |||||||||| ||| | ||||
Db        752 TATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGT 811

Qy       1094 TTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTA 1147
              ||||||||||||||||||||      |||||||||||| ||| ||||||| | ||||||
Db        812 TTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 871

Qy       1148 GACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTC 1207
              |||||| |||| || ||||||||||||||||||||||||||| |||||||||||||| ||
Db        872 GACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTT 931

Qy       1208 TTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              ||||||  |||  || || ||||||  ||| ||||  ||  ||  | |||
Db        932 CTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGTGGGTTT 980


RESULT 18
BY719852
LOCUS       BY719852                703 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  BY719852 RIKEN full-length enriched, in vitro fertilized eggs Mus
            musculus cDNA clone 7420437P03 5', mRNA sequence.
ACCESSION   BY719852
VERSION     BY719852.1  GI:27132969
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 703)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
```

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1..703
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7420437P03"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, in vitro fertilized eggs"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

ORIGIN

Query Match 46.6%; Score 612; DB 5; Length 703;
Best Local Similarity 97.5%; Pred. No. 1.7e-155;
Matches 629; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

```
Qy         68 CACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCG 127
              ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 CAGTCGCGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCG 80

Qy        128 TCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCG 187
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db         81 TCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGT-GGCG 139

Qy        188 CCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 CCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 199

Qy        248 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 259

Qy        308 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 319

Qy        368 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 379

Qy        428 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGC 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGC 439

Qy        488 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 547
              |||||||||||||||||||| |||||||||||||||||||||||  ||||||||||||||
Db        440 ACTATTTCTTTCATTACCTGNAATATTGATGGATTAGATGGATGNCATCTGCCCGAGAGG 499

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 GCTCGAGGNGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 559

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        560 NGTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACANTTATTACA 619

Qy        668 GGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 712
              ||  |||||||||||||||||||||||||||||||  ||||||||
Db        620 GGGTATGAAGAAGGATATTTCACAGCTATACTATTTGAGAAAGGA 664
```

RESULT 19
BB609556
LOCUS BB609556 712 bp mRNA linear EST 26-OCT-2001
DEFINITION BB609556 RIKEN full-length enriched, 18 days embryo Mus musculus cDNA clone 1110007K13 5', mRNA sequence.
ACCESSION BB609556
VERSION BB609556.1 GI:16451282
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 712)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.

FEATURES Location/Qualifiers
source 1. .712
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="1110007K13"
/dev_stage="18 days embryo"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, 18 days embryo"
/note="Site_1: XhoI; Site_2: SstI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAAGGATCCAAGAGCTCAATTAATTAATTAAACCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI."

ORIGIN

Query Match 46.4%; Score 608.6; DB 2; Length 712;
Best Local Similarity 95.6%; Pred. No. 1.5e-154;
Matches 670; Conservative 0; Mismatches 24; Indels 7; Gaps 4;

```
Qy         68 CACGAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCG 127
              ||   | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 CAGTCGCGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCG 70

Qy        128 TCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCG 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 TCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCG 130

Qy        188 CCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 CCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGC 190

Qy        248 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 TTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGAC 250

Qy        308 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 TGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG 310

Qy        368 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 CCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGAT 370

Qy        428 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGC 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 GCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGC 430

Qy        488 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGG 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 ACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCAAGAGG 490

Qy        548 GCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 GCTCGAGAGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGGA 550

Qy        608 GTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACA 667
              ||| |||||||||||||||| |||||| ||||||| |||||||||||||| ||| |||||
Db        551 GTTTTCCCCCCATACTGTGGCTACCTTAAGAAGAAAGCAGCCAGTTACACCATTTTTACA 610

Qy        668 GGTAATG-AAGAAGGATATTTCACAGCTATAC---TATTGAAGAAAGGAAGAGTG-AAAT 722
              ||||||| |||||||||||||||||||||| |   |    |  || ||||||||| ||||
Db        611 GGTAATGAAAGAAGGATATTTCACAGCTATTCCTATTGAAAGAAAGGGAAGAGTGAAAAT 670

Qy        723 TTAAAAGTCAGGAGATTATT--CCTTTTCCAAATACCAAAA 761
              ||||||||| |||||||| ||  | ||||||||||||||||
Db        671 TTAAAAGTCGGGAGATTTTTTCCTTTTTCCAAATACCAAAA 711
```

RESULT 20
AW321603
LOCUS AW321603 648 bp mRNA linear EST 25-JAN-2000
DEFINITION uo36b06.y1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:2631539 5' similar to TR:O95551 O95551 DJ30M3.3 ;, mRNA sequence.
ACCESSION AW321603
VERSION AW321603.1 GI:6751147
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 648)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

MGI:1025047
Seq primer: -40RP from Gibco
High quality sequence stop: 436.

FEATURES Location/Qualifiers
source 1. .648
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:2631539"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Query Match 46.2%; Score 606.8; DB 1; Length 648;
Best Local Similarity 95.8%; Pred. No. 4.5e-154;
Matches 620; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```
Qy        104 TGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGG 163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCCCGCAGGG 60

Qy        164 CCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAG 223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAG 120

Qy        224 AGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTC 283
              |||||||||||||||||||||||||||||||||||||||||||||||| | | ||||||
Db        121 AGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCNCNCNATGGTC 180

Qy        284 CCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAG 343
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CNCANCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAG 240

Qy        344 CTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCC 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCC 300

Qy        404 TATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCT 360

Qy        464 GGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTA 420

Qy        524 GATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGT 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGT 480

Qy        584 CCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGA 643
              |||||||||||||||||||||||||||| |  ||||||||||||||||||||||||||||||
Db        481 CCAGATGTGGTATTTCTACAGGAAGTTATACAGCCATACTGTGCCTACCTAAAGAAGAGA 540

Qy        644 GCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTG 703
              ||   |||||||||||||||||| |||||  | |||||| |||||||||| |||||||
Db        541 GCCAGCAGTTACACAATTATTACATGTAATTGAAAAGGATGTTTCACAGCTTTACTATTC 600

Qy        704 AAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCC 750
              ||||||| ||| ||| | |||| ||||| |||||| ||||| ||||
Db        601 GAGAAAGGGAGAATGAGACTTAAGAGTCACGAGATTTTTCCTGTTCC 647
```

RESULT 21
BI990829

```
LOCUS       BI990829                600 bp    mRNA    linear   EST 20-DEC-2001
DEFINITION  4081-88 Mouse E14.5 retina lambda ZAP II Library Mus musculus cDNA,
            mRNA sequence.
ACCESSION   BI990829
VERSION     BI990829.1  GI:17961846
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Mu,X., Zhao,S., Pershad,R., Hsieh,T.-F., Scarpa,A., Wang,S.W.,
            White,R.A., Beremand,P.D., Thomas,T.L., Gan,L. and Klein,W.H.
  TITLE     Gene expression in the developing mouse retina by EST sequencing
            and microarray analysis
  JOURNAL   Nucleic Acids Res. 29 (24), 4983-4993 (2001)
   PUBMED   11812828
COMMENT     Contact: Klein WH
            Department of Biochemistry and Molecular Biology
            University of Texas M.D. Anderson Cancer Center
            Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
            Tel: 713 792 3646
            Fax: 713 790 0329.
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="neural retina"
                     /dev_stage="embryonic day 14.5 post-fertilization"
                     /clone_lib="Mouse E14.5 retina lambda ZAP II Library"
ORIGIN
```

```
  Query Match             45.7%;  Score 600;  DB 3;  Length 600;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-152;
  Matches  600;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         89 GAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCG 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCG 60

Qy        149 GCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAG 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAG 120

Qy        209 GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGC 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGC 180

Qy        269 GACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTG 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTG 240

Qy        329 AGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCC 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCC 300

Qy        389 TTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTA 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTA 360

Qy        449 GAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGG 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGG 420

Qy        509 AATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGC 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGC 480

Qy        569 CTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCC 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCC 540

Qy        629 TACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTC 688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTC 600
```

RESULT 22
CN803230

```
LOCUS       CN803230                965 bp    mRNA    linear   EST 26-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_32551 Katze_MMBR Macaca mulatta cDNA clone
            IBIUW:12605 5' similar to Bases 1 to 965 highly similar to human
            TTRAP (Hs.210628), mRNA sequence.
ACCESSION   CN803230
VERSION     CN803230.1  GI:47699206
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 965)
  AUTHORS   Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
  TITLE     Analysis of the Macaca mulatta transcriptome and the sequence
```

```
            divergence between Macaca and human
  JOURNAL   Genome Biol. 6 (7), R60 (2005)
   PUBMED   15998449
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.04.02. 780 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 965   Std Error: 0.00
            Plate: CL000288  row: D  column: 11
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .965
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:12605"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match           45.2%;  Score 593.6;  DB 7;  Length 965;
  Best Local Similarity  81.4%;  Pred. No. 2e-150;
  Matches  715;  Conservative    0;  Mismatches  154;  Indels    9;  Gaps    2;

Qy        388 CTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTT 447
              ||   || || || |||||||||| |||||||| || ||  ||| |||| | |||| ||
Db          1 CTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTC 60

Qy        448 AGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTAC 504
              ||  ||| ||||||| |   |||| | |||| ||  |||||   | |  ||||||||
Db         61 TAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGACAGCATGTTCTGCTTCATTAC 120

Qy        505 CTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTC 564
              ||||||||||||||||||||||  |  ||||||| | |||||||||||||||||||||||
Db        121 CTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTC 180

Qy        565 CTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTG 624
              || | |||| ||||| || |||||||| |||||||||||||||||| |||||||| |
Db        181 CTACTTAGCGTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTA 240

Qy        625 TGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATA 684
              |  ||||||||||||||||||     |||    |||||||||||| ||||||||||||
Db        241 TAGCTACCTAAAGAAGAGAGCAAGTGATTATGAGATTATTACAGGTCATGAAGAAGGATA 300

Qy        685 TTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCC 744
              |||||||||||| | |||||||||  |||||||||||| ||||| || ||||||||||
Db        301 TTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCC 360

Qy        745 TTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAA 804
              |||||||| ||||||||||||||||||||||  |||| ||  |||||| | ||    |||||
Db        361 TTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAA 420

Qy        805 TGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAAT 864
              |||  |||||||||||||||||||||||||||||||||||  ||| ||||||||||||
Db        421 TGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAAT 480

Qy        865 AAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTAT 924
              |  || |||||||  ||| |    |||||||||||||||||||| ||| | || |||||
Db        481 GAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTAT 540

Qy        925 ATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGA 984
              |||||||||||||||||||| |||| |||||||| |||| ||||||||||||||||||  |
Db        541 ATTTGCAGGAGATACAAATCTAAGGGATCAAGAGGTTACCAAATGTGGTGGTTTACCCAA 600

Qy        985 CAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGA 1044
              |||| || | |||| |||||| ||||| ||||||||||||||||||||||||||||||||
Db        601 CAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGA 660

Qy       1045 TACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAAT 1104
              |||  ||   ||    || ||  | ||  ||||||||| ||||| ||||||||||||||
Db        661 TACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAGCTTCGTTTTGATCGAAT 720

Qy       1105 ATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGG 1158
              |||||||||      |||||||||||| ||| ||||||| | ||||||  |||||| ||||
Db        721 ATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTGGACCTTCTTGG 780

Qy       1159 GTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTT 1218
              |||||||||| |||||||||||||||| |||||||||||||  ||  ||||||  ||
Db        781 ATTGGAAAAACTCGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTTCTGTGCAATTT 840

Qy       1219 GAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              |||||| |||||| ||| ||||  ||  ||    |||
Db        841 AAATGTAATATTGTAAAATGCTTTTCAAATGTGAGTTT 878


RESULT 23
BG695251
LOCUS       BG695251                614 bp    mRNA    linear   EST 04-MAY-2001
DEFINITION  NISC_iv14c02.w2 Soares NMBP2 pituitary Mus musculus cDNA clone
            IMAGE:4317675 5', mRNA sequence.
ACCESSION   BG695251
VERSION     BG695251.1  GI:13954981
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 614)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
             DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
             MGI:1598443
            Plate: LLAM9922  row: E  column: 4
            Seq primer: T7 primer.
FEATURES             Location/Qualifiers
     source          1. .614
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4317675"
                     /tissue_type="pituitary gland"
                     /dev_stage="embryo, 14 dpc"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Soares NMBP2 pituitary"
                     /note="Organ: brain; Vector: pT7T3D-Pac; Site_1: NotI;
                     Site_2: EcoRI; 1st strand cDNA was primed with a NotI -
                     oligo(dT) primer
                     5'-AACTGGAAGAATTCGCGGCCGCGGCGCTTTTTTTTTTTTTTTTT-3';
                     double-stranded cDNA was ligated to EcoRI adaptors
                     5'-AATTCGGCACGAGG-3' AND 5'-CCTCGTGCCG-3'(Pharmacia),
                     digested with NotI and cloned into the NotI and EcoRI
```

sites of the pT7T3D-PacI vector. Library went through one round of normalization, and was constructed in the laboratory of M. Bento Soares (University of Iowa)."

ORIGIN

Query Match 45.1%; Score 591.8; DB 2; Length 614;
Best Local Similarity 99.5%; Pred. No. 5.5e-150;
Matches 604; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
Qy        593 GTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGT 652
              | ||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
Db          9 GGATTTCTCCAGGAAGTTATCCCCCC-TACTGTGCCTACCTAAAGAAGAGAGCAGCCAGT 67

Qy        653 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 127

Qy        713 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAAC 772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAAC 187

Qy        773 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTG 247

Qy        833 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 GAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAA 307

Qy        893 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGAT 367

Qy        953 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 CAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTA 427

Qy       1013 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATC 487

Qy       1073 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTT 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTT 547

Qy       1133 ATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGT 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 ATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGT 607

Qy       1193 GATCACT 1199
              |||||||
Db        608 GATCACT 614
```

RESULT 24
BF182465
LOCUS BF182465 621 bp mRNA linear EST 31-OCT-2000
DEFINITION 601804232F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035319 5', mRNA sequence.
ACCESSION BF182465
VERSION BF182465.1 GI:11060608
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 621)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9308 row: p column: 08
High quality sequence stop: 615.
FEATURES Location/Qualifiers
source 1. .621
/organism="Mus musculus"
/mol_type="mRNA"
/strain="mix FVB/N, C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4035319"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"
ORIGIN

Query Match 44.8%; Score 587.4; DB 2; Length 621;
Best Local Similarity 99.5%; Pred. No. 8.8e-149;
Matches 610; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
Qy         97 GGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCC 156
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 GGGGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCAGTTCCGATGCGGCGGAGCC 68

Qy        157 CGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCG 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCG 128

Qy        217 GGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCAC 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 GGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCAC 188

Qy        277 GATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTA 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 GATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTA 248

Qy        337 CTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTC 396
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        249 CTTCGAGCTGCCAGAGAACGA-CAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTC 307

Qy        397 CGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAG 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAG 367

Qy        457 TCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGA 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 TCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGA 427

Qy        517 TGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTT 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 TGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTT 487

Qy        577 GTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAA 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 GTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAA 547

Qy        637 GAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAG-AAGGATATTTCACAGCTA 695
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        548 GAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGCAAGGATATTTCACAGCTA 607

Qy        696 TACTATTGAAGAA 708
              |||||||||||||
```

```
Db        608 TACTATTGAAGAA 620


RESULT 25
CV119018
LOCUS       CV119018                734 bp    mRNA    linear   EST 30-AUG-2004
DEFINITION  AGENCOURT_31557860 NIH_MGC_270 Rattus norvegicus cDNA clone
            IMAGE:7444574 5', mRNA sequence.
ACCESSION   CV119018
VERSION     CV119018.1  GI:51640419
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 734)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Josef Lazar, Dr. Howard J. Jacob, Medical
            College of Wisconsin
             cDNA Library Preparation: Open Biosystems
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15744  row: h  column: 12
            High quality sequence start: 18
            High quality sequence stop: 735.
FEATURES             Location/Qualifiers
     source          1. .734
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="IMAGE:7444574"
                     /tissue_type="whole placenta, 2 pooled"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_270"
                     /note="Organ: placenta; Vector: pExpress-1; Site_1: EcoRV;
                     Site_2: NotI; Tissue was collected from two pooled
                     placentas from the 21st day of pregnancy. 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer,
                     double-stranded cDNA was cloned into the Not I and EcoRV
                     sites of pExpress-1. Library was size-selected for >1.25
                     kb fragments for an average insert size of 2.15 kb.
                     Library was normalized to Cot7. A non-normalized version
                     of this library is also available (NIH_MGC_269). Library
                     was constructed by Open Biosystems (Huntsville, AL). Note:
                     this is a Mammalian Gene Collection library"
ORIGIN

  Query Match             44.7%;  Score 586.8;  DB 7;  Length 734;
  Best Local Similarity   89.6%;  Pred. No. 1.3e-148;
  Matches  630;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy        226 GCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCC 285
              | |  |||||||||| |||||| ||||||||||||   | | ||||||  | |||  |||    |
Db         32 GGGATGGCTTCAGTCCCTGGGATTTGCGTTGGTGACGAGCTGCGACACTACGGTGGCTTC 91

Qy        286 CAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCT 345
              ||   |||||  ||||||| |||||||||| | |||||  |||||||||  ||||||||
Db         92 CACTTTCCTGTCGGAGAACAACTGGCAGACGAAAAAAGCGTTGAGCGCCTTCTTCGAGCA 151

Qy        346 GCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTA 405
              ||||||||||||||| || | |||| | ||||||||| || |||| |||||||||||| |||
Db        152 GCCAGAGAACGACCTAGCGCGGCCTCACCAGCCTCCGACATCCTCCAAGTCCGAGGACTA 211

Qy        406 TGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGG 465
              |||||||||||||||| |||||||||||||||| |||||||||||| ||||||||||||
Db        212 TGTTGATCTAACCAATGAGGATGCAAATGATACCACCATTTTAGAAACCAGTCCATCTGG 271

Qy        466 AACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGA 525
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        272 AACTCCTCTAGAAGATAGCAGCACTATCTCTTTCATTACCTGGAATATTGATGGATTAGA 331

Qy        526 TGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCC 585
              |||||||||||  || |||||||||||||||||||||||||||| ||||||||  |||||
Db        332 TGGATGCAATCTCCCAGAGAGGGCTCGAGGGGTGTGTTCCTGCCTTGCTTTGTACAGTCC 391

Qy        586 AGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGC 645
              ||  |||||||||||||||||||| ||  ||| | ||||||||||||||| |||||||||
Db        392 GGACGTGGTATTTCTACAGGAAGTGATTCCCTCCTACTGTGCCTACCTAAGGAAGAGAGC 451

Qy        646 AGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAA 705
              |  ||  |||||  |||||||||||||||||||||||||||||||||||||||| |||||||
Db        452 ACGCACTTACAACATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATATTATTGAA 511

Qy        706 GAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGAT 765
              |||||||||||||||||||||||||||||||| || ||||||||||||||||||||||||
Db        512 GAAAGGAAGAGTGAAATTTAAAAGTCAAGAAATTATTCCTTTTCCAAATACCAAAATGAT 571

Qy        766 GAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATC 825
              |||||||| || || ||||||||||||||||||||||||||||||||||||||||||||||
Db        572 GAGAAACCTTCTGTGTGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATC 631

Qy        826 CCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCT 885
              |||||||||||||| ||| ||||||||| |||||||||  ||||||||||||||| |
Db        632 CCATTTGGAGAGCACAAGANAACATTCTGCTGAACGAATAAATCAATTAAAAACTGTTTT 691

Qy        886 TGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTT 928
              |  |||||||||||||||| |||||||||| || |||||||||
Db        692 TCAAAAAATGCAAGAGGCTACAGATTCAACTACTGTTATATTT 734


RESULT 26
BM503555
LOCUS       BM503555                580 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  ih33e07.y1 Melton Mouse E16 5 Pancreas Library 2 M16B2 Mus musculus
            cDNA clone IMAGE:5679061 5' similar to TR:O95551 O95551 DJ30M3.3 ;,
            mRNA sequence.
ACCESSION   BM503555
VERSION     BM503555.1  GI:18665499
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Juliana Brown
            (brown@fas.harvard.edu)
```

```
            MGI:1956939 This sequence now available from the IMAGE consortium,
            for clone orders contact: info@image.llnl.gov
            Seq primer: -40RP from Gibco
            High quality sequence stop: 432.
FEATURES              Location/Qualifiers
     source           1. .580
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="ICR"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:5679061"
                      /sex="Both"
                      /tissue_type="Total pancreas"
                      /dev_stage="Embryonic day 16.5"
                      /lab_host="TOP10"
                      /clone_lib="Melton Mouse E16 5 Pancreas Library 2 M16B2"
                      /note="Organ: Pancreas; Vector: pBluescript II SK; Site_1:
                      NotI; Site_2: SalI; Library constructed using SuperScript
                      Plasmid Library kit (Life  Technologies). cDNA made by
                      oligo-dT priming.Size-selected by  column fractionation;
                      average insert size 1.06kb. Primary  library,
                      unamplified."
ORIGIN

  Query Match             44.2%;  Score 580;  DB 3;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 9e-147;
  Matches  580;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        587 GATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCA 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCA 60

Qy        647 GCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAG 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAG 120

Qy        707 AAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATG 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATG 180

Qy        767 AGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCC 240

Qy        827 CATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTT 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTT 300

Qy        887 GGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTA 360

Qy        947 AGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAA 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAA 420

Qy       1007 TTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTC 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTC 480

Qy       1067 AGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGG 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGG 540

Qy       1127 CACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAA 1166
              |||||||||||||||||||||||||||||||||||||||
Db         541 CACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAA 580


RESULT 27
BX444691
LOCUS       BX444691               1081 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX444691 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN005YN02 5-PRIME, mRNA sequence.
ACCESSION   BX444691
VERSION     BX444691.2  GI:47009162
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1081)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30780264.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DN005DG01QP1&c=3474.r.
FEATURES              Location/Qualifiers
     source           1. .1081
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DN005YN02"
                      /tissue_type="ADULT BRAIN"
                      /dev_stage="adult"
                      /clone_lib="Homo sapiens ADULT BRAIN"
                      /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                      was primed with a NotI-oligo(dT) primer. Five prime end
                      enriched, double-strand cDNA was digested with Not I and
                      cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                      vector. Library was not normalized."
ORIGIN

  Query Match             44.1%;  Score 579;  DB 5;  Length 1081;
  Best Local Similarity   76.6%;  Pred. No. 2e-146;
  Matches  784;  Conservative    3;  Mismatches  223;  Indels   13;  Gaps    6;

Qy        174 GGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGC 233
              ||  ||||  |  |  |  ||  |  | |  ||  |||     ||||||  | |||||| |
Db         46 GGMGGGCGCGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGAC 105

Qy        234 TTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCC 293
              |||  |||  |||  |||||  | |||  |   |  |||||  ||  |  |||  |   || |||
Db        106 TTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCC 165

Qy        294 TGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGA 353
              ||   ||||||||||||  |||  |  |  |  ||  ||||  |  ||||||||||||   ||  |  |
Db        166 TGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGG 225

Qy        354 ACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATC 413
              |  ||   |  |  ||||  |||   ||  ||  ||  ||  ||  ||  ||||||||||  |
Db        226 AGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACC 285

Qy        414 TAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTC 470
              |||||||  ||  ||  |||  ||||  |  ||||  ||    ||  |||  |||||||  |    ||||
Db        286 TAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTC 345

Qy        471 CTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGAT 530
              |  ||||  ||  ||||||    |  |||  ||||||||||||||||||||||||||   |
Db        346 AGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAA 405

Qy        531 GCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATG 590
              |||||||  |  ||||||||||||||||||||||  |  ||||||||||  ||  |||||||
Db        406 ACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATG 465
```

```
Qy       591 TGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCA 650
Db       466 TGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTA 525

Qy       651 GTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAG 710
Db       526 ATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAAT 585

Qy       711 GAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAA 770
Db       586 CAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAA 645

Qy       771 ACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATT 830
Db       646 ACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATT 705

Qy       831 TGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAA 890
Db       706 TGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGA 765

Qy       891 AAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAG 950
Db       766 AAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGG 825

Qy       951 ATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTT 1010
Db       826 ATCGAGAGGTTACCAGATGTGGTGGTTTACC-MACAACATTGTGGATGTCTGGGAGTTTT 884

Qy      1011 TAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGA 1070
Db       885 TGGGC-AACCTAAACATTGCCAGTATACATGGGATACACACATGAACTCTAATCTTGGAA 943

Qy      1071 TCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGG 1124
Db       944 TAACTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCAGARCAGCAGCAGAAGAGG 1003

Qy      1125 GGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGAT 1184
Db      1004 GACACATTATYCC-CGAAGTTTGGACCYTCTTGGATTAG-AAAACTGGACTGTWGTAGAT 1061

Qy      1185 TTC 1187
Db      1062 TYC 1064
```

RESULT 28
BX422491
LOCUS BX422491 904 bp mRNA linear EST 03-MAY-2004
DEFINITION BX422491 Homo sapiens FETAL LIVER Homo sapiens cDNA clone CS0DM007YE11 5-PRIME, mRNA sequence.
ACCESSION BX422491
VERSION BX422491.2 GI:46955237
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 904)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30766188.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DM007AC06QP1&c=3474.r.

FEATURES Location/Qualifiers
source 1. .904
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM007YE11"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 43.9%; Score 576; DB 5; Length 904;
Best Local Similarity 77.6%; Pred. No. 1.3e-145;
Matches 700; Conservative 7; Mismatches 192; Indels 3; Gaps 1;

```
Qy       214 CCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCC 273
Db         3 CGGGATGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCYGCGATGC 62

Qy       274 CACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGC 333
Db        63 CGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTC 122

Qy       334 CTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAA 393
Db       123 CTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGA 182

Qy       394 GTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGC 453
Db       183 GCCCAARACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAAT 242

Qy       454 CAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAA 510
Db       243 CAGCCCATCTGAAGATACTCAGCAARAAAAWGGCAGCATGTTCTCTCTCATTACCTGGAA 302

Qy       511 TATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCT 570
Db       303 TATTGATGGATTAGATCTAAACAAWCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTT 362

Qy       571 AGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTA 630
Db       363 AGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTA 422

Qy       631 CCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCAC 690
Db       423 CCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCAC 482

Qy       691 AGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCC 750
Db       483 AGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCC 542

Qy       751 AAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATT 810
Db       543 AAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCT 602

Qy       811 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACA 870
Db       603 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCA 662

Qy       871 ATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGC 930
Db       663 GTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGC 722

Qy       931 AGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGT 990
```

```
Db        723 AGGAGATACAAATCTAAGGGATCGAGAGGTTACCARATGTGGTGGTTTACCCAACAACAT 782

Qy        991 TTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAA 1050

Db        783 TGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACA 842

Qy       1051 AGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTT 1110

Db        843 AATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGTTTTGATCGAAWATTTTT 902

Qy       1111 CA 1112

Db        903 CA 904


RESULT 29
BX433489/c
LOCUS       BX433489                952 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX433489 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN005YN02 3-PRIME, mRNA sequence.
ACCESSION   BX433489
VERSION     BX433489.2  GI:47002503
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 952)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30775195.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAI056ZD12_CS05372_1&c=3474.r
            .
FEATURES             Location/Qualifiers
     source          1. .952
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             43.7%;  Score 573;  DB 5;  Length 952;
  Best Local Similarity   78.5%;  Pred. No. 8.5e-145;
  Matches  713;  Conservative    0;  Mismatches  186;  Indels    9;  Gaps    2;

Qy        373 CCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAA 432

Db        925 CGAACCTGAACCCATCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAACT 866

Qy        433 TGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCAC 489

Db        865 GNATTCCACACTTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCAT 806

Qy        490 TATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGC 549

Db        805 GTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGC 746

Qy        550 TCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGT 609

Db        745 TCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGT 686

Qy        610 TATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGG 669

Db        685 TATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGG 626

Qy        670 TAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAG 729

Db        625 TCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAG 566

Qy        730 TCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGT 789

Db        565 CCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGT 506

Qy        790 GAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACA 849

Db        505 GAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCA 446

Qy        850 TTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGA 909

Db        445 TGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGA 386

Qy        910 TTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATG 969

Db        385 GTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATG 326

Qy        970 TGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTG 1029

Db        325 TGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTG 266

Qy       1030 CCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCA 1089

Db        265 CCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACT 206

Qy       1090 TCGTTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAG 1143

Db        205 TCGTTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAG 146

Qy       1144 TTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGG 1203

Db        145 TTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGG 86

Qy       1204 GCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTT 1263

Db          85 TCTTCTGTGCAACTTAGATATAATATTGTAAAATGCTTTTCAAGTGGGGTTTTGCCCTGA 26

Qy       1264 TGTTAGCA 1271

Db          25 TTGTTGCA 18


RESULT 30
BY708937
LOCUS       BY708937                936 bp    mRNA    linear   EST 16-DEC-2002
DEFINITION  BY708937 RIKEN full-length enriched, adult male tongue Mus musculus
            cDNA clone 2310002D07 5', mRNA sequence.
ACCESSION   BY708937
VERSION     BY708937.1  GI:27120132
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 936)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
```

Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers

source 1. .936
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="2310002D07"
/sex="male"
/tissue_type="tongue"
/dev_stage="adult"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, adult male tongue"
/note="Site_1: XhoI; Site_2: SstI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. "

ORIGIN

Query Match 43.6%; Score 572.6; DB 5; Length 936;
Best Local Similarity 93.7%; Pred. No. 1.1e-144;
Matches 683; Conservative 0; Mismatches 36; Indels 10; Gaps 8;

```
Qy        73 GGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGG 132
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         6 GCGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGA-CCTGCGGCGATGGCGTCTGG 64

Qy       133 CAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGA 192
             |||||||||||||||||||| | |   ||| ||||||||||||||||||||||||||||||
Db        65 CAGCAGTTCCGATGCGGCGGACCGC---AGGCNGGCAGGGCGGGCGGCGTCGGCGCCCGA 121

Qy       193 AGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGC 252
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       122 AGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGC 181

Qy       253 GTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCA 312
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       182 GTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCA 241

Qy       313 GACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCG 372
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       242 GACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCG 301

Qy       373 CCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAA 432
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       302 -CAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAA 360

Qy       433 TGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTAT 492
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTAT 420

Qy       493 TTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCG 552
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       421 TTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGA-GGCTCG 479

Qy       553 AGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTAT 612
             ||||||||| |||||||||||||||||  |||||||||||||||||||||||||||||||||
Db       480 AGGGGTGTG-TCCTGCCTAGCTTTGTATAG-NCAGATGTGGTATTTCTACAGGAAGTTAT 537

Qy       613 CCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAA 672
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       538 -CCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAA 596

Qy       673 TGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 732
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       597 TGAAGAAGGATATTTCACAGCTATACTATTGAAGAAA-GAAGAGTGAAATTTAAAAGTCA 655

Qy       733 GGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAG 792
             |||||||||||||||||| | | |  |||| |      |||    ||  || | | |
Db       656 GGAGATTATTCCTTTTCAATACAAAATGATGAGAACCTCTATGCGTAATGTGGTTGGGTG 715

Qy       793 TTTGGGTGG 801
              ||  | |
Db       716 AATGATTTG 724
```

```
RESULT 31
BX337141
LOCUS       BX337141               1018 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX337141 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI036YJ02 5-PRIME, mRNA sequence.
ACCESSION   BX337141
VERSION     BX337141.2  GI:46271144
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1018)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30337585.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI036DE01QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .1018
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI036YJ02"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match           43.6%;  Score 572.4;  DB 5;  Length 1018;
  Best Local Similarity   72.2%;  Pred. No. 1.3e-144;
  Matches  729;  Conservative   29;  Mismatches  248;  Indels    4;  Gaps    2;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | ||:||    |  |   |  || |   | | ||  |||     ||||||  | |||||
Db          7 GCCTGGMGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 66

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db         67 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 126

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||  || |
Db        127 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 186

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   ||||  |||   ||   ||   || || || |||||||||
Db        187 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 246

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| |||||||  |   |
Db        247 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 306

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||    | ||| ||||||||||||||||||||||||||||||
Db        307 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 366

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  ||||||| | |||||||||||||||||||||||||| | |||||||||| || ||||
Db        367 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 426

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||||| |||||||| | |  ||||||||||||||||| ||
Db        427 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 486

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| |||||||||||||||||||||||||||| | |||||||
Db        487 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 546

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        547 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 606

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   ||||||||  ||||||||||||||||||
Db        607 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 666

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||:|||||  ||| ||||||||:|||| |  || |||||||  ||| |
Db        667 ATTTGGAGAGCAMCAGAGGGCATGCTGCGGAAMGAATGAATCAGTTAAAAATGGTTTTAA 726

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              |||||| |||||||||||||| ||| | || ||||||||||||||||||||||||| |||
Db        727 AGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 786

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||:   | ||: || | ||||  ||||||
Db        787 GGGATCGAGAGGTTACCAGATGTGGTGGTTTAMAMAAAAAMATTGTGGATGTATGGGAAG 846

Qy       1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
              |||| || :||:|||||||||::|||||||||||||||:  ||   ||  : || ||
Db        847 TTTTTGGGMAAMCTAAACATTGMMAGTATACATGGGATAMACAAATGAACTMTAATCTTG 906

Qy       1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAG-CAGAAGAGGGG 1126
              |:||  :|||||:|| | |   ||||||||||:||||||||||:|::|  :|:||| | |
Db        907 GRATAAMTGCTGMTTGTKAAATTCGTTTTGATMGAATATTTTTMARRGAGMGVAGAAGAG 966

Qy       1127 CACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTG 1176
                |  :|:   :    || |  : | |||| |:: ||||:  |:  :|
Db        967 GGRCAAWTWTTMMCSGAGTTTGGGMCTCTTGGATWRAAAAAHGGGVTGKG 1016

RESULT 32
CN642527
LOCUS       CN642527               1087 bp    mRNA    linear   EST 12-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_6530 Katze_MMPL2 Macaca mulatta cDNA clone IBIUW:4699
            5' similar to Bases 1 to 1084 highly similar to human TTRAP
            (Hs.210628), mRNA sequence.
ACCESSION   CN642527
VERSION     CN642527.1  GI:47153537
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 1087)
  AUTHORS   Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
  TITLE     Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
  JOURNAL   Genome Biol. 6 (7), R60 (2005)
   PUBMED   15998449
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
```

```
            Email: cmagness@illumigen.com
            Sequenced on 2003.12.24. 840 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 1087   Std Error: 0.00
            Plate: CL000066  row: D  column: 10
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .1087
                      /organism="Macaca mulatta"
                      /mol_type="mRNA"
                      /strain="Indian"
                      /db_xref="taxon:9544"
                      /clone="IBIUW:4699"
                      /sex="male"
                      /dev_stage="newborn infant"
                      /lab_host="E. coli SOLR"
                      /clone_lib="Katze_MMPL2"
                      /note="Organ: placenta; Vector: Uni-ZAP XR; Site_1: EcoR
                      I; Site_2: Xho I; Created from Stratagene ZAP-cDNA
                      Synthesis kit (catalog #200400)  and ZAP-CDNA Gigapack III
                      Gold Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match             43.0%;  Score 563.8;  DB 7;  Length 1087;
  Best Local Similarity   81.7%;  Pred. No. 2.9e-142;
  Matches  678;  Conservative    0;  Mismatches  143;  Indels    9;  Gaps    2;

Qy        436 TACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTAT 492
              | | |||| ||   ||  ||| ||||||| |   ||||  | |||| ||  |||||   |
Db          1 TTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGACAGCATGTT 60

Qy        493 TTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCG 552
              |  ||||||||||||||||||||||||||||||  |  |||||||| | ||||||||||||
Db         61 CTGCTTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCG 120

Qy        553 AGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTAT 612
              |||||||||||||| | |||| ||||| || |||||||||| ||||||||||||||||||
Db        121 AGGGGTGTGTTCCTACTTAGCGTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTAT 180

Qy        613 CCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAA 672
              ||||||||  | |  |||||||||||||||||||      |||    |||||||||||| |
Db        181 TCCCCCATATTATAGCTACCTAAAGAAGAGAGCAAGTGATTATGAGATTATTACAGGTCA 240

Qy        673 TGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCA 732
              |||||||||||||||||||||||| | |||||||||  |||||||||||| ||||| ||
Db        241 TGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCA 300

Qy        733 GGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAG 792
              |||||||||||||||||  |||||||||||||||||||||  |||| ||   ||||||
Db        301 AGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAA 360

Qy        793 TTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTC 852
              | ||   ||||||||  |||||||||||||||||||||||||||||||||||  ||| |
Db        361 TGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGC 420

Qy        853 TGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTC 912
              |||||||||||| |  || |||||||  ||| |    |||||||||||||||||||| ||
Db        421 TGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTC 480

Qy        913 AACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGG 972
              | | || |||||||||||||||||||||||||| |||| |||||||| |||| |||||||||
Db        481 AGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCAAGAGGTTACCAAATGTGG 540

Qy        973 TGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCA 1032
              |||||||||  |||||  || | |||| ||||||  ||||| ||||||||||||||||||||
Db        541 TGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCA 600

Qy       1033 GTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCG 1092
              |||||||||||||||  ||   ||    || ||  | ||  |||||||||  ||||| |||
Db        601 GTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAGCTTCG 660

Qy       1093 TTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTT 1146
              |||||||||||||||||||||      |||||||||||| ||| ||||||| | ||||||
Db         661 TTTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTT 720

Qy       1147 AGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCT 1206
              |||||| |||| |||||||||| |||||||||||||||| ||||||||||||||  ||
Db         721 GGACCTTCTTGGATTGGAAAAACTCGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCT 780

Qy       1207 CTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              ||||||  ||  ||||| |||||| ||| ||||  ||  ||    |||
Db         781 TCTGTGCAATTTAGATGTAATATTGTAAAATGCTTTTCANATGTGAGTTT 830


RESULT 33
BF101875
LOCUS       BF101875                911 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  601753033F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3980457 5',
            mRNA sequence.
ACCESSION   BF101875
VERSION     BF101875.1  GI:10884401
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 911)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9176  row: b  column: 10
            High quality sequence stop: 678.
FEATURES              Location/Qualifiers
     source           1. .911
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="FVB/N"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:3980457"
                      /tissue_type="tumor, biopsy sample"
                      /dev_stage="3 months, virgin"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Mam1"
                      /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Library constructed by Life Technologies. Investigator
                      providing samples: Gilbert Smith, NIH"
ORIGIN

  Query Match             42.7%;  Score 560.8;  DB 2;  Length 911;
  Best Local Similarity   98.0%;  Pred. No. 1.8e-141;
  Matches  631;  Conservative    0;  Mismatches    7;  Indels    6;  Gaps    6;

Qy        662 ATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAA 721
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAA 60

Qy        722 TTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGC 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGC 120

Qy        782 GTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACC 841
              |||||||||||| |||||||||||||| ||||||||||||||||||| ||||||||||||
```

```
Db        121 GTAAATGTGAG-TTGGGTGGAAATGAA-TTTGCCTTATGACATCCCA-TTGGAGAGCACC 177

Qy        842 AGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAG 901

Db        178 AGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAG 237

Qy        902 GCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTT 961

Db        238 GCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTT 297

Qy        962 ATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCT 1021

Db        298 ATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCT 357

Qy       1022 AAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCT 1081

Db        358 AAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCT 417

Qy       1082 TATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAA 1141

Db        418 TATAAGCATCG-TTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAA 476

Qy       1142 AGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201

Db        477 AGTTTAGACCTTGTCGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 536

Qy       1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCT-TTACAC 1260

Db        537 GGGCTCTTGTGCACCTTGAATGTAGTATCGTGAAAAGCTTCCACTTGGCAGCTGTTACAC 596

Qy       1261 GTTTGTTAG-CACTAGTTCTGAATTTGTGTAGGTCTCAACCTTT 1303

Db        597 GTTTGTTAGACACTAGTTCTGAATTTGTGTAGGTCTCAACCCTT 640
```

```
RESULT 34
BM468826
LOCUS       BM468826                 972 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6445782 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5587154
            5', mRNA sequence.
ACCESSION   BM468826
VERSION     BM468826.1  GI:18517868
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 972)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12356  row: d  column: 03
            High quality sequence stop: 707.
FEATURES             Location/Qualifiers
     source          1. .972
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5587154"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_92"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN
```

```
  Query Match           42.7%;  Score 560.8;  DB 3;  Length 972;
  Best Local Similarity   77.5%;  Pred. No. 1.8e-141;
  Matches  747;  Conservative    0;  Mismatches  202;  Indels   15;  Gaps    5;

Qy        264 GATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAG 323

Db          5 GCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGG 64

Qy        324 CCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCA 383

Db         65 CTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAA 124

Qy        384 CGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCA 443

Db        125 CCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCA 184

Qy        444 TTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCA 500

Db        185 CTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCA 244

Qy        501 TTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGT 560

Db        245 TTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGT 304

Qy        561 GTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCAT 620

Db        305 GTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCAT 364

Qy        621 ACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAG 680

Db        365 ATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAG 424

Qy        681 GATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTA 740

Db        425 GATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTA 484

Qy        741 TTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTG 800

Db        485 TTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAACGTGTCAG 544

Qy        801 GAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAAC 860

Db        545 GAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAAC 604

Qy        861 GAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGG 920

Db        605 GAATGAATCAGTTAAAAATGGTTTTGAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAG 664

Qy        921 TTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTAC 980

Db        665 TTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTAC 724

Qy        981 CTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACAT 1040

Db        725 CCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACAT 784

Qy       1041 GGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT-AAGCATCGTTTTGAT 1099

Db        785 GGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAAACTTCGTTTTGAT 844

Qy       1100 CGAATATTTTTCAGAGCAG-------AAGAGGGGCACCTTATTCCTCAAAGTTTAG--AC 1150

Db        845 CGAATATTTTTCAGAGCAGCAGCAGAAAGAGGGACACATTATTTCCCGGAATTTGGGACC 904

Qy       1151 CTTGTTGGGTTGGAAAAACTGGACTGTGG--TAGATTTCCGAGTGATCACTGGGGGCTCT 1208

Db        905 CTTCTGGATTAGAAAAAACTGGACTGTGGGTAGAATTTCCTAAGGATCACTGGGGGTCTT 964
```

```
Qy       1209 TGTG 1212
              |||
Db        965 TCTG 968


RESULT 35
BM926092
LOCUS       BM926092                1046 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6649780 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764440
            5', mRNA sequence.
ACCESSION   BM926092
VERSION     BM926092.1  GI:19376471
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1046)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12817  row: o  column: 01
            High quality sequence start: 4
            High quality sequence stop: 632.
FEATURES             Location/Qualifiers
     source          1. .1046
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5764440"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_114"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
ORIGIN

  Query Match             42.6%;  Score 559;  DB 3;  Length 1046;
  Best Local Similarity   77.7%;  Pred. No. 5.8e-141;
  Matches  729;  Conservative    0;  Mismatches  196;  Indels   13;  Gaps    4;

Qy        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376
              ||||  || |||| | ||||||||||||  || | | | ||    | | |   ||||
Db         36 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 95

Qy        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436
              |||   ||   ||   || || || |||||||||| ||||||| || ||  ||| ||||
Db         96 CCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 155

Qy        437 ACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATT 493
              | |||| ||   ||  ||| ||||||| |   ||||  | |||| || ||||||   |
Db        156 TCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTC 215

Qy        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553
              ||| ||||||||||||||||||||||||||||  |  ||||||| | ||||||||||||
Db        216 TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 275

Qy        554 GGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATC 613
              |||||||||||||| | |||||||||| || ||||||||| ||||||||||||||||||
Db        276 GGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATT 335

Qy        614 CCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAAT 673
              |||||||| | |  |||||||||||||||| ||   | |||    ||||||||||| ||
Db        336 CCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCAT 395

Qy        674 GAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAG 733
              |||||||||||||||||||||||| | |||||||||  |||||||||||| ||||| ||
Db        396 GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAA 455

Qy        734 GAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT 793
              |||||||||||||||||| |||||||||||||||||||||||  |||| ||  |||||| |
Db        456 GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAT 515

Qy        794 TTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCT 853
              ||   |||||||| ||||||||||||||||||||||||||||||||||||  ||| ||
Db        516 GTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCT 575

Qy        854 GCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCA 913
              |||||||||| |  || |||||||  ||| |    |||||| |||||||||||||| |||
Db        576 GCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGGAAGAGGCTCCAGAGTCA 635

Qy        914 ACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGT 973
              | || |||||||||||||||||||||||| |||| |||| ||| |||| || |||||||
Db        636 GCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGT 695

Qy        974 GGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAG 1033
              ||||||||  ||||| || | |||| |||||  | |||| ||||||||||||||||||||
Db        696 GGTTTACCCAACAACATTGTGGATGTCTGGGGAGTTTTGGGCAAACCTAAACATTGCCAG 755

Qy       1034 TATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGT 1093
              ||||||||||||||  ||   ||    || ||  | ||  |||||||| ||| | ||||
Db        756 TATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGT 815

Qy       1094 TTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTA 1147
              ||||||||||||||||||||      |||||||||||| ||| ||||||| | ||| ||
Db        816 TTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGGTTG 875

Qy       1148 GA-CCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCT 1206
              || |||| |||| |   |||||||||||||| || | | || |       |  |||| |
Db        876 GACCCTTCTTGGATAANAAAAACTGGACTGGGGGAAAATTTCCTAATGGATCCTGGGGGT 935

Qy       1207 CTT---GTGCACCTTGAATGTAGTATTGTGAAAAGCTT 1241
              |||   | |||| || ||  ||  ||||| ||| ||||
Db        936 CTTCTGGGGCACTTTAAAAATAAAATTGTAAAAGGCTT 973


RESULT 36
BX337905
LOCUS       BX337905                1067 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX337905 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI052YN13 5-PRIME, mRNA sequence.
ACCESSION   BX337905
VERSION     BX337905.2  GI:46273926
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1067)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30339657.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
```

end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DI052CG07QP1&c=3474.r.

FEATURES Location/Qualifiers
source 1. .1067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI052YN13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 42.2%; Score 553.2; DB 5; Length 1067;
Best Local Similarity 73.8%; Pred. No. 2.2e-139;
Matches 722; Conservative 6; Mismatches 246; Indels 4; Gaps 2;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    |  |   | || | | || |||    ||||| | |||||
Db         41 GCCTGGWGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 100

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        101 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 160

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||   || |
Db        161 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 220

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    |  | |   |||| |||   ||   ||  || || || |||||||||
Db        221 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 280

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |   |
Db        281 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 340

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||   | ||| ||||||||||||||||||||||||||||||
Db        341 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 400

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||||| | |||||||||||||||||||||||||| | ||||||||||| || ||||
Db        401 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 460

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        461 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 520

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| ||||||||||||||||||||||||||| | |||||||
Db        521 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 580

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||| ||||||||||||||
Db        581 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 640

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
              |||||||  |||| ||  |||||| | ||   |||||||| ||||||||||||||||||
Db        641 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 700

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
              ||||||||||||||||  ||| |||||||||||| |  || |||||||  ||| |
Db        701 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAA 760
```

```
Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947
              ||||||| |||||||||||||| ||| | || |||||||||||||||||||||||| |||
Db        761 AGAAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 820

Qy        948 GAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAAT 1007
              | |||| ||| |||| || ||||||||||||||||  ||||| || | |||| ||||| |
Db        821 GGGATCGAGAGGTTACCAGATGTGGTGGTTTACC-AACAACATTGTGGATGTCTGGGAGT 879

Qy        1008 TTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA 1067
               |||| ||||||||||||||||||||||||||||||:||| ||  ||   :|
Db         880 TTTTKGGCAAACCTAAACATTGCCAGTATACATGGGRTACACAAATGAACTCTHAATCTT 939

Qy        1068 GGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGC 1127
               |||           | | | | |||||||||||:| | ||| : ||    | |||:|||
Db         940 GGASTAACTGCTGCTGTTAACTTCGTTTTGATCRATTTTTTCHGAGCTGCGCAGARGGGG 999

Qy        1128 ACCTTATTCCTCAAAGTT 1145
               || :  | | | | |||
Db        1000 GCCCWTATTCCCGAGGTT 1017
```

RESULT 37
BX338160
LOCUS BX338160 981 bp mRNA linear EST 08-APR-2004
DEFINITION BX338160 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI056YC23 5-PRIME, mRNA sequence.
ACCESSION BX338160
VERSION BX338160.2 GI:46283046
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 981)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 2, 2003 this sequence version replaced gi:30345671.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DI056AB12QP1&c=3474.r.

FEATURES Location/Qualifiers
source 1. .981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI056YC23"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 42.1%; Score 552.6; DB 5; Length 981;
Best Local Similarity 77.7%; Pred. No. 3.2e-139;
Matches 702; Conservative 3; Mismatches 193; Indels 5; Gaps 3;

```
Qy        216 GGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCA 275
              ||   ||   |||||| |||| |||  |||  ||||| | ||| |   | |||||  ||
Db         75 GGTGAAAACAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCG 134
```

```
Qy        276 CGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCT 335
Db        135 CAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCT 194

Qy        336 ACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGT 395
Db        195 ACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGC 254

Qy        396 CCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCA 455
Db        255 CCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCA 314

Qy        456 GTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATA 512
Db        315 GCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATA 374

Qy        513 TTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAG 572
Db        375 TTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAG 434

Qy        573 CTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACC 632
Db        435 CTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACC 494

Qy        633 TAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAG 692
Db        495 TAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAG 554

Qy        693 CTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAA 752
Db        555 CTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAA 614

Qy        753 ATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTT 812
Db        615 GTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTT 674

Qy        813 GCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAAT 872
Db        675 GCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGT 734

Qy        873 TAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAG 932
Db        735 TAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAG 794

Qy        933 GAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTT 992
Db        795 GAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTACCCAACAACATTG 854

Qy        993 TTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAG 1052
Db        855 TGGATGTCTGGGAGTTTTTGGGC-AACCTAAACATTGCCAGTATACATGGGATACACAAA 913

Qy       1053 CAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCA 1112
Db        914 TGAACTCTAATCT-TGGAATAMTGCTGCTTGTAAACTTCGTTTTGATCGAATATTTTTCA 972

Qy       1113 GAG 1115
Db        973 VAS 975
```

```
RESULT 38
BX358707
LOCUS       BX358707                942 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX358707 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI042YL19 5-PRIME, mRNA sequence.
ACCESSION   BX358707
VERSION     BX358707.2  GI:46306560
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 942)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 5, 2003 this sequence version replaced gi:30378261.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI042CF10QP1&c=3474.r.
FEATURES             Location/Qualifiers
     source          1. .942
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI042YL19"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

Query Match 42.0%; Score 551.4; DB 5; Length 942;
Best Local Similarity 78.7%; Pred. No. 6.7e-139;
Matches 684; Conservative 2; Mismatches 174; Indels 9; Gaps 2;

```
Qy        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376
Db         32 CAGAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGA 91

Qy        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436
Db         92 CCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 151

Qy        437 ACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATT 493
Db        152 TCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTC 211

Qy        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553
Db        212 TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGA 271

Qy        554 GGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATC 613
Db        272 GGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATT 331

Qy        614 CCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAAT 673
Db        332 CCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCAT 391

Qy        674 GAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAG 733
Db        392 GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAA 451

Qy        734 GAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT 793
Db        452 GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAAT 511

Qy        794 TTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCT 853
Db        512 GTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCT 571

Qy        854 GCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCA 913
```

```
                ||||||||||| |  || |||||||  ||| |    ||||||||||||||||||||| |||
Db        572 GCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCA 631

Qy        914 ACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGT 973
                | || ||||||||||||||||||||||||| |||| |||| ||| |||| || |||||||
Db        632 GCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGT 691

Qy        974 GGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAG 1033
                ||||||||  ||||| || | |||| |||||| ||||| ||||||||||||||||||||
Db        692 GGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAACATTGCCAG 751

Qy       1034 TATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGT 1093
                |||||||||||||||  ||    ||    || ||  | ||  ||||||||| ||| | ||||
Db        752 TATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGTAAACTTCGT 811

Qy       1094 TTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTA 1147
                |||||||||||||||||||      |||||||||||| ||| ||||||| | ||||||
Db        812 TTTGATCGAATATTTTTCAGAGCAGCAGCAGAAGAGGGACACATTATTCCCCGAAGTTTG 871

Qy       1148 GACCTTGTTGGGTTGGAAAAACTGGACTG 1176
                |||:|  | | | | ||||: |  ||
Db        872 GACHTCTTGGAWTAGAAAAAWGGGMTGTG 900
```

RESULT 39
BY737520
LOCUS BY737520 645 bp mRNA linear EST 17-DEC-2002
DEFINITION BY737520 RIKEN full-length enriched, 11 days pregnant adult female placenta Mus musculus cDNA clone I530029B02 5', mRNA sequence.
ACCESSION BY737520
VERSION BY737520.1 GI:27150647
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 645)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .645
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I530029B02"
/sex="female"
/tissue_type="placenta"
/dev_stage="11 days pregnant adult"
/clone_lib="RIKEN full-length enriched, 11 days pregnant adult female placenta"
ORIGIN

Query Match 42.0%; Score 551; DB 5; Length 645;
Best Local Similarity 93.8%; Pred. No. 7.7e-139;
Matches 594; Conservative 0; Mismatches 37; Indels 2; Gaps 2;

```
Qy        271 CCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAG 330
                ||||  |  ||      ||| ||||| |      | ||          ||||||||||||
Db         15 CCCCCGGGAGGCGGGGCGCGCCCTGCAGCCTTGTGCCTTTGCTTTTCAGAAAGCCCTGAG 74

Qy        331 CGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTT 390
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTT 134

Qy        391 CAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGA 450
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGA 194

Qy        451 AGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAA 510
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 AGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAA 254

Qy        511 TATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCT 570
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 TATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCT 314

Qy        571 AGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTA 630
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 AGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTA 374

Qy        631 CCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCAC 690
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 CCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCAC 434

Qy        691 AGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCC 750
                |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        435 AGCTATACTATTGAAGAAAGGAAGAGTGAAATTT-AAAGTCAGGAGATTATTCCTTTTCC 493

Qy        751 AAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATT 810
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 -AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATT 552

Qy        811 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACA 870
                |||||||||||||||||| |||||||||||||||||||||||||||||||||| |
Db        553 TTGCCTTATGACATCCCATNTGGAGAGCACCAGAGAACATTCTGCGGAACGAAAATAGAC 612

Qy        871 ATTAAAAACTGTTCTTGGAAAAATGCAAGAGGC 903
                |||||||||||| |||||||||||||  ||||||
Db        613 ATTAAAAACTGGTCTTGGAAAAATGNCAGAGGC 645


RESULT 40
BF471995
LOCUS       BF471995                637 bp    mRNA    linear   EST 04-DEC-2000
DEFINITION  UI-M-BH3-awo-f-01-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
            UI-M-BH3-awo-f-01-0-UI 5', mRNA sequence.
ACCESSION   BF471995
VERSION     BF471995.1  GI:11541178
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 637)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
            should be noted that Bento Soares is generating a small number of
            additional specialized non-redundant arrays of BMAP cDNAs whose
            availability will be considered under appropriate and limited
            collaborative arrangements
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .637
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-BH3-awo-f-01-0-UI"
                     /dev_stage="27-32 days"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NIH_BMAP_M_S4"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_M_S4 library is a subtracted library of a series,
                     ultimately derived from a mixture of individually tagged
                     normalized libraries from ten regions of the mouse brain
                     (cerebellum, brain stems, olfactory bulbs, hypothalamus,
                     cortex, amygdala, basal ganglia, pineal gland, striatum,
                     hipoccampus) after a series of subtractions to reduce the
                     representation of cDNAs from which ESTs had already been
                     generated. The following serially subtracted libraries
                     were generated in this process: NIH_BMAP_M_S4,
                     NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
                     NIH_BMAP_M_S2, NIH_BMAP_M_S1.  The subtracted library
                     (NIH_BMAP_M_S4) was constructed as follows:  PCRamplified
                     cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
                     NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
                     was used as  a driver in a hybridization with a pool of
                     the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
                     libraries in the form of  single-stranded circles.  The
                     remaining single-stranded circles  (subtracted library)
                     was purified by hydroxyapatite column chromatography,
                     converted to double-stranded circles and electroporated
                     into DH10B  bacteria (LifeTechnologies) to generate the
                     NIH_BMAP_M_S4 library.  This procedure has been previously
                     described (Bonaldo, Lennon and Soares, Genome Research
                     6:791-806, 1996)"
ORIGIN

  Query Match             41.5%;  Score 545;  DB 2;  Length 637;
  Best Local Similarity   99.8%;  Pred. No. 3.3e-137;
  Matches  556;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy         756 CCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCC 815
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCC 60

Qy         816 TTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAA 875
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAA 120

Qy         876 AAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAG 935
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAG 180

Qy         936 ATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTG 995
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTG 240

Qy         996 ATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAA 1055
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAA 300

Qy        1056 ATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAG 1115
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAG 360

Qy        1116 CAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACT 1175
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACT 420

Qy        1176 GTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAA 1235
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAA 480

Qy        1236 AAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCT 1295
               |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AAGC-TCCCACTTGCAGCTTTACACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCT 539

Qy        1296 CAACCTTTCAGGACATC 1312
               |||||||||||||||||
Db         540 CAACCTTTCAGGACATC 556


RESULT 41
BG800090
LOCUS       BG800090                600 bp    mRNA    linear   EST 20-DEC-2001
DEFINITION  2112-54 Mouse E14.5 retina lambda ZAP II Library Mus musculus cDNA,
            mRNA sequence.
ACCESSION   BG800090
VERSION     BG800090.1  GI:17946942
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Mu,X., Zhao,S., Pershad,R., Hsieh,T.-F., Scarpa,A., Wang,S.W.,
            White,R.A., Beremand,P.D., Thomas,T.L., Gan,L. and Klein,W.H.
  TITLE     Gene expression in the developing mouse retina by EST sequencing
            and microarray analysis
  JOURNAL   Nucleic Acids Res. 29 (24), 4983-4993 (2001)
   PUBMED   11812828
COMMENT     Contact: Klein WH
            Department of Biochemistry and Molecular Biology
            University of Texas M.D. Anderson Cancer Center
            Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
            Tel: 713 792 3646
            Fax: 713 790 0329.
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="neural retina"
                     /dev_stage="embryonic day 14.5 post-fertilization"
                     /clone_lib="Mouse E14.5 retina lambda ZAP II Library"
ORIGIN

  Query Match             41.4%;  Score 543.8;  DB 2;  Length 600;
  Best Local Similarity   97.5%;  Pred. No. 7e-137;
  Matches  582;  Conservative    0;  Mismatches   11;  Indels    4;  Gaps    3;

Qy         73 GGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGC-GTCTG 131
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |
Db          4 GCGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCAGTCTG 63

Qy        132 GCAGCAGTTCCGATGC--GGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCC 189
              |||||||||||||||  |  |    || ||||||||||||||||||||||||||||||||
Db         64 GCAGCAGTTCCGATGCAGGCGGAGCCCAGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCC 123

Qy        190 CGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTT 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTT 183

Qy        250 TGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTG 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 TGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTG 243

Qy        310 GCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCC 369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCC 303

Qy        370 GCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGC 429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGC 363

Qy        430 AAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCAC 489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        364 AAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCANCAC 423

Qy        490 TATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGC 549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 TATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGC 483

Qy        550 TCGA-GGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAG 608
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        484 TCGAGGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACANGAAG 543

Qy        609 TTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTA 665
              |||||||||||||||| |||||||||||||| ||||||||||||||||||||||||||
Db        544 TTATCCCCCCATACTGNGCCTACCTAAAGAANAGAGCAGCCAGTTACACAATTATTA 600


RESULT 42
BU179107
```

```
LOCUS       BU179107                883 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7984768 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6174958
            5', mRNA sequence.
ACCESSION   BU179107
VERSION     BU179107.1  GI:22693091
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13548  row: o  column: 23
            High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6174958"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

  Query Match             41.4%;  Score 542.8;  DB 5;  Length 883;
  Best Local Similarity   79.2%;  Pred. No. 1.5e-136;
  Matches  697;  Conservative    0;  Mismatches  172;  Indels   11;  Gaps    4;

Qy        300 AGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACC 359
              |||||||||| ||| | | |  || |||| | ||||||||||  || | | | ||
Db          4 AGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGA 63

Qy        360 AAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCA 419
              |  | |   |||| |||   ||   ||   || || || |||||||||| |||||||
Db         64 GCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCA 123

Qy        420 ACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAG 476
              | || ||  ||| |||| | |||| ||   ||  ||| |||||||| |   ||||  | ||
Db        124 ATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAG 183

Qy        477 AAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATC 536
              || || ||||||   | ||| |||||||||||||||||||||||||| |  |||||
Db        184 AAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATC 243

Qy        537 TGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTAT 596
              || | |||||||||||||||||||||| | |||||||||| || |||||||||| |||
Db        244 TGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATAT 303

Qy        597 TTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACA 656
              |||||||||||||||| |||||||| | |  ||||||||||||||| ||   | |||
Db        304 TTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATG 363

Qy        657 CAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAG 716
              |||||||||||| |||||||||||||||||||||||||| | |||||||| |||||
Db        364 AGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAG 423

Qy        717 TGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGC 776
```

```
Db        424 TGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTT 483

Qy        777 TATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGA 836

Db        484 TATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGA 543

Qy        837 GCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGC 896

Db        544 GCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGC 603

Qy        897 AAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAG 956

Db        604 AAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAG 663

Qy        957 AAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCA 1016

Db        664 AGGTTACCAGATGTGGTGGTTTACCCAACAACATTGTGGATGTCTGGGAGTTTTTGGGCA 723

Qy       1017 AACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTG 1076

Db        724 AACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTG 783

Qy       1077 CTGCTTATAAGCATCG-TTTTGATCGAATATTTTTCAG------AGCAGAAGAGGGGCAC 1129

Db        784 CTGCTTGTAAACTTCGTTTTTGATCGAATATTTTTCCGAGCAGCAGCAGAAGAGGGACAC 843

Qy       1130 CTTATT-CCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAA 1168

Db        844 TTTATTCCCCCGAAGTTTGGACCTTCTTGGATTAAAAAAA 883


RESULT 43
CN641671
LOCUS       CN641671               1090 bp    mRNA    linear   EST 12-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_5159 Katze_MMBR Macaca mulatta cDNA clone IBIUW:5970
            5' similar to Bases 11 to 990 highly similar to human TTRAP
            (Hs.210628), mRNA sequence.
ACCESSION   CN641671
VERSION     CN641671.1  GI:47152681
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 1090)
  AUTHORS   Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
  TITLE     Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
  JOURNAL   Genome Biol. 6 (7), R60 (2005)
   PUBMED   15998449
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2003.12.04. 732 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 1090   Std Error: 0.00
            Plate: CL000043  row: E  column: 07
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .1090
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:5970"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match             41.4%;  Score 542.8;  DB 7;  Length 1090;
  Best Local Similarity   78.6%;  Pred. No. 1.6e-136;
  Matches  662;  Conservative    0;  Mismatches  177;  Indels    3;  Gaps    1;

Qy         283 CCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGA 342

Db           1 CCCTTGAAAACTGGCCGAGAACGACTGGGAGATGGAAAGCCTCTCAAACTCCTACTTCGA 60

Qy         343 GCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGC 402

Db          61 GCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGCCCTGAAACCATCTCTGAGCCCAAGAC 120

Qy         403 CTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATC 462

Db         121 CTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATC 180

Qy         463 TGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGG 519

Db         181 TGAAGATACTCAGCAAGAAAATGACAGCATGTTCTGCTTCATTACCTGGAATATTGATGG 240

Qy         520 ATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTA 579

Db         241 ATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCGTTGTA 300

Qy         580 TAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAA 639

Db         301 CAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCTTATTATAGCTACCTAAAGAA 360

Qy         640 GAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACT 699

Db         361 GAGAGCAAGTGATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAAT 420

Qy         700 ATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAA 759

Db         421 GTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAA 480

Qy         760 AATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTAT 819

Db         481 AATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTAT 540

Qy         820 GACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAAC 879

Db         541 GACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAAT 600

Qy         880 TGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATAC 939

Db         601 GGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATAC 660

Qy         940 AAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGC 999

Db         661 AAATCTAAGGGATCAAGAGGTTACCAAATGTGGTGGTTTACCCAACAACATTGTGGATGT 720

Qy        1000 CTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAA 1059

Db         721 CTGGGAGTTTTTGGGCAAACCTAAACATTGCCAGTATACATGGGATACACAAATGAACTC 780

Qy        1060 CAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGA 1119

Db         781 TAATCTTGGAATAACTGCTGCTTGTAAGCTTCCTTTTGATCGAATATTTTTCAGAGCAGC 840

Qy        1120 AG 1121
```

```
Db        841 AG 842

RESULT 44
CN801806
LOCUS       CN801806                922 bp    mRNA    linear   EST 26-MAY-2004
DEFINITION  ILLUMIGEN_MCQ_37066 Katze_MMBR Macaca mulatta cDNA clone
            IBIUW:15450 5' similar to Bases 18 to 922 highly similar to human
            TTRAP (Hs.210628), mRNA sequence.
ACCESSION   CN801806
VERSION     CN801806.1  GI:47697782
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 922)
  AUTHORS   Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proll,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
  TITLE     Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
  JOURNAL   Genome Biol. 6 (7), R60 (2005)
   PUBMED   15998449
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.05.13. 514 Q20 bases.
            PCR PRimers
            FORWARD: CCCTCACTAAAGGGAACAAAA
            BACKWARD: CACTATAGGGCGAATTGGGTA
            Insert Length: 922   Std Error: 0.00
            Plate: CL000291  row: H  column: 09
            Seq primer: CCCTCACTAAAGGGAACAAAA
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .922
                     /organism="Macaca mulatta"
                     /mol_type="mRNA"
                     /strain="Indian"
                     /db_xref="taxon:9544"
                     /clone="IBIUW:15450"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Katze_MMBR"
                     /note="Organ: brain; Vector: Uni-ZAP XR; Site_1: EcoR I;
                     Site_2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
                     kit (catalog #200400)  and ZAP-CDNA Gigapack III Gold
                     Cloning Kit (Catalog #200450)"
ORIGIN

  Query Match           41.1%;  Score 538.6;  DB 7;  Length 922;
  Best Local Similarity   76.4%;  Pred. No. 2.1e-135;
  Matches  690;  Conservative    0;  Mismatches  204;  Indels    9;  Gaps    2;

Qy        275 ACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCC 334
              | || |   | | |  |||||   ||||| || ||||| | |   |  || |||| | ||
Db          3 ATGAGGCCTCTCCGTTTCCTGGACGAGAAGGAGTGGCACATGACCAGGGCTCTGATCTCC 62

Qy        335 TACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAG 394
              |  |||||||  || | | | ||   | | |  |||| |||  ||  ||  ||  ||
Db         63 TGATTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGCCCTGAAACCATCTCTGAG 122

Qy        395 TCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCC 454
              || || |||||||||| |||||||| || ||  ||| |||| | |||| ||   ||  |
Db        123 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATC 182

Qy        455 AGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAAT 511
              || ||||||| |   ||||  | |||| ||  |||||    |  | ||||||||||||||
Db        183 AGCCCATCTGAAGATACTCAGCAAGAAAATGACAGCATGTTCTGCTTCATTACCTGGAAT 242

Qy        512 ATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTA 571
              |||||||||||| ||  |  |||||||  | |||||||||||||||||||||||||| | ||
Db        243 ATTGATGGATTAAATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTA 302

Qy        572 GCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTAC 631
              || ||||| || |||||||| |||||||||||||||||| |||||||| | |  ||||
Db        303 GCGTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTAC 362

Qy        632 CTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACA 691
              ||||||||||||||      |||    |||||||||||| ||||||||||||||||||||
Db        363 CTAAAGAAGAGAGCAAGTGATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACA 422

Qy        692 GCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCA 751
              |||||| | |||||||||  |||||||||||| ||||| || ||||||||||||||||||
Db        423 GCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCA 482

Qy        752 AATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTT 811
              | |||||||||||||||||||  |||| ||  |||||| | ||   ||||||||  ||
Db        483 AGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTT 542

Qy        812 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAA 871
              |||||||||||||||||||||||||||||||||  ||| ||||||||||||| |  ||
Db        543 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAG 602

Qy        872 TTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCA 931
              |||||||  ||| |   | ||||||||||||||||||| ||| | || ||||||||||||
Db        603 TTAAAAATGGTTTTAAAGAGAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCA 662

Qy        932 GGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTT 991
              |||||||||  || |||| |||||||| |||| |||||||||| |||||||  ||||| ||
Db        663 GGAGATACATATCTAAGGGATCAAGAGGTTACCAAATGTGGTCGTTTACCCAACAACATT 722

Qy        992 TTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAA 1051
              | |||| |||||| |||||  |||||||||||||||  ||||||||||||||||||   |
Db        723 GTGGATGTCTGGGAGTTTTTGAGCAAACCTAAACATTTGCAGTATACATGGGATACATCA 782

Qy       1052 GCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTTC 1111
                ||    || ||    || |||||||||| ||||| ||||||||||||||||||||||
Db        783 ATGAACTCTAATCTGGAAATACCTGCTGCTTGTAAGCTTCGTTTTGATCGAATATTTTTC 842

Qy       1112 AGAG------CAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 1165
              | ||      |||||   ||   | ||  ||| | ||| ||  ||  || |||| ||||||
Db        843 AAAGCACCAACAGAACGAGGAACCATTTATCCCCGAAGCTTGTACTTTCTTGGATTGGAA 902

Qy       1166 AAA 1168
              |||
Db        903 AAA 905

RESULT 45
CB545297
LOCUS       CB545297                639 bp    mRNA    linear   EST 01-APR-2003
DEFINITION  AMGNNUC:NRDG1-00108-F11-A nrdg1 (10855) Rattus norvegicus cDNA
            clone nrdg1-00108-f11 5', mRNA sequence.
ACCESSION   CB545297
VERSION     CB545297.1  GI:29429238
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 639)
  AUTHORS   Amgen EST Program.
  TITLE     Amgen Rat EST Program
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
```

```
           Tel: 805 447-4881
           Plate: 00108  row: f  column: 11.
FEATURES             Location/Qualifiers
     source          1. .639
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="nrdg1-00108-f11"
                     /tissue_type="Dorsal Root Ganglia"
                     /clone_lib="nrdg1 (10855)"
                     /note="Vector: pSPORT1; Site_1: Sal1; Site_2: Not1; rat
                     dorsal root ganglia"
ORIGIN

  Query Match             40.0%;  Score 524.2;  DB 6;  Length 639;
  Best Local Similarity   89.3%;  Pred. No. 1.6e-131;
  Matches  565;  Conservative    0;  Mismatches   68;  Indels    0;  Gaps    0;

Qy        175 GGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCT 234
              | ||||| |||||||||| ||||||| || |||||||||| | ||||||||||||||||||
Db          7 GCCGGCGGCGGCGCCCGCAGCAGCAGAGACGGAGGAGGATCAGGTGAAGAGGCGGCGGCT 66

Qy        235 TCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCT 294
              ||||| |||||| |||||||||||  | | |||||| | ||| |||   |||   ||||
Db         67 TCAGTCCCTGGGATTTGCGTTGGTGACGAGCTGCGACACTACGGTGGCTTCCACTTTCCT 126

Qy        295 GCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAA 354
              | ||||||| |||||||||| | |||||  ||||||||| |||||||| |||||||||
Db        127 GTCGGAGAACAACTGGCAGACGAAAAAAGCGTTGAGCGCCTTCTTCGAGCAGCCAGAGAA 186

Qy        355 CGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCT 414
              ||||| || | |||| | ||||||||| || |||| |||||||||| ||||||||||
Db        187 CGACCTAGCGCGGCCTCACCAGCCTCCGACATCCTCCAAGTCCGAGGACTATGTTGATCT 246

Qy        415 AACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCT 474
              |||||| |||||||||||||||| |||||||||| |||||||||||||||||||||||
Db        247 AACCAATGAGGATGCAAATGATACCACCATTTTAGAAACCAGTCCATCTGGAACTCCTCT 306

Qy        475 AGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAA 534
              |||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||
Db        307 AGAAGATAGCAGCACTATCTCTTTCATTAGCTGGAATATTGATGGATTAGATGGATGCAA 366

Qy        535 TCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGT 594
              ||| || |||||||||||||||||||||||||||||| |||||||| ||||| || |||||
Db        367 TCTCCCAGAGAGGGCTCGAGGGGTGTGTTCCTGCCTTGCTTTGTACAGTCCGGACGTGGT 426

Qy        595 ATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTA 654
              |||||||||||||| || ||| | |||||||||||||| ||||||||||  || |||
Db        427 ATTTCTACAGGAAGTGATTCCCTCCTACTGTGCCTACCTAAGGAAGAGAGCACGCACTTA 486

Qy        655 CACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAG 714
              ||  ||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        487 CAACATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATATTATTGAAGAAAGGAAG 546

Qy        715 AGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCT 774
              |||||||||||| |||| || ||||||||||||||||||||||||||||||||||||||
Db        547 AGTGAAATTTAAAGGTCAAGAAATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCT 606

Qy        775 GCTATGCGTAAATGTGAGTTTGGGTGGAAATGA 807
              || || |||||||||||||||||||||||||||
Db        607 TCTGTGTGTAAATGTGAGTTTGGGTGGAAATGA 639


RESULT 46
AA824213
LOCUS       AA824213                 531 bp    mRNA    linear   EST 17-FEB-1998
DEFINITION  vy20h08.r1 Stratagene mouse macrophage (#937306) Mus musculus cDNA
            clone IMAGE:1296063 5', mRNA sequence.
ACCESSION   AA824213
VERSION     AA824213.1  GI:2894081
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:677111
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 464.
FEATURES             Location/Qualifiers
     source          1. .531
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1296063"
                     /tissue_type="macrophage"
                     /dev_stage="WEHI-3 cell line"
                     /lab_host="SOLR (kanamycin resistant)"
                     /clone_lib="Stratagene mouse macrophage (#937306)"
                     /note="Organ: blood; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. WEHI-3 cell line. Average insert size: 1.5 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             39.6%;  Score 520;  DB 1;  Length 531;
  Best Local Similarity   99.8%;  Pred. No. 2.1e-130;
  Matches  531;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        751 AAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATT 810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATT 60

Qy        811 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACA 870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACA 120

Qy        871 ATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGC 930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGC 180

Qy        931 AGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGT 990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGT 240

Qy        991 TTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAA 1050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAA 300

Qy       1051 AGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTT 1110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTTTT 360

Qy       1111 CAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACT 1170
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGAGCAGAAGAGGGGCA-CTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAAAAACT 419

Qy       1171 GGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATT 1230
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 GGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATT 479

Qy       1231 GTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACTAGTTCTGAA 1282
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACTAGTTCTGAA 531


RESULT 47
CC248786
LOCUS       CC248786                517 bp    mRNA    linear   GSS 13-MAY-2003
DEFINITION  XK529 BayGenomics Gene Trap Library pGT0Lxf Mus musculus cDNA, mRNA
            sequence.
ACCESSION   CC248786
VERSION     CC248786.1  GI:30585518
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 517)
  AUTHORS   BayGenomics.
  TITLE     http://baygenomics.ucsf.edu/
  JOURNAL   Unpublished (2001)
COMMENT     Contact: BayGenomics
            Bay Area Functional Genomics Consortium (BayGenomics)
            Email: info@baygenomics.ucsf.edu
            Sequence tag generated by 5' RACE of total RNA from gene trap ES
            cell line. ES cell lines harboring insertion mutation of target
            gene are available upon request from BayGenomics. Annotation
            information available from http://baygenomics.ucsf.edu/cgi-bin/
            BaySearch.py?OPTION=EXACT&TYPE=CELL_LINE&KEY=XK529
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .517
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129 ola"
                     /db_xref="taxon:10090"
                     /sex="Male"
                     /cell_type="Embryonic stem cell"
                     /clone_lib="BayGenomics Gene Trap Library pGT0Lxf"
                     /note="Vector: pGT0Lxf"
ORIGIN

  Query Match             39.4%;  Score 517;  DB 9;  Length 517;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-129;
  Matches  517;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        442 CATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCAT 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCAT 60

Qy        502 TACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTG 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTG 120

Qy        562 TTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATA 180

Qy        622 CTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGG 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGG 240

Qy        682 ATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTAT 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTAT 300

Qy        742 TCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGG 801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGG 360

Qy        802 AAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACG 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACG 420

Qy        862 AATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGT 921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGT 480

Qy        922 TATATTTGCAGGAGATACAAATTTAAGAGATCAAGAA 958
              |||||||||||||||||||||||||||||||||||||
Db        481 TATATTTGCAGGAGATACAAATTTAAGAGATCAAGAA 517


RESULT 48
BI150788
LOCUS       BI150788                961 bp    mRNA    linear   EST 05-JUL-2001
DEFINITION  602914879F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5065531 5',
            mRNA sequence.
ACCESSION   BI150788
VERSION     BI150788.1  GI:14610789
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 961)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11176  row: m  column: 20
            High quality sequence stop: 785.
FEATURES             Location/Qualifiers
     source          1. .961
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="Czech II"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5065531"
                     /tissue_type="spontaneous tumor, metastatic to mammary.
                     Stem cell origin."
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu29"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
ORIGIN

  Query Match             39.3%;  Score 516.2;  DB 2;  Length 961;
  Best Local Similarity   83.1%;  Pred. No. 2.8e-129;
  Matches  684;  Conservative    0;  Mismatches    3;  Indels  136;  Gaps    3;

Qy        506 TGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCC 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCC 60

Qy        566 TGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGT 625
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCCTAGCTTTGTACAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGT 120

Qy        626 GCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATAT 685
              |||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 GCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAG---------------- 163

Qy        686 TTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCT 745

Db        164 ------------------------------------------------------------ 163

Qy        746 TTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAAT 805
                                                  ||||||||||||||||||
Db        164 ------------------------------------------GTGAGTTTGGGTGGAAAT 181

Qy        806 GAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATA 865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATA 241

Qy        866 AGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATA 925
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 AGTCAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATA 301

Qy        926 TTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGAC 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGAC 361

Qy        986 AACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGAT 1045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 AACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGAT 421

Qy       1046 ACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATA 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATA 481

Qy       1106 TTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTGGAA 1165
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        482 TTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGATCTTGTTGGGTTGGAA 541

Qy       1166 AAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTA 1225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 AAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTGAATGTA 601

Qy       1226 GTATTGTGAAAAGCTTCC----------------CACTTGCAGCTTTACACGTTTGTTAG 1269
              ||||||||||||||||||                ||||||||||||||||||||||||||
Db        602 GTATTGTGAAAAGCTTCCCACTTCCAGCTTTACACACTTGCAGCTTTACACGTTTGTTAG 661

Qy       1270 CACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              |||||||||||||| |||||||||||||||||||||||||||||
Db        662 CACTAGTTCTGAA-TTGTGTAGGTCTCAACCTTTCAGGACATC 703
```

RESULT 49
CZ905189
LOCUS CZ905189 682 bp mRNA linear GSS 29-JUL-2005
DEFINITION NAISTrap_26v1023 NAIST Gene Trap Library v6.4-UPAT-egfp3ic Mus musculus cDNA, mRNA sequence.
ACCESSION CZ905189
VERSION CZ905189.1 GI:71493507
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 682)
AUTHORS Shigeoka,T., Kawaichi,M. and Ishida,Y.
TITLE Suppression of nonsense-mediated mRNA decay permits unbiased gene trapping in mouse embryonic stem cells
JOURNAL Nucleic Acids Res. 33 (2), e20 (2005)
PUBMED 15687378
COMMENT Contact: Ishida Y
Division of Gene Function in Animals
Nara Institute of Science and Technology (NAIST)
8916-5 Takayama-cho, Ikoma-shi, Nara 630-0192, Japan
Tel: 81 743 72 5531
Fax: 81 743 72 5539
Email: ishiday@bs.naist.jp

Sequence tag generated by 3' RACE using total RNA from gene-trapped ES-cell line.
Class: Gene Trap.
FEATURES Location/Qualifiers
source 1. .682
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x 129/SvJae"
/db_xref="taxon:10090"
/sex="male"
/cell_type="Embryonic Stem Cell"
/clone_lib="NAIST Gene Trap Library v6.4-UPAT-egfp3ic"
/note="Vector: pUPAT-egfp3ic"
ORIGIN

Query Match 39.0%; Score 512.2; DB 10; Length 682;
Best Local Similarity 82.2%; Pred. No. 3.1e-128;
Matches 555; Conservative 0; Mismatches 119; Indels 1; Gaps 1;

```
Qy        434 GATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATT 493
              ||||| ||||||  | | | ||| || ||||||||||||||||||||||||| |||||||
Db          1 GATACNACCATTCNACANGGCAGNCCNTCTGGAACTCCTCTAGAAGATAGCANCACTATT 60

Qy        494 TCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553
              |||||||||||| |||||||||||||||||| ||||||||| ||||||||| ||||||||
Db         61 TCTTTCATTACCGGGAATATTGATGGATTACATGGATGCAANCTGCCCGAGNGGGCTCGA 120

Qy        554 GGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATC 613
              |||||||||||||||||| |||||||| | ||| ||||||||||||||| | | | ||
Db        121 GGGGTGTGTTCCTGCCTANCTTTGTATNGNCCATATGTGGTATTTCTACNCGCAACTNTC 180

Qy        614 CCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAAT 673
              ||||||| ||| | ||||||||| ||||| || |||| |||||| |||||||||| |||
Db        181 CCCCCATNCTGNGNCTACCTAAACAAGAGNGCCGCCANTTACACNATTATTACAGGNAAT 240

Qy        674 GAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAG 733
              ||| |||||| |||| ||  |||||||||||||| |||||||||||||| |||||||
Db        241 GAACAAGGATTTTTCNCANTTATACTATTGAAGAANGGAAGAGTGAAATTNAAAAGTCCT 300

Qy        734 GAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT 793
              || |||||||||||||||| |||||||| | ||| ||||||| |||||  ||||||||||
Db        301 GANATTATTCCTTTTCCAANTACCAAAANGNTGANAAACCTGNTATGCNNAAATGTGAGT 360

Qy        794 TTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCT 853
              | |||||||||||||||| |||| ||||| |||| | ||| ||| || |||||| ||||
Db        361 TAGGGTGGAAATGAATTTTNCCTTNTGACANCCCANTGGGANAGCNCCNGAGAACCTTCT 420

Qy        854 GCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCA 913
              |||| |||||||| ||  || || |||||||||||||  | |||||| || ||||||||
Db        421 GCGGNACGAATAACACGNTTTAACACTGTTCTTGGAANCANGCAAGACGCNCCAGATTCT 480

Qy        914 ACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGT 973
              ||||||| ||||||| ||||| ||||||||| |||||||| | || || || || | ||
Db        481 ACCACGGGTATATTTNCAGGANATACAAATTNAAGAGATCNANAANTTNTCNAANGCGGN 540

Qy        974 GGTTTACCTGACAACGT-TTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCA 1032
              ||    | |||| |||| |    |||| |||| |||| || ||    | | | ||||| |
Db        541 GGAAANCNTGACNACGTGTGAAGATGTNTGGGNATTTATANGCCNCCGTGATCATTGGCG 600

Qy       1033 GTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCG 1092
              || |||| |||||||||||||||| |  || | || |||||||||||| | |||||||
Db        601 GTNTACAGGGGATACGAAAGCAAATANCCCCCNCGGGTTCCCTGCTGCTTNTTAGCATCG 660

Qy       1093 TTTTGATCGAATATT 1107
              |||| ||| ||  ||
Db        661 TTTTNATCCAAATTT 675
```

RESULT 50
AA208842
LOCUS AA208842 544 bp mRNA linear EST 18-FEB-1997
DEFINITION mw73d02.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:676323 5', mRNA sequence.

```
ACCESSION   AA208842
VERSION     AA208842.1  GI:1806779
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 544)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:416027
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 491.
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:676323"
                     /tissue_type="Liver"
                     /lab_host="DH10B"
                     /clone_lib="Soares mouse NML"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATCTTTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
ORIGIN

  Query Match             39.0%;  Score 511.8;  DB 1;  Length 544;
  Best Local Similarity   98.9%;  Pred. No. 3.7e-128;
  Matches  537;  Conservative    0;  Mismatches    2;  Indels    4;  Gaps    2;

Qy        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGC--A 63

Qy        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 123

Qy        437 ACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCT 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 ACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCT 183

Qy        497 TTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGG 556
              |||||||||||||||||||||||||||||||||||||||||||  |  |||||||
Db        184 TTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGACG--TCGAGGG 241

Qy        557 GTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCC 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCC 301

Qy        617 CCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAA 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAA 361

Qy        677 GAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAG 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAG 421

Qy        737 ATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTG 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTG 481

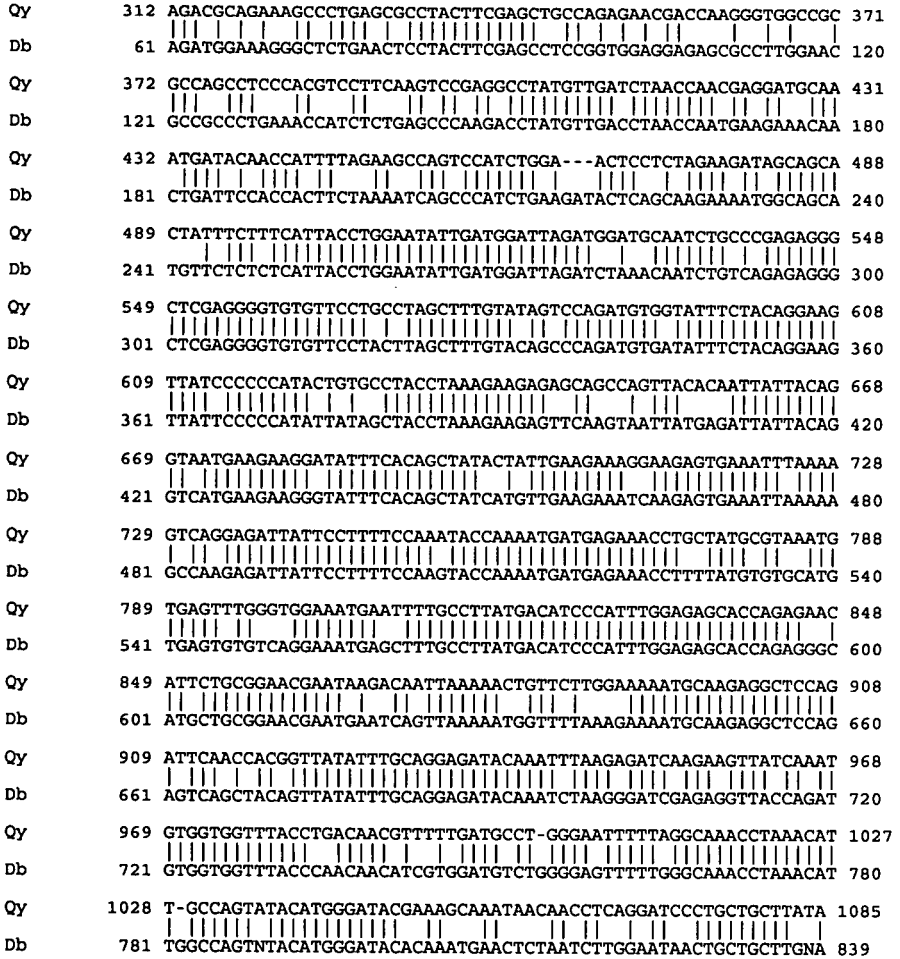
Qy        797 GGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCG 856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCG 541

Qy        857 GAA 859
              |||
Db        542 GAA 544
```

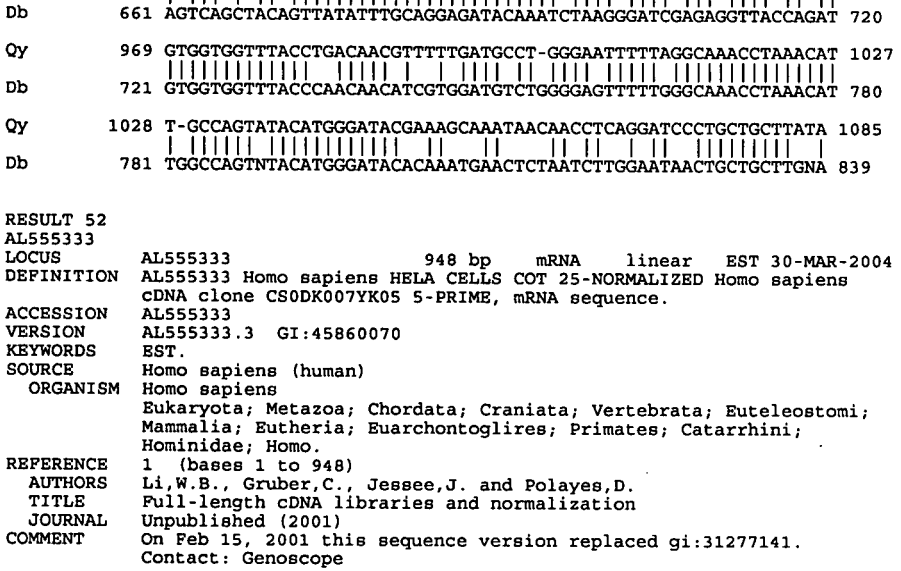
```
RESULT 51
CR765451
LOCUS       CR765451                846 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp469E2434_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469E2434 5', mRNA sequence.
ACCESSION   CR765451
VERSION     CR765451.1  GI:52605526
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 846)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,
            Deutschenbaur,S., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix
            (Martinsried/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp469E2434) is available at
            the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
            Berlin, Germany. Please contact RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469E2434
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .846
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp469E2434"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="469 (synonym: pkid1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

  Query Match             38.6%;  Score 505.8;  DB 7;  Length 846;
  Best Local Similarity   77.5%;  Pred. No. 1.8e-126;
  Matches  650;  Conservative    0;  Mismatches  184;  Indels    5;  Gaps    3;

Qy        252 CGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGC 311
              | | ||| | | | |||||  || |  |||  |   || |||||   ||||||||||||
Db          1 CCTCGGTCGCGAGCTGCGATACCGCTGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGG 60
```



RESULT 52
AL555333
LOCUS AL555333 948 bp mRNA linear EST 30-MAR-2004
DEFINITION AL555333 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens cDNA clone CS0DK007YK05 5-PRIME, mRNA sequence.
ACCESSION AL555333
VERSION AL555333.3 GI:45860070
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 948)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31277141.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3474.r
For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CS0DK007AF03QP1&c=3474.r.
FEATURES Location/Qualifiers
source 1. .948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK007YK05"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN



```
Qy        873 TAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAG 932
              ||||||  |||  |    ||||||||||||||||||| ||| | || ||||||||||||||
Db        741 TAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAG 800

Qy        933 GAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTT 992
              ||||||||:|| ||:| |||| ||| |||| || ||||||||:|||||  ||||: ||
Db        801 GAGATACAWATCTARGGGATCGAGAGGTTACCAGATGTGGTGGKTTACC-AACAAMATTG 859

Qy        993 TTGATGCCTGGGAATTTTTAGGCAAACCTAA 1023
              |:|||: ||||||  |||||  |||||:| |||
Db        860 TKGATKTCTGGGAGTTTTTGGGCAAMCTTAA 890


RESULT 53
CR767800
LOCUS       CR767800                749 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp469C058_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469C058 5', mRNA sequence.
ACCESSION   CR767800
VERSION     CR767800.1  GI:52610302
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ottenwaelder,B., Obermaier,B.,
            Deutschenbaur,S., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert. Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix
            (Martinsried/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp469C058) is available at
            the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
            Berlin, Germany. Please contact RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469C058
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Pongo pygmaeus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9600"
                     /clone="DKFZp469C058"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="469 (synonym: pkid1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

  Query Match           38.3%;  Score 503;  DB 7;  Length 749;
  Best Local Similarity  82.3%;  Pred. No. 1e-125;
  Matches  591;  Conservative    0;  Mismatches  121;  Indels    6;  Gaps    1;

Qy        545 AGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAG 604
              |||||||||||||||||||||| | |||||||||| || ||||||||| ||||||||||
Db          7 AGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAG 66

Qy        605 GAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATT 664
              |||||||| |||||||| | |  ||||||||||||||  ||   | |||    ||||||
Db         67 GAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGTTCAAGTAATTATGAGATTATT 126

Qy        665 ACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTT 724
              |||||| |||||||| |||||||||||||| | |||||||||  |||||||||||
Db        127 ACAGGTCTTGAAGAAGGGTATTTCACAGCTATCATGTTGAAGAAATCAAGAGTGAAATTA 186

Qy        725 AAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 784
              ||||| || ||||||||||||||||||||| ||||||||||||||||||||  |||| ||
Db        187 AAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTG 246

Qy        785 AATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 844
              ||||||||  ||   ||||||||  ||||||||||||||||||||||||||||||||||
Db        247 CATGTGAGTGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGA 306

Qy        845 GAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCT 904
              |  ||| ||||||||||||| |  || |||||||  |||| |     ||||||||||||||
Db        307 GGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCT 366

Qy        905 CCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATC 964
              ||||| ||| | || ||||||||||||||||||||| ||| |||| |||| ||| |||| |
Db        367 CCAGAGTCAGCTACAGTTATATTTGCAGGAGATACGAATCTAAGGGATCGAGAGGTTACC 426

Qy        965 AAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAA 1024
              | |||||||||||||||  ||||| |  | |||| |||||| ||||| ||||||||||||
Db        427 AGATGTGGTGGTTTACCCAACAACATCGTGGATGTCTGGGAGTTTTTGGGCAAACCTAAA 486

Qy       1025 CATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTAT 1084
              |||||||||||||||||||||||  ||    ||     || ||  | ||  |||||||| |
Db        487 CATTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAATAACTGCTGCTTGT 546

Qy       1085 AAGCATCGTTTTGATCGAATATTTTTCAGA------GCAGAAGAGGGGCACCTTATTCCT 1138
              |||| |||||||||||||||||||||||      |||||||||| ||| || ||||
Db        547 AAGCTTCGTTTTGATCGAATATTTTTCAGAGCAGCCGCAGAAGAGGGACACATTGTTCCC 606

Qy       1139 CAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCAC 1198
              |||||||| |||||| |||| || ||||||||||||||||||||||||| |||||||||
Db        607 CAAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTAGATTTCCTAGTGATCAC 666

Qy       1199 TGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
              ||||| ||  |||||| |||  || || |||||| ||| ||||  ||  ||    |||
Db        667 TGGGGTCTTCTGTGCAACTTANATATAATATTGTAAAATGCTTTTCAAATGTGAGTTT 724


RESULT 54
CX166335
LOCUS       CX166335                810 bp    mRNA    linear   EST 23-DEC-2004
DEFINITION  HESC2_39_B09.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone
            IMAGE:7469923 5', mRNA sequence.
ACCESSION   CX166335
VERSION     CX166335.1  GI:56796415
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 810)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15771  row: d  column: 17
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 810.
```

```
FEATURES             Location/Qualifiers
     source          1. .810
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7469923"
                     /sex="male"
                     /tissue_type="human embryonic stem cells differentiated to
                     an early endodermal cell type"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_258"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early endodermal cell type. Cell line
                     id and NIH Registry designation is BG01. Positive for
                     GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP
                     expression. Passage number 40. cDNA primed using oligo-dT
                     primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and
                     cloned into the EcoRV/NotI sites of pExpress-1. This
                     primary library is non-normalized (normalized primary
                     library is NIH_MGC_259). It was constructed by Express
                     Genomics (Frederick, MD). Sequence ends have been trimmed
                     to exclude vector and regions below Phred quality 16.
                     Three-prime sequences are presented as their reverse
                     complement and have been trimmed to exclude polyA. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             37.7%;  Score 495.2;  DB 8;  Length 810;
  Best Local Similarity   77.5%;  Pred. No. 1.4e-123;
  Matches  626;  Conservative    0;  Mismatches  178;  Indels    4;  Gaps    2;

Qy        201 AGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGG 260
              | | ||  |||     |||||| | |||||| |||| |||  |||  |||||| | ||| |
Db          4 AAGAGGGCGAGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCG 63

Qy        261 GGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGA 320
                | |||||  || |  |||  |   || |||||   |||||||||||| ||| | | |
Db         64 CAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAA 123

Qy        321 AAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTC 380
              || |||| | ||||||||||||  || | | | ||    | | |   ||||  |||
Db        124 GGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTG 183

Qy        381 CCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAA 440
              ||   ||   || || || |||||||||| |||||||| || ||  ||| |||| | |
Db        184 AAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCA 243

Qy        441 CCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTT 497
              ||| ||   ||  ||| ||||||| |   ||||  | |||| || ||||||   | |||
Db        244 CCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTC 303

Qy        498 TCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGG 557
              ||||||||||||||||||||||||||||  | |||||||| | ||||||||||||||||
Db        304 TCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGG 363

Qy        558 TGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCC 617
              ||||||||| | |||||||||| || |||||||| ||||||||||||||||||| ||||
Db        364 TGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCC 423

Qy        618 CATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAG 677
              |||| | |  |||||||||||||||| ||    | |||    |||||||||||| ||||||
Db        424 CATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAG 483

Qy        678 AAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGA 737
              |||||||||||||||||| | |||||||||  |||||||||||| ||||| || ||||
Db        484 AAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGA 543

Qy        738 TTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGG 797
              |||||||||||||| |||||||||||||||||||  |||| ||  |||||| | ||
Db        544 TTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGT 603

Qy        798 GTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGG 857
              ||||||||  |||||||||||||||||||||||||||||||||||  ||| ||||||
Db        604 CAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGG 663

Qy        858 AACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCA 917
              ||||||| |  || |||||||  ||| |    |||||||||||||||||||| ||| | |
Db        664 AACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTA 723

Qy        918 CGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTT 977
              | |||||||||||||||||||||||| |||| |||| ||| |||| || ||||||||||
Db        724 CAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTT 783

Qy        978 TACCTGACAACGTTTTTGATGCCTGGGA 1005
              ||||  ||||| || | |||| ||||||
Db        784 TACC-CACAACATTGTGGATGTCTGGGA 810


RESULT 55
BG740396
LOCUS       BG740396                870 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602634171F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4779449 5',
            mRNA sequence.
ACCESSION   BG740396
VERSION     BG740396.1  GI:14051049
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 870)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10636  row: m  column: 18
            High quality sequence stop: 826.
FEATURES             Location/Qualifiers
     source          1. .870
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4779449"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn3"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

  Query Match             37.5%;  Score 492.6;  DB 2;  Length 870;
  Best Local Similarity   77.7%;  Pred. No. 7.5e-123;
  Matches  647;  Conservative    0;  Mismatches  179;  Indels    7;  Gaps    4;

Qy        244 GGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAA 303
              ||  ||||| | ||| |   | |||||  || |  |||  |   || |||||   |||||
Db          1 GGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAA 60

Qy        304 CGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGG 363
              ||||||| ||| | | |  || |||| | ||||||||||||  || | | | ||    |
Db         61 CGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGC 120
```

```
Qy        364 GTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGA 423
Db        121 CTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGA 180

Qy        424 GGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGA 480
Db        181 AGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAA 240

Qy        481 TAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCC 540
Db        241 TGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTC 300

Qy        541 CGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCT 600
Db        301 AGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCT 360

Qy        601 ACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAAT 660
Db        361 ACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGAT 420

Qy        661 TATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAA 720
Db        421 TATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAA 480

Qy        721 ATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATG 780
Db        481 ATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATG 540

Qy        781 CGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCAC 840
Db        541 TGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCAC 600

Qy        841 CAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGA 900
Db        601 CAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGA 660

Qy        901 GGCTCCAGATTCAACCACGGTTAT-ATTTGCAGGAGATACAAATTTAAGAGATCAAGAAG 959
Db        661 GGCTCCAGAGTCAGCTACAGTTATAATTCGCAGGAGATACAAATCTAAGGGATCGAGAGG 720

Qy        960 TTATCAAATGTGGTGGTTTA--CCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAA 1017
Db        721 TTACCAGATGTGGTGGTTTACCCCAAACAACATTGTGGATGTCTGGG-ATTTTTGGGCAA 779

Qy       1018 ACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGA 1070
Db        780 ACCTAAACATTGCCAGTATACCTGGGATACCAAATTGAACTCTAATCCTTGGA 832
```

RESULT 56
BM553049

```
LOCUS       BM553049               1192 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6542413 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742798
            5', mRNA sequence.
ACCESSION   BM553049
VERSION     BM553049.1  GI:18791437
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1192)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12761  row: i  column: 07
            High quality sequence stop: 747.
FEATURES             Location/Qualifiers
     source          1. .1192
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5742798"
                     /tissue_type="medulla"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_119"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source normal medulla from
                     anonymous male age 27.  Library  is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed  upon
                     cloning).  Average insert size 1.3 kb, insert size range
                     0.9-3 kb.  Library is normalized and enriched for
                     full-length clones and  was constructed by C. Gruber
                     (Invitrogen).  Research Genetics  tracking code 013. Note:
                     this is a NIH_MGC Library."
ORIGIN
```

Query Match 37.3%; Score 489.8; DB 3; Length 1192;
Best Local Similarity 73.9%; Pred. No. 4.8e-122;
Matches 664; Conservative 0; Mismatches 227; Indels 8; Gaps 3;

```
Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
Db         74 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 133

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
Db        134 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 193

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
Db        194 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 253

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
Db        254 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 313

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
Db        314 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 373

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db        374 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 433

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
Db        434 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 493

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
Db        494 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 553

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
Db        554 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 613

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
Db        614 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 673

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCC 827
Db        674 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 733

Qy        828 ATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTTG 887
```

```
Db        734 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTTA 793

Qy        888 GAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAA 947

Db        794 AGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 853

Qy        948 GAGATCAAGAAGTTATCA-AATGTGGTGGTTTACCTGACAACGTTTTTGATGCC----TG 1002

Db        854 GGGATCGAGAGGTTACCACGATGTGGTGGTTTACCCACAACATTGTGGGATGTCCTGGGG 913

Qy       1003 GGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACA 1061

Db        914 AGTTTTTGGGGGCAACCCTAAACATTGCCCAGGTATACATGGGGATACCACAAATGAAA 972


RESULT 57
BU159911
LOCUS       BU159911                870 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7933863 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6144208
            5', mRNA sequence.
ACCESSION   BU159911
VERSION     BU159911.1  GI:22673821
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 870)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13468  row: n  column: 17
            High quality sequence stop: 677.
FEATURES             Location/Qualifiers
     source          1. .870
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6144208"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match           37.0%;  Score 485.6;  DB 5;  Length 870;
  Best Local Similarity   77.1%;  Pred. No. 6.1e-121;
  Matches  642;  Conservative    0;  Mismatches  182;  Indels    9;  Gaps    4;

Qy        220 GAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGAT 279

Db          1 GAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGT 60

Qy        280 GGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTT 339

Db         61 GGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTT 120

Qy        340 CGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGA 399

Db        121 CGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAA 180

Qy        400 GGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCC 459

Db        181 GACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCC 240

Qy        460 ATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGA 516

Db        241 ATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGA 300

Qy        517 TGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTT 576

Db        301 TGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTT 360

Qy        577 GTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAA 636

Db        361 GTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAA 420

Qy        637 GAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTAT 696

Db        421 GAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTAT 480

Qy        697 ACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATAC 756

Db        481 AATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTAC 540

Qy        757 CAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCT 816

Db        541 CAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCT 600

Qy        817 TATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAA 876

Db        601 TATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAA 660

Qy        877 AACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGA 936

Db        661 AATGGTTTTAAAGAAANATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGA 720

Qy        937 TACAAATTTAAGAGATCAAGAAGTTATC--AAATGTGGTGGTTTACCTGACAACGTTTTT 994

Db        721 TACAAATCTAAGGGATCGAGAGGTTACCCAGATGTGNGTGGTTTACCCAACAACATTGTG 780

Qy        995 GATGCCTGGGAATTTTTAGG---CAAACCTAAACATT-GCCAGTATACATGGG 1043

Db        781 GATGTCTGGGGAGTTTTTGGGGCAAACCCTANACATTGGCCAGTATACATGGG 833


RESULT 58
AA163045
LOCUS       AA163045                493 bp    mRNA    linear   EST 16-FEB-1997
DEFINITION  mr28g09.r1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:598816
            5', mRNA sequence.
ACCESSION   AA163045
VERSION     AA163045.1  GI:1739008
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 493)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
```

```
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:364248
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 480.
FEATURES              Location/Qualifiers
     source           1. .493
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:598816"
                      /sex="male"
                      /tissue_type="Spleen"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
                      /clone_lib="Soares mouse 3NbMS"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. RNA
                      provided by Dr. Bertrand Jordan. Library went through
                      three rounds of normalization, and was constructed by
                      Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match           36.5%;  Score 479.4;  DB 1;  Length 493;
  Best Local Similarity   99.4%;  Pred. No. 2.5e-119;
  Matches  492;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy        443 ATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATT 60

Qy        503 ACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGT 562
              |||||||||||||||||||||||||||||||||||||||||||| |  ||||||||||||||
Db         61 ACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGACG--TCGAGGGGTGTGT 118

Qy        563 TCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCCATAC 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 TCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCCATAC 178

Qy        623 TGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGA 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 TGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGA 238

Qy        683 TATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATT 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATT 298

Qy        743 CCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGA 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 CCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGA 358

Qy        803 AATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGA 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 AATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGA 418

Qy        863 ATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTT 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 ATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTT 478

Qy        923 ATATTTGCAGGAGAT 937
              |||||||||||||||
Db        479 ATATTTGCAGGAGAT 493


RESULT 59
```

```
AU015298/c
LOCUS       AU015298                587 bp    mRNA    linear   EST 15-OCT-1998
DEFINITION  AU015298 Mouse two-cell stage embryo cDNA Mus musculus cDNA clone
            J0710B09 3', mRNA sequence.
ACCESSION   AU015298
VERSION     AU015298.1  GI:3370089
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 587)
  AUTHORS   Ko,M.S.H, Kitchen,J.R., Wang,X., Threat,T.A., Sun,T., DePalma,G.E.,
            Liang,Y., Kargul,G.J., Sharara,R. and Doi,H.
  TITLE     Systematic analyses of genes expressed in 2-cell stage mouse
            embryos (The ERATO/Doi Project at Wayne State University)
            (Ko,M.S.H. et al.)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Hirofumi Doi
            Doi Bioasymmetry Project, ERATO
            Japan Science and Technology Corporation (JST)
            WBG Marive East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan
            Email: hd@bioa.jst.go.jp.
FEATURES              Location/Qualifiers
     source           1. .587
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="J0710B09"
                      /dev_stage="two-cell stage embryo"
                      /clone_lib="Mouse two-cell stage embryo cDNA"
ORIGIN

  Query Match           36.2%;  Score 474.4;  DB 1;  Length 587;
  Best Local Similarity   95.1%;  Pred. No. 6.2e-118;
  Matches  509;  Conservative    0;  Mismatches   24;  Indels    2;  Gaps    2;

Qy        780 GCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACAT-CCCATTTGGAGAGC 838
              |||||||||||||||||| |||||||||| |||||||| ||||||||||||||||||||
Db        580 GCGTAAATGTGAGTTTGGGTGNAAATGAATTTTCCCTTATGACATCCCCATTTGGAGAGC 521

Qy        839 ACCAGAGAACATTCTGCGG-AACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCA 897
              |||  |||||||||||||| ||||||||| ||||||||||||||||  | | ||||
Db        520 ACCCNAGAACATTCTGCGGAAACGAATAANACAATTAAAAACTGTTTCTTGGAAAAATGC 461

Qy        898 AGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGA 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||  | | |
Db        460 AGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTAGAGATCAAGAA 401

Qy        958 AGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAA 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 GTTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAA 341

Qy       1018 ACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGC 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 ACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGC 281

Qy       1078 TGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 TGCTTATAAGCATCGTTTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCC 221

Qy       1138 TCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCA 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 TCAAAGTTTAGACCTTGTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCA 161

Qy       1198 CTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTA 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 CTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTA 101

Qy       1258 CACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        100 CACGTTTGTTAGCACTAGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 46


RESULT 60
CB579756
LOCUS       CB579756                615 bp    mRNA    linear   EST 03-APR-2003
DEFINITION  AMGNNUC:NRDG1-00175-E9-A nrdg1 (10855) Rattus norvegicus cDNA clone
            nrdg1-00175-e9 5', mRNA sequence.
ACCESSION   CB579756
VERSION     CB579756.1  GI:29523797
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 615)
  AUTHORS   Amgen EST Program.
  TITLE     Amgen Rat EST Program
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00175  row: e  column: 9.
FEATURES             Location/Qualifiers
     source          1. .615
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="nrdg1-00175-e9"
                     /tissue_type="Dorsal Root Ganglia"
                     /clone_lib="nrdg1 (10855)"
                     /note="Vector: pSPORT1; Site_1: Sal1; Site_2: Not1; rat
                     dorsal root ganglia"
ORIGIN

  Query Match             36.1%;  Score 474;  DB 6;  Length 615;
  Best Local Similarity   86.8%;  Pred. No. 8.1e-118;
  Matches  538;  Conservative    0;  Mismatches   70;  Indels   12;  Gaps    1;

Qy         73 GGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGG 132
              | |||| |||  | ||||||||||||| |||||||| |||||||| ||||||||||||||
Db          8 GCGCGGAAAGTGGGGTGAAGAGCGGGTATTTTGAGGAGACCCTGCAGCGATGGCGTCTGG 67

Qy        133 CAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGA 192
              |||||||||||||||||||||| | |||| ||          ||||| |||||||||
Db         68 CAGCAGTTCCGATGCGGCGGAGTCGGCAGAGC-----------CGGCGGCGGCGCCCGC 115

Qy        193 AGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGC 252
              ||||||| || |||||||||| | |||||||||||||||||||||| |||||| |||||
Db        116 AGCAGCAGAGACGGAGGAGGATCAGGTGAAGAGGCGGCGGCTTCAGTCCCTGGGATTTGC 175

Qy        253 GTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCA 312
              |||||||   | | |||||| | ||| |||   |||   |||||  |||||| |||||||
Db        176 GTTGGTGACGAGCTGCGACACTACGGTGGCTTCCACTTTCCTGTCGGAGAACAACTGGCA 235

Qy        313 GACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCG 372
              |||| | |||||  |||||||| |||||||| |||||||||||||| || | |||| |
Db        236 GACGAAAAAAGCGTTGAGCGCCTTCTTCGAGCAGCCAGAGAACGACCTAGCGCGGCCTCA 295

Qy        373 CCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAA 432
              ||||||||| || |||| |||||||||| |||||||||||||||||| ||||||||||
Db        296 CCAGCCTCCGACATCCTCCAAGTCCGAGGACTATGTTGATCTAACCAATGAGGATGCAAA 355

Qy        433 TGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTAT 492
              |||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        356 TGATACCACCATTTTAGAAACCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTAT 415

Qy        493 TTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCG 552
              |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||
Db        416 CTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTCCCAGAGAGGGCTCG 475

Qy        553 AGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTAT 612
              |||||||||||||||||| |||||||| ||||| || ||||||||||||||||||| ||
Db        476 AGGGGTGTGTTCCTGCCTTGCTTTGTACAGTCCGGACGTGGTATTTCTACAGGAAGTGAT 535

Qy        613 CCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAA 672
              ||| | |||||||||||||||| |||||||||  || |||||  ||||||||||||||
Db        536 TCCCTCCTACTGTGCCTACCTAAGGAAGAGAGCACGCACTTACAACATTATTACAGGTAA 595

Qy        673 TGAAGAAGGATATTTCACAG 692
              ||||||||||||||||||||
Db        596 TGAAGAAGGATATTTCACAG 615


RESULT 61
BG740339
LOCUS       BG740339                932 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602635289F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4780318 5',
            mRNA sequence.
ACCESSION   BG740339
VERSION     BG740339.1  GI:14050992
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 932)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10639  row: a  column: 23
            High quality sequence stop: 759.
FEATURES             Location/Qualifiers
     source          1. .932
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4780318"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn3"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

  Query Match             36.0%;  Score 472;  DB 2;  Length 932;
  Best Local Similarity   79.4%;  Pred. No. 3.2e-117;
  Matches  597;  Conservative    0;  Mismatches  150;  Indels    5;  Gaps    3;

Qy        324 CCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCA 383
              | |||| | ||||||||||||  || | | | ||    | | |  ||||  |||   |
Db          1 CTCTGAACTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAA 60

Qy        384 CGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCA 443
              |    ||   || || || |||||||||| |||||||| || ||  ||| |||| | ||||
Db         61 CCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCA 120

Qy        444 TTTTAGAAGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCA 500
              ||   ||  ||| |||||||  |   ||||  | |||| || ||||||    | ||| |||
Db        121 CTTCTAAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCA 180

Qy        501 TTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGT 560
              |||||||||||||||||||||||||  |  ||||||| | ||||||||||||||||||||
```

```
Db        181 TTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGT 240

Qy        561 GTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCAT 620

Db        241 GTTCCTACTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCAT 300

Qy        621 ACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAG 680

Db        301 ATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAG 360

Qy        681 GATATTTCACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTA 740

Db        361 GATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTA 420

Qy        741 TTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTG 800

Db        421 TTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAG 480

Qy        801 GAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAAC 860

Db        481 GAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAAC 540

Qy        861 GAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGG 920

Db        541 GAATGAATCAGTTAAAAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAG 600

Qy        921 TTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTAC 980

Db        601 TTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGTGGTGGTTTAC 660

Qy        981 CTGACAACGTTTTTGATGCCTGGGAATT-TTTAGGCAAACCTAAACATTGCCAG-TATAC 1038

Db        661 CCAACAACATTGTGGATGTCTGGGAGTTCCTTGGGCAAACCTAAACATTGCCAGTTATAC 720

Qy       1039 ATGGGATACGAAAGCAAATAACAACCTCAGGA 1070

Db        721 ATGGGATACACAAATGAACTCTAATCTTGGGA 752
```

RESULT 62
AA681654
LOCUS AA681654 471 bp mRNA linear EST 05-DEC-1997
DEFINITION vr43d12.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone IMAGE:1123415 5', mRNA sequence.
ACCESSION AA681654
VERSION AA681654.1 GI:2663794
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 471)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:612751
High quality sequence stop: 469.
FEATURES Location/Qualifiers
source 1. .471
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:1123415"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/clone_lib="Knowles Solter mouse 2 cell"
/note="Organ: embryo; Vector: pBluescribe (modified); Site_1: MluI; Site_2: SalI; Cloned unidirectionally from mRNA prepared from 13,500 2-cell stage embryos. Primer: SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'. cDNAs were cloned into the MluI/SalI sites of a modified pBluescribe vector using commercial linkers (NEB). Average insert size: 1.2 kb."
ORIGIN

Query Match 35.8%; Score 469.4; DB 1; Length 471;
Best Local Similarity 99.8%; Pred. No. 1.3e-116;
Matches 470; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        638 AAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATA 697

Db          1 AAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATA 60

Qy        698 CTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACC 757

Db         61 CTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACC 120

Qy        758 AAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTT 817

Db        121 AAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTT 180

Qy        818 ATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAA 877

Db        181 ATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAA 240

Qy        878 ACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGAT 937

Db        241 ACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGAT 300

Qy        938 ACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGAT 997

Db        301 ACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACATGACAACGTTTTTGAT 360

Qy        998 GCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAAT 1057

Db        361 GCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAAT 420

Qy       1058 AACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTT 1108

Db        421 AACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTT 471
```

RESULT 63
AA636415
LOCUS AA636415 468 bp mRNA linear EST 22-OCT-1997
DEFINITION vq86d01.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone IMAGE:1109185 5', mRNA sequence.
ACCESSION AA636415
VERSION AA636415.1 GI:2560194
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Murinae; Mus.
REFERENCE 1 (bases 1 to 468)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

```
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:607353.
FEATURES             Location/Qualifiers
     source          1. .468
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J x DBA/2J F1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1109185"
                     /tissue_type="embryo"
                     /dev_stage="2-cell"
                     /lab_host="DH10B"
                     /clone_lib="Knowles Solter mouse 2 cell"
                     /note="Organ: embryo; Vector: pBluescribe (modified);
                     Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
                     mRNA prepared from 13,500 2-cell stage embryos. Primer:
                     SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'.  cDNAs
                     were cloned into the MluI/SalI sites of a modified
                     pBluescribe vector using commercial linkers (NEB).
                     Average insert size: 1.2 kb."
ORIGIN

  Query Match           35.7%;  Score 468;  DB 1;  Length 468;
  Best Local Similarity  100.0%;  Pred. No. 3.2e-116;
  Matches  468;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        638 AAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATA 697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATA 60

Qy        698 CTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACC 757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACC 120

Qy        758 AAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTT 817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTT 180

Qy        818 ATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAA 877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAA 240

Qy        878 ACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGAT 937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGAT 300

Qy        938 ACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGAT 997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTGAT 360

Qy        998 GCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAAT 1057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACGAAAGCAAAT 420

Qy       1058 AACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATA 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATA 468


RESULT 64
CK952220
LOCUS       CK952220                680 bp    mRNA    linear   EST 15-MAR-2004
DEFINITION  4091710 BARC 10BOV Bos taurus cDNA clone 10BOV29_F21 5', mRNA
            sequence.
ACCESSION   CK952220
VERSION     CK952220.1  GI:45466600
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Sonstegard, T.S., Van Tassell, C.P., Matukumalli, L.K., Harhay,
            G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
  TITLE     Production of EST from cDNA libraries derived from immunologically
            activated bovine gut
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Tad S. Sonstegard
            Bovine Functional Genomics Laboratory
            Animal and Natural Resources Institute
            Bdlg. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
            Tel: 3015048416
            Fax: 3015048414
            Email: tads@anri.barc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            0.000925 using options -trim_alt " -trim_fasta. Vector identified
            by cross_match using options -minmatch 12 -minscore 18
            Plate: 29  row: F  column: 21
            Seq primer: CCCAGTCACGACGTTGTAAAACG
            High quality sequence stop: 680.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /strain="Holstein"
                     /db_xref="taxon:9913"
                     /clone="10BOV29_F21"
                     /sex="Male"
                     /tissue_type="Pooled"
                     /dev_stage="Multiple"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="BARC 10BOV"
                     /note="Organ: Small Intestine; Vector: pAgen-1; Site_1:
                     EcoRV; Site_2: NotI; Equimolar amounts of mRNA extracted
                     from proximal jejunums of 18 and 21 wk old steers, and
                     distal ileums of 14 day old calves. proximal jejunum
                     exposed to C. oncophora for 3 and 6 weeks, and distal
                     ileum exposed to C. parvum for 7 days"
ORIGIN

  Query Match           35.5%;  Score 466;  DB 7;  Length 680;
  Best Local Similarity  81.5%;  Pred. No. 1.3e-115;
  Matches  554;  Conservative    0;  Mismatches  120;  Indels    6;  Gaps    1;

Qy        509 AATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGC 568
              ||||||||||||||  ||      ||||||  |  ||||||||||||||||||||||||||
Db          1 AATATTGATGGATTGGACATGAACAATCTCCTAGAGAGGGCTCGAGGGGTGTGTTCCTAT 60

Qy        569 CTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCC 628
              ||  |||||||  ||  ||  ||||||  ||||||||||||||||||  ||  |||||||||  |  ||||
Db         61 TTAACTTTGTACAGCCCTGATGTGATATTTCTACAGGAAGTCATTCCCCCATATTATGCC 120

Qy        629 TACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTC 688
              ||||||||||||  ||||    |||||||  ||||||||||||  ||||||||||||||||
Db        121 TACCTAAAGAAGAAAGCAAGTAGTTACAAGATTATTACAGGTCGTGAAGAAGGATATTTC 180

Qy        689 ACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTT 748
              |||||||||  |  |||||||||  |||||||||  ||||||||  ||  ||  ||||||||||||
Db        181 ACAGCTATAATGTTGAAGAAATCAAGAGTGAAGTTTAAAAGCCAAGAAATTATTCCTTTT 240

Qy        749 CCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAA 808
              ||||||||  ||||||||||||||||  |  ||  ||  ||||||||  ||  |  |||||||||
Db        241 CCAAATACGCAAATGATGAGAAACCTTTTGTGTGTTCATGTGAGTGTGTCAGGAAATGAA 300
```

```
Qy        809 TTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGA 868
              ||||||||||||| || |||||||| |||||||||   ||   ||  |||||||| |
Db        301 CTTTGCCTTATGACTTCTCATTTGGAAAGCACCAGAGGGCACGCTAAGGAACGAATGAAT 360

Qy        869 CAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTT 928
              || || |||| ||| | | ||||||||||||||||||  ||| | || |||||||||
Db        361 CAGTTGAAAATGGTTTTAGAGAAAATGCAAGAGGCTCCAGGATCAGCTACAGTTATATTT 420

Qy        929 GCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAAC 988
              || |||||||| ||| | || |||||||||||||  ||||||||||||||||||  |||||
Db        421 GCTGGAGATACGAATCTCAGGGATCAAGAAGTTACCAAATGTGGTGGTTTACCCAACAAC 480

Qy        989 GTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAGTATACATGGGATACG 1048
              |||| |||| |||||| ||||| |||||||||||  ||||||||||||||||||||||||
Db        481 ATTTTGGATGTCTGGGAGTTTTTGGGCAAACCTAAGCATTGCCAGTATACATGGGATACA 540

Qy       1049 AAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTTTTGATCGAATATTT 1108
              ||   ||    || ||  | ||  ||||| || ||| | ||||||||||||||||||||
Db        541 CAAATGAACTCTAATCTTGGAATAGCTGCTAATTGTAAACTTCGTTTTGATCGAATATTT 600

Qy       1109 TTCAG------AGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTTGTTGGGTTG 1162
              || ||      |||||||| || ||| ||||||| |||||||| ||||| ||||| ||
Db        601 TTTAGAGCTGCAGCAGAAGGTGGCCACATTATTCCCCAAAGTTTGGACCTGCTTGGGCTG 660

Qy       1163 GAAAAACTGGACTGTGGTAG 1182
              ||||||||||||||||||||
Db        661 GAAAAACTGGACTGTGGTAG 680
```

RESULT 65
BM554324
LOCUS BM554324 1035 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6546794 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742466 5', mRNA sequence.
ACCESSION BM554324
VERSION BM554324.1 GI:18793827
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1035)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12760 row: k column: 11
High quality sequence stop: 738.

FEATURES Location/Qualifiers
source 1. .1035
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5742466"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 35.5%; Score 465.2; DB 3; Length 1035;
Best Local Similarity 72.5%; Pred. No. 2.4e-115;
Matches 629; Conservative 0; Mismatches 235; Indels 4; Gaps 2;

```
Qy        151 GGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGA 210
              |||    | ||    ||    ||  | | | || |   |  | | | | | | ||  ||
Db        169 GGAAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGGAGAAGAGGGCGA 228

Qy        211 GGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGA 270
              |     |||||| | |||||| |||| ||| ||| ||||| | ||| |   | |||||
Db        229 GCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGA 288

Qy        271 CCCCACGATGGTCCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAG 330
              || |  ||| |  || ||||| |||||||||||| ||| | | |  || ||||
Db        289 TGCCGCAGTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAA 348

Qy        331 CGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTT 390
              | |||||||||||| || | | | ||    | | |  |||| ||| |  || ||||
Db        349 CTCCTACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTC 408

Qy        391 CAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGA 450
              || || || |||||||||| |||||||| || ||  ||| |||| | |||| || |
Db        409 TGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAA 468

Qy        451 AGCCAGTCCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTG 507
              | ||| |||||| |   |||| | |||| || ||||||   | ||| ||||||||||
Db        469 AATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTG 528

Qy        508 GAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTG 567
              |||||||||||||||||| | |||||||| | |||||||||||||||||||||||||||
Db        529 GAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTA 588

Qy        568 CCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGC 627
              | |||||||||| || ||||||||| |||||||||||||||||||| |||||||| | |
Db        589 CTTAGCTTTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAG 648

Qy        628 CTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTT 687
              |||||||||||||||| ||   | |||    |||||||||||| ||||||||||||||||
Db        649 CTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTT 708

Qy        688 CACAGCTATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTT 747
              |||||||||| | |||||||||  |||||||||||| ||||| || ||||||||||||||
Db        709 CACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTT 768

Qy        748 TCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGA 807
              ||| | |||||||||||||||||||  |||| ||  |||||| | ||   ||||||||
Db        769 TCCNAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGA 828

Qy        808 ATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAG 867
              |||||||||||||||||||||||||||||||||||||||||  ||| ||||||  |  || |
Db        829 GCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAG-GCATGCTGCGGGAACGATGAA 887

Qy        868 ACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATT 927
              || ||||||    ||   |||||    |  ||  |||||| ||| | || | ||||||
Db        888 TCAGTTAAAATGGTTTTAAAGAAAATGCAAGAAGCTTCCAGAGTCAGCTACAGNTATATT 947

Qy        928 TGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAA 987
              |||||||||||||||| ||| |||| |||||||| | ||||| | || ||  |  ||||
Db        948 TGCAGGAGATACAAATCTAAAGGATCGAGAAGTTACCCAATGTTGGGGGTTTACCCACAA 1007

Qy        988 CGTTTTTGATGCCTGGGAATTTTTAGGC 1015
              | || | ||      || |||||| |||
Db       1008 CATTGTGGAATGTCTGGGATTTTTTGGC 1035
```

RESULT 66
BM555041
LOCUS BM555041 1103 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6545705 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5737150

```
            5', mRNA sequence.
ACCESSION   BM555041
VERSION     BM555041.1  GI:18795166
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1103)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12746  row: m  column: 23
            High quality sequence stop: 678.
FEATURES             Location/Qualifiers
     source          1. .1103
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5737150"
                     /tissue_type="duodenal adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_88"
                     /note="Organ: small intestine; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                     oligo-dT primed. Average insert size 1.767 kb. Library
                     enriched for full-length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match           35.0%;  Score 459.6;  DB 3;  Length 1103;
  Best Local Similarity  73.6%;  Pred. No. 8.2e-114;
  Matches  641;  Conservative    0;  Mismatches  220;  Indels   10;  Gaps    4;

Qy        171 GGCGGGCGGCGTCGGCGCCCGAAGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGC 230
              | | || ||    | |   |  || |  | | ||  |||      |||||| | |||||
Db         46 GCCTGGAGGGCGGGAGGGAGGCGGCGGAGGAAGAGGGCGAGCCTGAGGTGAAAAAGCGGC 105

Qy        231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCG 290
              | |||| |||  |||  ||||| | ||| |   | |||||  || |  |||  |   ||
Db        106 GACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCTCAGTGCT 165

Qy        291 TCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350
              |||||   |||||||||||| ||| | | |  || |||| | ||||||||||||   || |
Db        166 TCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAGCCTCCGG 225

Qy        351 AGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTG 410
              | | ||    | | |  |||| |||   ||   ||   || || || ||||||||||
Db        226 TGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 285

Qy        411 ATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGA---A 467
              | |||||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||| |   |
Db        286 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCTGAAGATA 345

Qy        468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATG 527
              |||  | |||| || ||||||    | ||| ||||||||||||||||||||||||||||
Db        346 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGATTAGATC 405

Qy        528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAG 587
              |  |||||||  | |||||||||||||||||||||||| | ||||||||||  || ||||
Db        406 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTACAGCCCAG 465

Qy        588 ATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAG 647
              ||||| |||||||||||||||||| |||||||| | |  ||||||||||||||| ||
Db        466 ATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAA 525

Qy        648 CCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGA 707
              | |||    |||||||||||| ||||||||||||||||||||||||||| | |||||||
Db        526 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 585

Qy        708 AAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767
              ||  |||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        586 AATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGA 645

Qy        768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACAT-CC 826
              |||||||  |||| ||  |||||| | ||   ||||||||  ||||||||||||||| ||
Db        646 GAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCCTTATGACATCCC 705

Qy        827 CATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACTGTTCTT 886
              ||||||||||||||||||  ||| ||||||||||||| |  ||  ||||||  ||| |
Db        706 CATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGNTAAAAATGGTTTTA 765

Qy        887 GGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACAAATTTA 946
              |||||||||||  |||||||| ||| |    |||||||||||||||||||||||| ||
Db        766 AAGAAAATGCAAGAAGCTCCAGAGTCACCTTACGTTATATTTGCAGGAGATACAAATCTA 825

Qy        947 AGAG--ATCAAGAAGTTATCAAATGTGGTGGTTTACCTGACAACGTTTTTG----ATGCC 1000
              || |     |    ||   | ||||||||| ||  |  |||||  || ||
Db        826 AGGGGATCGAGAGGGTACCCGAATGTGGTGGGTTTACCCACAACCATTGTGGAATGGCTG 885

Qy       1001 TGGGAATTTTTAGGCAAACCTAAACATTGCC 1031
              || | ||||| |||||||||||| |||||
Db        886 GGGAATTTTTTGGGCAAACCTAAACTTTGCC 916


RESULT 67
BU169945
LOCUS       BU169945                883 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7913097 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6024760
            5', mRNA sequence.
ACCESSION   BU169945
VERSION     BU169945.1  GI:22683929
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13235  row: m  column: 17
            High quality sequence stop: 413.
FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6024760"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
```

ORIGIN

Query Match 35.0%; Score 459.4; DB 5; Length 883;
Best Local Similarity 78.3%; Pred. No. 8.7e-114;
Matches 577; Conservative 0; Mismatches 156; Indels 4; Gaps 2;

```
Qy        279 TGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACT 338
              |||  |   ||  |||||   ||||||||||||| ||| | | |  || |||| | |||||||
Db          5 TGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACT 64

Qy        339 TCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCG 398
              ||||||  || | | | ||     |  | |    ||||  |||   ||    ||    || ||
Db         65 TCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCA 124

Qy        399 AGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTC 458
              || |||||||||| ||||||||  || ||   ||| |||| | |||| ||    ||  ||| |
Db        125 AGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCC 184

Qy        459 CATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTG 515
              ||||||  |    ||||  | ||||  || ||||||    | ||| ||||||||||||||||||
Db        185 CATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTG 244

Qy        516 ATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTT 575
              ||||||||||  |  |||||||| | ||||||||||||||||||||||| | ||||||
Db        245 ATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTT 304

Qy        576 TGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAA 635
              ||||  || |||||||||  |||||||||||||||||||||  |||||||| | |  ||||||||
Db        305 TGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAA 364

Qy        636 AGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTA 695
              ||||||||  ||    | |||     |||||||||||| ||||||||||||||||||||||||
Db        365 AGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTA 424

Qy        696 TACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATA 755
              || | |||||||||  ||||||||||| ||||| || ||||||||||||||||||||| ||
Db        425 TAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTA 484

Qy        756 CCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCC 815
              ||||||||||||||||||  |||| ||  |||||| | ||   ||||||||  |||||||
Db        485 CCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGCC 544

Qy        816 TTATGACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAA 875
              |||||||||||||||||||||||||||||||  ||| ||||||||||||  |  || ||||
Db        545 TTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAA 604

Qy        876 AAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAG 935
              |||  ||| |    |||||||||||||||||||||||| ||| | || ||||||||||||||||
Db        605 AAATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAG 664

Qy        936 ATACAAATTTAAGAGATCAAGAAGTTATCAAAT-GTGGTGGTTTACCTGACAACGTTTTT 994
              |||||||| |||| |||| ||| | || || || | || ||||||||  |||| ||
Db        665 ATACAAATCTAAGGGATCGAGAGGGTACCAGATGGGGGGGGTTTACCCAACAACATTGGG 724

Qy        995 GATGCCTGGGAATTTTT 1011
              |||| ||||| | ||||
Db        725 GATGGCTGGGGAGTTTT 741
```

RESULT 68
AA476073
LOCUS AA476073 547 bp mRNA linear EST 18-JUN-1997
DEFINITION vh26g06.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone IMAGE:876634 5', mRNA sequence.
ACCESSION AA476073
VERSION AA476073.1 GI:2203924
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 547)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:516114
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 493.
FEATURES Location/Qualifiers
source 1. .547
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:876634"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares_mammary_gland_NbMMG"
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 34.9%; Score 458; DB 1; Length 547;
Best Local Similarity 96.4%; Pred. No. 1.8e-113;
Matches 534; Conservative 0; Mismatches 10; Indels 10; Gaps 6;

```
Qy         78 GGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGGCAGCA 137
              || | |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          3 GGGAACAGCGTGAAGAGCGGGTGTTTTGAGGGGA-CCTGCGGCGATGGCGTCTGGCAGCA 61

Qy        138 GTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGAAGCAG 197
              ||||||||||||||| ||     |||  ||||||||||||||||||  ||||||||||
Db         62 GTTCCGATGCGGCGGCGC---GAGGCGCGGCAGGGCGGGCGGCGTCG--GCCCGAAGCAG 116

Qy        198 CACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGG 257
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 CACAGGCCGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGG 176

Qy        258 TGGGGGGATGCGACCCCAC-GATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACG 316
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        177 TGGGGGGATGCGACCCCACAGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACG 236

Qy        317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGC--A 294

Qy        377 CCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGAT 354

Qy        437 ACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCT 496
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 ACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTATTTCT 414

Qy        497 TTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGG 556
                |||||||||||||||||||||||||||||||||||||||||||||||||| || |||||||
Db        415 TTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGACGG-TCGAGGG 473

Qy        557 GTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCC 616
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCC 533

Qy        617 CCATACTGTGCCTA 630
                ||||||||||||||
Db        534 CCATACTGTGCCTA 547


RESULT 69
CB577890
LOCUS       CB577890                627 bp    mRNA    linear   EST 03-APR-2003
DEFINITION  AMGNNUC:NRDG1-00195-C6-A nrdg1 (10855) Rattus norvegicus cDNA clone
            nrdg1-00195-c6 5', mRNA sequence.
ACCESSION   CB577890
VERSION     CB577890.1  GI:29521931
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 627)
  AUTHORS   Amgen EST Program.
  TITLE     Amgen Rat EST Program
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Dan Fitzpatrick
            Amgen, Inc
            One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
            Tel: 805 447-4881
            Plate: 00195  row: c  column: 6.
FEATURES             Location/Qualifiers
     source          1. .627
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="nrdg1-00195-c6"
                     /tissue_type="Dorsal Root Ganglia"
                     /clone_lib="nrdg1 (10855)"
                     /note="Vector: pSPORT1; Site_1: Sal1; Site_2: Not1; rat
                     dorsal root ganglia"
ORIGIN

  Query Match             34.8%;  Score 456;  DB 6;  Length 627;
  Best Local Similarity   86.4%;  Pred. No. 6.7e-113;
  Matches  520;  Conservative    0;  Mismatches   70;  Indels   12;  Gaps    1;

Qy         73 GGGCGGGAAGCAGCGTGAAGAGCGGGTGTTTTGAGGGGACCCTGCGGCGATGGCGTCTGG 132
              | |||| ||| | |||||||||||||| |||||||| |||||||| ||||||||||||||
Db         38 GCGCGGAAAGTGGGGTGAAGAGCGGGTATTTTGAGGAGACCCTGCAGCGATGGCGTCTGG 97

Qy        133 CAGCAGTTCCGATGCGGCGGAGCCCGCAGGGCCGGCAGGGCGGGCGGCGTCGGCGCCCGA 192
              |||||||||||||||||||| | |||| ||            ||||| |||||||||
Db         98 CAGCAGTTCCGATGCGGCGGAGTCGGCAGAGC-----------CGGCGGCGGCGCCCGC 145

Qy        193 AGCAGCACAGGCGGAGGAGGACCGGGTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGC 252
              ||||||| || |||||||||| | ||||||||||||||||||||| |||||| |||||
Db        146 AGCAGCAGAGACGGAGGAGGATCAGGTGAAGAGGCGGCGGCTTCAGTCCCTGGGATTTGC 205

Qy        253 GTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCA 312
              |||||||  | | |||||| | ||| |||   |||   |||||  ||||||| |||||||
Db        206 GTTGGTGACGAGCTGCGACACTACGGTGGCTTCCACTTTCCTGTCGGAGAACAACTGGCA 265

Qy        313 GACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCG 372
              |||| | ||||| |||||||| |||||||| |||||||||||||| || | |||| |
Db        266 GACGAAAAAAGCGTTGAGCGCCTTCTTCGAGCAGCCAGAGAACGACCTAGCGCGGCCTCA 325

Qy        373 CCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATGCAAA 432
              ||||||||| || |||| |||||||||| |||||||||||||||||| ||||||||||
Db        326 CCAGCCTCCGACATCCTCCAAGTCCGAGGACTATGTTGATCTAACCAATGAGGATGCAAA 385

Qy        433 TGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTAT 492
              |||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        386 TGATACCACCATTTTAGAAACCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCACTAT 445

Qy        493 TTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCG 552
              |||||||||||||||||||||||||||||||||||||||||||| || ||||||||||
Db        446 CTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTCCCAGAGAGGGCTCG 505

Qy        553 AGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTAT 612
              |||||||||||||||||| |||||||| |||| || |||||||||||||||||||| ||
Db        506 AGGGGTGTGTTCCTGCCTTGCTTTGTACAGTCCGGACGTGGTATTTCTACAGGAAGTGAT 565

Qy        613 CCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAA 672
              ||| | |||||||||||||||| ||||||||||  || |||||  ||||||||||||||
Db        566 TCCCTCCTACTGTGCCTACCTAAGGAAGAGAGCACGCACTTACAACATTATTACAGGTAA 625

Qy        673 TG 674
              ||
Db        626 TG 627


RESULT 70
CB600618
LOCUS       CB600618                546 bp    mRNA    linear   EST 03-APR-2003
DEFINITION  AGENCOURT_13036035 NIH_MGC_176 Mus musculus cDNA clone
            IMAGE:30311958 5', mRNA sequence.
ACCESSION   CB600618
VERSION     CB600618.1  GI:29518474
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDCM96  row: c  column: 07
            High quality sequence stop: 542.
FEATURES             Location/Qualifiers
     source          1. .546
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30311958"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_176"
                     /note="Organ: kidney; Vector: pDNR-LIB; Site_1: SfiI
                     (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); cDNA made
                     by oligo-dT priming and directionally cloned. 5' and 3'
                     adaptors were used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 0.5 kb
                     size fraction. Library created in the laboratory of M.
                     Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
ORIGIN
```

```
  Query Match           34.7%;  Score 455.8;  DB 6;  Length 546;
  Best Local Similarity  99.6%;  Pred. No. 7.3e-113;
  Matches  457;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        854 GCGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCA 913
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGAACGAATAAGACAATTAAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCA 60

Qy        914 ACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGT 973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCACGGTTATATTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGT 120

Qy        974 GGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAG 1033
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTTTACCTGACAACGTTTTTGATGCCTGGGAATTTTTAGGCAAACCTAAACATTGCCAG 180

Qy       1034 TATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGT 1093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATACATGGGATACGAAAGCAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGT 240

Qy       1094 TTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTT 1153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTGATCGAATATTTTTCAGAGCAGAAGAGGGGCACCTTATTCCTCAAAGTTTAGACCTT 300

Qy       1154 GTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGC 1213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTGGGTTGGAAAAACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGGCTCTTGTGC 360

Qy       1214 ACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACT 1273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTTACACGTTTGTTAGCACT 420

Qy       1274 AGTTCTGAATTTGTGTAGGTCTCAACCTTTCAGGACATC 1312
              |||||||||||||||||||||||||||| |||||||||
Db        421 AGTTCTGAATTTGTGTAGGTCTCAACCTTCCAGGACATC 459
```

```
RESULT 71
CN298924
LOCUS       CN298924                729 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000424496911 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN298924
VERSION     CN298924.1  GI:47315338
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 729   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, cell lines H1, H7, and
                     H9"
                     /clone_lib="GRN_ES"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from undifferentiated hES cell lines H1 (p32), H7 (p29),
                     and H9 (p26) maintained in feeder-free conditions"
ORIGIN

  Query Match           34.6%;  Score 454;  DB 7;  Length 729;
  Best Local Similarity  77.5%;  Pred. No. 2.5e-112;
  Matches  563;  Conservative    0;  Mismatches  160;  Indels    3;  Gaps    1;

Qy        224 AGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTC 283
              |||| | ||||| ||| ||| ||||| | ||| |   | |||||  || |  |||
Db          4 ACGCGTCCGCTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCAGTGGCT 63

Qy        284 CCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAG 343
              |   || |||||   |||||||||||| ||| | | |  || |||| | ||||||||||
Db         64 CAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTACTTCGAG 123

Qy        344 CTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCC 403
              |  || | | | ||    | | |  |||| |||   ||    ||   || || || ||
Db        124 CCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAGACC 183

Qy        404 TATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCT 463
              |||||||| |||||| || ||  ||| |||| | |||| ||   ||  ||| ||||||
Db        184 TATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGCCCATCT 243

Qy        464 GGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGA 520
              |   |||| | |||| || ||||||   | ||| |||||||||||||||||||||||
Db        244 GAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGGA 303

Qy        521 TTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTAT 580
              |||||| | ||||||| | ||||||||||||||||||||||||| | ||||||||||
Db        304 TTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCTTTGTAC 363

Qy        581 AGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAG 640
              || ||||||||| |||||||||||||||||||| |||||||| | |  ||||||||||||
Db        364 AGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTAAAGAAG 423

Qy        641 AGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTA 700
              ||| ||   | |||   ||||||||||||| |||||||||||||||||||||||||| |
Db        424 AGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 483

Qy        701 TTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAA 760
              |||||||||  |||||||||||| ||||| || |||||||||||||||||| |||||||
Db        484 TTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGTACCAAA 543

Qy        761 ATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATG 820
              |||||||||||||  |||| ||  ||||||   ||  ||||||||  ||||||||||||
Db        544 ATGATGAGAAACCTTTATGTGTGCATGTGAACGTGTCAGGAAATGAGCTTTGCCTTATG 603

Qy        821 ACATCCCATTTGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAACT 880
              |||||||||||||||||||||||||  ||| ||||||||||||  |   || |||||||
Db        604 ACATCCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAAAATG 663

Qy        881 GTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGGAGATACA 940
              ||| |    |||||||||||||||||||| ||| | || ||||||||||||||||||||
Db        664 GTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCAGGAGATACA 723

Qy        941 AATTTA 946
              ||| ||
Db        724 AATCTA 729
```

```
RESULT 72
BI258848
LOCUS       BI258848                767 bp    mRNA    linear   EST 17-JUL-2001
DEFINITION  602969663F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5109303 5',
            mRNA sequence.
ACCESSION   BI258848
VERSION     BI258848.1  GI:14815606
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 767)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11264  row: m  column: 16
            High quality sequence stop: 758.
FEATURES             Location/Qualifiers
     source          1. .767
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5109303"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_12"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
ORIGIN

  Query Match             34.5%;  Score 452.8;  DB 2;  Length 767;
  Best Local Similarity   77.0%;  Pred. No. 5.3e-112;
  Matches  591;  Conservative    0;  Mismatches  172;  Indels    5;  Gaps    3;

Qy        218 GTGAAGAGGCGGCGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACG 277
              ||||| | |||||| |||| |||  |||  ||||| | ||| |    | |||||  || |
Db          1 GTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCGCAAGCTGCGATGCCGCA 60

Qy        278 ATGGTCCCCAGCGTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTAC 337
              |||  |   || ||||||   ||||||||||||| ||| | |  |  || |||| | |||||
Db         61 GTGGCTCAGTGCTTCCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCCTAC 120

Qy        338 TTCGAGCTGCCAGAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCC 397
              |||||||  || | | | ||    | | |   ||||  |||   ||  ||  || ||
Db        121 TTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAACGCCGACCTGAAACCATCTCTGAGCCC 180

Qy        398 GAGGCCTATGTTGATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGT 457
              ||  |||||||||| |||||||| || ||  ||| |||| | |||| ||   ||  |||
Db        181 AAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAAATCAGC 240

Qy        458 CCATCTGGA---ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATT 514
              |||||||  |   ||||  | |||| || ||||||    | ||| ||||||||||||||||
Db        241 CCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATT 300

Qy        515 GATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCT 574
              ||||||||||||  |  ||||||| | ||||||||||||||||||||||||| | |||||
Db        301 GATGGATTAGATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCCTACTTAGCT 360

Qy        575 TTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTA 634
              |||||  || |||||||| |||||||||||||||||||| |||||||| | |  |||||||
Db        361 TTGTACAGCCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTA 420

Qy        635 AAGAAGAGAGCAGCCAGTTACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCT 694
              |||||||||  ||   | |||    |||||||||||| ||||||||||||||||||||||
Db        421 AAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCT 480

Qy        695 ATACTATTGAAGAAAGGAAGAGTGAAATTTAAAAGTCAGGAGATTATTCCTTTTCCAAAT 754
              ||| | |||||||||  |||||||||||| ||||| || ||||||||||||||||||| |
Db        481 ATAATGTTGAAGAAATCAAGAGTGAAATTAAAAAGCCAAGAGATTATTCCTTTTCCAAGT 540
```

```
Qy        755 ACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGC 814
              |||||||||||||||||||||  |||| ||  ||||||  | ||   ||||||||  |||||
Db        541 ACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATGAGCTTTGC 600

Qy        815 CTTATGACATCCCATTTGGAGAGCACCAGAGAACATTC-TGCGGAACGAATAAGACAATT 873
              |||||||||||||||||||||||||||||||  ||| | ||||||||||||  |  || ||
Db        601 CTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTTGCGGAACGAATGAATCAGTT 660

Qy        874 AAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGG 933
              | ||   |||  |    | ||||||||||||||||||| ||| | || |||||| |||||||
Db        661 ACAACTGGTTTTAAAGACAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATA-TTGCAGG 719

Qy        934 AGATACAAATTTAAGAGATCAAGAAGTTATCAAATGTGGTGGTTTACC 981
              |||||||| | || | |||| ||| |||  || |||||| |||| |||
Db        720 AGATACAAGTCTAGGGGATCGAGAGGTTCCCAGATGTGGGGGTTAACC 767
```

RESULT 73
BE290901

```
LOCUS       BE290901                525 bp    mRNA    linear   EST 13-JUL-2000
DEFINITION  601084205F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3498336 5',
            mRNA sequence.
ACCESSION   BE290901
VERSION     BE290901.1  GI:9172261
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 525)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8553  row: j  column: 01
            High quality sequence stop: 500.
FEATURES             Location/Qualifiers
     source          1. .525
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3498336"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam6"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
ORIGIN

  Query Match             34.5%;  Score 452.2;  DB 2;  Length 525;
  Best Local Similarity   97.5%;  Pred. No. 6.9e-112;
  Matches  512;  Conservative    0;  Mismatches    8;  Indels    5;  Gaps    5;

Qy        413 CTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCT 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCT 60

Qy        473 CTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGC 532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 CTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGC 120

Qy        533 AATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTG 592
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        121 AATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGC-TTGTATAGTCCAGATGTG 179

Qy        593 GTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGT 652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGT 239

Qy        653 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATTGAAGAAAGGA 712
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        240 TACACAATTATTACAGGTAATGAAGAAGGATATTTCACAGCTATACTATGGAAGAAAGGA 299

Qy        713 AGAGTGAAATTTAAAAGTCAGGAGATTATTCC-TTTTCCAAATACCAAAATGATGAGAAA 771
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        300 AGAGTGAAATTTAAAAGTCAGGAGATTATTCCATTTTCCAAATACCAAAATGATGAGAAA 359

Qy        772 CCTGCTATGCGTAA-ATGTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATT 830
              ||||||||||||||| |||||||| ||||||||||||||| ||||||||||||||||||
Db        360 CCTGCTATGCGTAACATGTGAGTCAGGGTGGAAATGAATTTGGCCTTATGACATCCCATT 419

Qy        831 TGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAA-AAACTGTTCTTGGA 889
              |||||||||||||||||||||||||||||||||||||||||||| |||| |||||||||
Db        420 GGGAGAGCACCAGAGAACATTCTGCGGAACGAATAAGACAATTAACAAACGGTTCTTGGA 479

Qy        890 AAA-ATGCAAGAGGCTCCAGATTCAACCACGGTTATATTTGCAGG 933
              ||| |||||||||||||||||||| |||||||||||||||| ||
Db        480 AAACATGCAAGAGGCTCCAGATTCAGCCACGGTTATATTTGCCGG 524
```

RESULT 74
CA976386
LOCUS CA976386 1046 bp mRNA linear EST 06-JAN-2003
DEFINITION AGENCOURT_8862577 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:6437549 5', mRNA sequence.
ACCESSION CA976386
VERSION CA976386.1 GI:27509040
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1046)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13954 row: i column: 06
High quality sequence stop: 348.
FEATURES Location/Qualifiers
source 1. .1046
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:6437549"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam2"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"
ORIGIN

Query Match 34.4%; Score 451.2; DB 6; Length 1046;
Best Local Similarity 94.7%; Pred. No. 1.6e-111;
Matches 521; Conservative 0; Mismatches 23; Indels 6; Gaps 5;

```
Qy        230 CGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGC 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGGATGCGACCCCACGATGGTCCCCAGC 60

Qy        290 GTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCA 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTCCTGCGGGAGAACGACTGGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCA 120

Qy        350 GAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTT 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGAACGACCAAGGGTGGCCGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTT 180

Qy        410 GATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACT 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATCTAACCAACGAGGATGCAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACT 240

Qy        470 CCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGA 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTCTAGAAGATAGCAGCACTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGA 300

Qy        530 TGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGAT 589
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        301 TGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGCCCACAT 360

Qy        590 GTGGTATTTCTACAGGAAGTTATCCCCCCATACTGTGCCTACCTAAAGAAGA-GAGCAGC 648
              |||||||||||||||||||||||||||||| |||||||||||||||| |||||| |||||||
Db        361 GTGGTATTTCTACAGGAAGTTATCCCCCCCTACTGTGCCTACCTAGAGAAGAGGAGCAGC 420

Qy        649 CAGTTACACAATTATTA-CAGGTAATGAAGAAGGATATTTCACAGCTATACTAT-TGAAG 706
              |||||||||||||||| | || |||| |||||||||||||| ||||||||||  |
Db        421 CAGTTACACAATTATTACCGGGGAATGGAGAAGGATATTTCACCGCTATACTATCCGCCC 480

Qy        707 AAAGGAAGAGTGAAATTTAAAAGTCAGG-AGATTATTCC--TTTTCCAAATACCAAAATG 763
              | |||||||| |||||||||| |||| |||||||||| ||| ||||  || | ||
Db        481 AGAGGAAGAGCGAAATTTAAAACGCAGGCAGATTATTCCGTTTTCCCAATACCCCACCTG 540

Qy        764 ATGAGAAACC 773
              ||||||||||
Db        541 ATGAGAAACC 550
```

RESULT 75
BI648868
LOCUS BI648868 691 bp mRNA linear EST 12-SEP-2001
DEFINITION 603275826F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5316055 5', mRNA sequence.
ACCESSION BI648868
VERSION BI648868.1 GI:15563104
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 691)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

```
           http://image.llnl.gov
           Plate: LLAM11801  row: d  column: 08
           High quality sequence stop: 478.
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129,C57BL/6J,FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5316055"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="10 months"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam3"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                     Reference for transgenic model: Xu et al., Nature Genetics
                     22, 37-43 (1999)."
ORIGIN

  Query Match             34.3%;  Score 449.8;  DB 3;  Length 691;
  Best Local Similarity   87.5%;  Pred. No. 3.4e-111;
  Matches  603;  Conservative    0;  Mismatches   57;  Indels   29;  Gaps    9;

Qy        309 GGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGC 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCAGACGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGC 60

Qy        369 CGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATG 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGCCAGCCTCCCACGTCCTTCAAGTCCGAGGCCTATGTTGATCTAACCAACGAGGATG 120

Qy        429 CAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCA 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAATGATACAACCATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGATAGCAGCA 180

Qy        489 CTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGG 548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTATTTCTTTCATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGG 240

Qy        549 CTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAG 608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCGAGGGGTGTGTTCCTGCCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAG 300

Qy        609 TTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAG 668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATCCCCCCATACTGTGCCTACCTAAAGAAGAGAGCAGCCAGTTACACAATTATTACAG 360

Qy        669 GTAATGAAGAAGGATATTTCACAGCTATACTATTGAA-GAAAGGAAGAGTGAAA-TTTAA 726
              |||||||||||||||||||||||||||||||||||| |||||||||||||||| |||||
Db        361 GTAATGAAGAAGGATATTTCACAGCTATACTATTGAACGAAAGGAAGAGTGAAATTTTAA 420

Qy        727 AAGTCAGGAGATTATTCCTTTT--CCAAATACCAAAATGATGAGAAACCTGCTATGCGTA 784
              |||||||||||||||||||||  | ||   | | ||||||||||||||||||||||||||
Db        421 AAGTCAGGAGATTATTCCTTTTTCCAAATACCAAGAATGATGAGAAACCTGCTATGCGTA 480

Qy        785 AAT--GTGAGTTTGGGTGGAAATGAATTTTG-CCTTATGACATCCCATTT----GGAGAG 837
              |||     |||| ||  ||||||||||||||| ||||||||| ||||||||    |||
Db        481 AATTGTGAAGTTGGGTGGGAAATGAATTTTGCCCTTATGACTTCCCATTTGGGAGGAAGT 540

Qy        838 CACCAGAGAACATTCTGCGGAACGAATAAGACAATTAAAAAC--------------TGT 882
              ||||  |||| |   | |     |    ||   |||||  |||              |
Db        541 CACCCAAGAAAACCATTCCTGGCCGGGAACGGAATTAGGAACAACTTAAAAACTTGGTCC 600

Qy        883 TCTTGGAAAAATGCAAGAGGCTCC--AGATTCAACCACGGTTATATTTGCAGGAGATACA 940
              |    ||||| ||| |||||||||     | || ||||||| |||||||| |||||| |
Db        601 TTGGCGAAAATTGCCAGAGGCTCCCAGATTCCACCCACGGTAATATTTGCGGGAGATTCC 660

Qy        941 AATTTAAG-AGATCAAGAAGTTATCAAAT 968
              |||||| | ||||||||| ||| |||| |
Db        661 AATTTAGGAAGATCAAGAGGTTTTCAATT 689
```

```
Search completed: March 11, 2006, 14:18:47
Job time : 5846 secs
```

THIS PAGE BLANK (USPTO)

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☑ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)